

- Hadley, R. F. (1994). Systematicity revisited: reply to Chater & Christiansen and Niklasson & van Gelder. *Mind and Language*, 9, 431-443.
- Karmiloff-Smith, A. (1992). *Beyond modularity*. Cambridge, Mass.: MIT Press. Traducción española, *Más allá de la modularidad*. Madrid: Alianza, 1994.
- Karmiloff-Smith, A., Scerif, G. & Ansari, D. (2003). Double Dissociations in Developmental Disorders? Theoretically Misconceived, Empirically Dubious. *Cortex*, 39, 161-163.
- Legrenzi, P., Girotto, V. & Johnson-Laird, P. N. (1993). Focussing in reasoning and decision-making. *Cognition*, 49, 37-66.
- Leslie, A. M. (1991). The theory of mind impairment in autism: Evidence for a modular mechanism of development? En A. Whithen (Ed.), *Natural theories of mind*. Oxford: Blackwell.
- Leslie, A. M. (1992). Pretense, autism and the «theory-of- mind» module. *Current directions in Psychological Science*, 1, 18-21.
- Mareschal, D. & Shultz, T. R. (1996). Generative connectionist networks and constructivist cognitive development. *Cognitive Development*, 11, 571-603.
- Orr, H. A. (2003). Darwinian Storytelling. *The New York Review of Books*. Vol L (3), 17-20.
- Pinker, S. (1994). *The Language Instinct*. New York: William Morrow. Versión española en Alianza Psicología, 1996.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human*. London: Viking.
- Quartz, S. R. (1993). Nativism, neural networks, and the plausibility of constructivism. *Cognition*, 48, 123-144.
- Smolensky, P. (1989). Connectionism and Constitutive Structure. En R. Pfeifer y otros (Eds.), *Connectionism in Perspective*. Amsterdam: Elsevier Science Publishers B. V.
- Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, 50, 431-446
- Spelke, E. S. & Newton, E. L. (1998). Nativism, Empiricism, and the Development of Knowledge. En R. M. Lerner (Ed.), *Handbook of Child Development. Vol. 1: Theoretical Models of Human Development*. New York: Wiley and Sons.

Las paradojas de la modularidad

José Manuel Igoa
Universidad Autónoma de Madrid

La Modularidad de la Mente (MM) de Jerry Fodor es, probablemente, una de las obras más célebres de la bibliografía de las ciencias cognitivas, hasta el punto de ser quizá el libro de psicología teórica más citado de las dos últimas décadas, sobre todo en psicología cognitiva. Sin embargo, ha sido una obra poco influyente, en la mayor parte de las áreas de estudio de las disciplinas cognitivas, en lo que hace al seguimiento de sus propuestas por parte de los investigadores en estas materias. Cumplidos veinte años de su publicación original, creo que difícilmente se puede discutir este balance. A la hora de buscar una explicación a estos hechos, se podrían invocar razones muy diversas, y muy probablemente sería difícil llegar a conclusiones inequívocas o conseguir un acuerdo general. Esta es la primera paradoja que quiero resaltar en este escrito. Se trata de una paradoja que bien pudiéramos llamar «externa», ya que afecta a la acogida de la obra

Correspondencia: Departamento de Psicología Básica. Facultad de Psicología. Universidad Autónoma de Madrid. Campus de Cantoblanco. Ctra. de Colmenar, km. 15. 28049 Madrid. Correo electrónico: josemanuel.igoa@uam.cs

y a las reacciones que provocó y sigue provocando entre los investigadores de las ciencias del conocimiento. A tal efecto, MM ha sido una obra tal vez más admirada que denostada, provocadora y sugerente a la vez que polémica y pesimista. Si la juzgamos por su fecundidad, es decir, por su capacidad de suscitar teorías o modelos capaces de generar investigaciones empíricas, no hay duda de que ha tenido un peso considerable en la psicología. Sin embargo, ello no impide reconocer que hay pocos «fodorianos» a carta cabal en el panorama actual de la ciencia cognitiva. Y quizá no podría ser de otro modo, teniendo en cuenta las razones que expondré a lo largo de este artículo.

En esta breve reseña conmemorativa, me propongo señalar y comentar otras dos paradojas, «internas» a la propia propuesta de Fodor de arquitectura cognitiva recogida en MM, que quizá puedan ayudar a entender esta primera paradoja externa y a explicar por qué la propuesta de Fodor ha tenido el curioso efecto de volverse contra sí misma, dando paso a diversas propuestas de arquitectura modular que poco tienen en común con las ideas originales del autor. Una primera razón, nada paradójica por lo demás, es que Fodor ha tenido la rara virtud de articular un concepto de modularidad sumamente restrictivo en lo que respecta a los criterios constitutivos de los sistemas modulares y, consiguientemente, en lo que afecta a la porción «modular» del sistema cognitivo humano, que es mínima en comparación con la porción no modular. Como es bien sabido, Fodor condena además los sistemas no modulares, responsables de la fijación de creencias no demostrativas, al limbo de los procesos científicamente inabordables. Así pues, y por decirlo en términos metafóricos, la modularidad fodoriana ha tenido la consecuencia de «quitar el suelo bajo los pies» de buena parte de la investigación cognitiva, que se afana precisamente en estudiar esos procesos de carácter central. La lógica reacción de los damnificados por esta propuesta ha sido, como no podía ser menos, la de buscar formas alternativas de arquitectura modular más benignas con la mayor parte de los procesos cognitivos sometidos a indagación. Y aquí es donde radica la primera de las dos paradojas internas que me propongo analizar y que puede formularse del siguiente modo: la arquitectura modular de Fodor ha resultado ser la menos modular de todas las arquitecturas modulares hoy en día disponibles. Por ponerlo en términos más crudos, lo que Fodor propone en MM es, en su mayor parte, una arquitectura *no modular* del sistema cognitivo humano. Con esto, Fodor logra dejar insatisfechos a todos: a los partidarios de las arquitecturas globales, que ven sus teorías arrojadas al pozo insondable de los procesos centrales, y a los partidarios de las arquitecturas genuinamente modulares, que ven a sus módulos privados de las propiedades que les otorgan carta de naturaleza. Se diría que en el justo medio en el que Fodor se sitúa *no* está la virtud.

Haciendo un ejercicio de interpretación más o menos arriesgado, podría decirse que la propuesta restrictiva de arquitectura modular de Fodor obedece a una calculada estrategia de ambigüedad, consistente en definir los módulos en virtud de dos características tan independientes entre sí como el encapsulamiento informativo y la especificidad de dominio. Fodor se sirve de la proposición de que un sistema cognitivo es modular si está informativamente encapsulado y es específico de dominio (en otras palabras: si A y $B \rightarrow C$; A y $B \rightarrow C$, A y no $B \rightarrow$

no C, no A y B \rightarrow no C) para caracterizar aquellos sistemas que, a su juicio, merecen la consideración de modulares, esto es, los sistemas perceptivos o sistemas de entrada, excluyendo todos los demás. Esta estrategia genera ambigüedad porque deja sin aclarar expresamente cuál de estas dos propiedades es realmente constitutiva de los módulos. Si tomamos la especificidad de dominio como el rasgo definitorio de los módulos, habremos optado por una arquitectura modular definida en términos de un *racionalismo epistemológico*, como el propio Fodor ha subrayado en su reciente libro *The mind doesn't work that way* (Fodor, 2000), o sea, por una arquitectura de módulos de conocimiento a la Chomsky (1980) o paquetes de información de base innata que definen dominios peculiares de actuación. Esta parece haber sido la opción preferida por la mayoría de los autores que defienden alguna versión «post-fodoriana» de arquitectura modular: desde Gardner (1983), con su teoría de las inteligencias múltiples, y Carey y Spelke (1994), con su propuesta de modularidad de los sistemas de conocimiento, de inspiración chomskyana, hasta otros autores como Karmiloff-Smith (1992), con su teoría de la modularidad emergente, e incluso Jackendoff (1987, 2002), defensor de lo que él mismo denomina la «modularidad representacional». Es de justicia señalar que son más las cuestiones que separan a estos autores que las que les unen. La propiedad de los módulos en la que radican las principales discrepancias entre estas teorías es la del innatismo y, en consecuencia, la modificabilidad de los sistemas de procesamiento modulares a partir de la experiencia. Según esta propiedad, encontramos teorías emergentistas o desarrollistas (si se permite el término), como las de Gardner o Karmiloff-Smith, frente a otras de claro sesgo innatista, como la de Carey y Spelke y la de Jackendoff.

Sin embargo, no es menos cierto que todas estas teorías, que en parte surgieron como reacción a la propuesta original de Fodor, comparten el énfasis en la especificidad de dominio como propiedad constitutiva de los módulos, en detrimento del encapsulamiento informativo. Esto es especialmente cierto de las teorías de Gardner, Carey y Spelke y Karmiloff-Smith, y de forma más matizada de la de Jackendoff. A propósito de esta última, que quizá es menos conocida entre los psicólogos cognitivos en comparación con las otras que he citado, hay que señalar que Jackendoff defiende un tipo de arquitectura modular que se articula en dos niveles de análisis de los sistemas cognitivos: uno macroscópico o molar y otro microscópico o molecular. El primero individualiza los sistemas modulares de acuerdo con un criterio basado principalmente en la especificidad de dominio, lo que arroja un censo de unos pocos módulos grandes (al estilo de las facultades fodorianas) y de límites inciertos que incluye las facultades del lenguaje, la visión y la música, entre otras. Es importante resaltar que estos sistemas no se conciben en modo alguno sistemas perceptivos (ni siquiera el sistema de procesamiento visual está pensado exclusivamente en esos términos) y tampoco son sistemas encapsulados, al menos según Fodor define la propiedad del encapsulamiento, sino que se comunican entre sí de forma sistemática o *arquitectónicamente definida* (es decir, por la propia constitución de su arquitectura) en varios puntos. Por otra parte, en un nivel microscópico o molecular, la arquitectura modular de Jackendoff está constituida por unidades de procesamiento que, a la vez que son específicas de dominio (en la medida en que cada una de ellas posee una

base de datos peculiar que le es propia), se hallan informativamente encapsuladas. Estos procesadores definidos representacionalmente, que Jackendoff denomina «módulos de integración» (las estructuras fonológica y sintáctica serían ejemplos paradigmáticos de la facultad del lenguaje y el esbozo primario y la estructura 2¹/₂ D de la facultad de la visión) se comunican unos con otros a través de otra clase de módulos que les sirven de interfaz y que tienen acceso a ciertas representaciones privilegiadas de los dominios (i.e. de los módulos de integración) que conectan. Entre las consecuencias más dignas de mención de la arquitectura modular de Jackendoff se deben destacar dos: por un lado, se trata de una arquitectura mucho más flexible que la de Fodor y globalmente mucho más «modularizada»; y por otro, y derivado de lo anterior, es una arquitectura en la que queda totalmente desdibujada la distinción que Fodor tanto subraya entre la percepción (modular) y la cognición (no modular).

Me he detenido brevemente a exponer algunas de las principales características de la teoría modular de Jackendoff para mostrar hasta qué punto una propuesta que se dice deudora y se considera próxima a la de Fodor termina siendo, a la postre, bastante diferente. En mi opinión, la clave de ello reside, en buena medida, en la primacía que esta teoría concede a la especificidad de dominio sobre el encapsulamiento informativo a la hora de caracterizar los sistemas modulares de procesamiento. Con todo, la teoría de la modularidad representacional de Jackendoff es una teoría sobre los *procesos cognitivos*, y no sólo sobre los *estados epistémicos*, y por ello se adhiere a la idea de que los procesos cognitivos son computacionales, en el sentido clásico del término, es decir, operaciones formales que se ejecutan sobre representaciones sintácticamente estructuradas (Jackendoff, 1987). En este particular, la arquitectura modular de Jackendoff se sitúa en la misma tradición de pensamiento que la de Fodor (y que éste denomina *racionalismo computacional*) y en una posición asimismo muy próxima a otra de las concepciones post-fodorianas de arquitectura mental, la denominada teoría de la «modularidad masiva», que defienden, entre otros autores, Cosmides y Tooby (1994), Sperber (1994) y Pinker (1997).

La modularidad masiva es una teoría de la arquitectura mental que ha sido caracterizada como una síntesis afortunada de la teoría computacional de la mente y la biología evolucionista neodarwiniana (de ahí el apelativo de Nueva Síntesis con que también se la conoce). La idea central de esta teoría es que la selección natural ha favorecido la emergencia de estructuras cognitivas específicas de dominios altamente restringidos dedicadas a resolver problemas muy particulares. Estas estructuras cognitivas están realizadas en mecanismos de cómputo que sobrevienen en estructuras neuronales fijas y, consecuentemente, innatas. Bajo esta óptica, la mente humana queda caracterizada, en palabras de Pinker, como «un sistema de órganos de computación, diseñados por selección natural, para resolver la clase de problemas a los que se enfrentaron nuestros antepasados cazadores-recolectores, en particular, la comprensión y manipulación de objetos, animales, plantas y otros congéneres» (1997, p. 21). El grado de especificidad de los módulos en esta teoría es tal que ha llegado a proponerse hasta un sistema de procesamiento específicamente diseñado para reconocer conductas de engaño en situaciones de intercambio social, el famoso módulo de «detección

de mentirosos». El resultado es una concepción masiva de la modularidad, esto es, una mente superpoblada de módulos sin cabida para los sistemas centrales. Lo señalado en los últimos párrafos deja bien patente la paradoja de que la teoría de Fodor es, a fin de cuentas, la menos modular de las teorías modulares. No obstante, los efectos de esta paradoja son más estéticos que sustantivos, pues a fin de cuentas poco importa saber quién ha llegado más lejos en su propuesta de modularización de la mente. Mucho más importante, en cambio, es ver cómo se conjugan los principios que sustentan las diferentes propuestas de arquitectura mental. En este caso, se trata de examinar la congruencia entre las distintas variedades de modularidad (que a efectos de la presente discusión podemos reducir a dos, una modularidad restringida, como la que propone Fodor, y una modularidad masiva, como la de la Nueva Síntesis), de un lado, y el postulado del carácter computacional de los procesos cognitivos, principio fundamental al que todas las arquitecturas modulares declaran mantenerse fieles. Y aquí es donde surge una segunda paradoja interna a la teoría modular de Fodor que, si se toma en serio, arroja un sombrío panorama sobre la posibilidad de hallar una explicación naturalista (o causal) del funcionamiento de los procesos mentales.

Esta segunda paradoja descansa en dos premisas: por una parte, la premisa de que los procesos computacionales son intrínsecamente *locales*, es decir, sólo se pueden definir sobre dominios de muy corto alcance; y por otra, la premisa de que una buena parte de nuestras operaciones mentales está dedicada a la realización de inferencias no demostrativas que, por su propia naturaleza, necesitan acceder a un amplio rango de fuentes y tipos de información (imposible de delimitar a priori, en la mayoría de los casos). La primera premisa se basa en la idea de que los procesos mentales, en tanto que computacionales, operan sobre la sintaxis de las representaciones mentales. La sintaxis, por su parte, es una propiedad local de las representaciones, en la medida en que (1) se refiere a la disposición e interdependencia de las partes constituyentes de su estructura y (2) determina, en su proyección externa, un conjunto de relaciones sistemáticas entre representaciones que comparten constituyentes. Según esto, la proposición *Juan leyó un libro de Chomsky* tiene dos clases de propiedades sintácticas: (1) unas que determinan la identidad de los constituyentes que forman la proposición y sus relaciones (*Juan-agente, un libro de Chomsky – tema, leyó-predicado*); y (2) otras que determinan las relaciones (semánticas) de la proposición con otras proposiciones lógicamente asociadas a ella (relaciones tales como la contradicción *–Juan no leyó un libro de Chomsky–* o el entrañamiento *–Juan leyó un libro*). La segunda premisa se justifica por la dependencia contextual de muchas, si no todas, las operaciones de inferencia *no demostrativa* (obviamente, las inferencias demostrativas o deductivas no poseen la característica de ser operaciones dependientes del contexto). Por dependencia contextual se entiende la necesidad de recurrir al contexto global de una determinada creencia, y en caso extremo al sistema total de creencias del individuo, para establecer su grado de confirmación. Volviendo al ejemplo anterior, las inferencias que puedan formarse a partir de la proposición *Juan leyó un libro de Chomsky*, como la de que «Juan leyó un libro de lingüística» (inferencia no demostrativa, por lo demás, ya que podría tratarse de un libro de política) o la de que «Juan se aburrió/disfrutó leyendo un libro de

Chomsky», requieren acceder a información contextual (el conocimiento u opinión que uno tenga sobre la identidad de Chomsky, el grado de interés de sus obras, etc.).

A partir de estas dos premisas, Fodor llega a la conclusión de que sólo los procesos locales, o, lo que es lo mismo, modulares, pueden caracterizarse con propiedad en términos computacionales, dado que la sintaxis de las representaciones sobre las que operan tales procesos es una propiedad local de las mismas. En otras palabras, la sintaxis de una representación no autoriza, ni tampoco basta, por sí sola para realizar inferencias de carácter global. La identificación de procesos locales con procesos modulares requiere, por su parte, una justificación adicional, a saber, el postulado de que tales procesos modulares manejan exclusivamente información «excéntrica», esto es, disponen de una base de datos restringida a la que se accede de forma automática a partir de la representación de entrada al módulo. Aquí volvemos a constatar la necesidad de la especificidad de dominio como propiedad constitutiva de los módulos. La paradoja resultante queda expresada, pues, en la necesidad de restringir las explicaciones computacionales al funcionamiento de los módulos. En el caso de los sistemas no modulares, en cambio, la sintaxis de las representaciones que manejan (y, por tanto, los procesos de cómputo) no resulta suficiente para explicar el trabajo que desempeñan. Si toda la mente humana fuera modular, entonces, repitiendo el título del libro de Pinker, podríamos saber *Cómo funciona la mente*, pero si sólo una parte de la mente humana lo es, sólo sabremos, como afirma Fodor, que probablemente *La mente no funciona de esa manera* (es decir, de manera computacional). La modularidad fodoriana da el golpe de gracia al mentalismo computacional.

La presentación de estos hechos como paradojas no deja de ser un recurso retórico. Es más que probable que para Fodor no se trate en modo alguno de paradojas (excepto quizá la primera de las que he enunciado aquí, que denominaba «externa», pero incluso eso es dudoso). Quizá sólo lo sean para quien haga una interpretación superficial y, por eso mismo, errónea de la modularidad fodoriana. Sin embargo, para terminar de entender la argumentación de Fodor, sólo queda por encajar una última pieza en la argumentación que vengo exponiendo en este artículo. Esa pieza es la respuesta a la pregunta de por qué Fodor considera que los módulos masivos (o *ipso facto* cualquier módulo encargado de llevar a cabo procesos de fijación de creencias no demostrativas) no pueden ser modulares, por muy específico que sea su dominio de operación, mientras que los sistemas de entrada sí lo son. El problema de los primeros radica en definir la entrada que pone en marcha a cada módulo, es decir, en cómo decidir qué información es propia (i.e. pertenece al dominio) de un sistema modular u otro. Este problema se presenta porque, según la concepción de Fodor, los módulos deben estar sintonizados con propiedades de la entrada sensorial o, lo que es lo mismo, deben responder a propiedades definidas en un vocabulario perceptivo. Por consiguiente, si un hipotético módulo sólo es sensible a información definida en términos abstractos o conceptuales y no guarda correlaciones sistemáticas con propiedades de algún tipo de entrada perceptiva, no habrá modo de activar los procesos que le son característicos. Esto explica por qué la noción fodoriana de módulo va estrecha-

mente ligada a la de sistema de entrada y permite entender las dificultades que entraña cualquier intento de modularizar los sistemas centrales. Sin embargo, la coherencia que exhibe Fodor en su forma de caracterizar los módulos no está exenta de problemas. Un problema que el propio autor ha reconocido recientemente (Fodor, 2000) es el de descubrir qué clase de propiedades son las que activan el módulo dedicado a la percepción del lenguaje, suponiendo que haya un conjunto de ellas que sea a la vez necesario y suficiente para tal fin. Por lo que se sabe actualmente de los procesos de percepción del habla, es prácticamente seguro que no existen en la señal propiedades invariantes necesarias y suficientes para la identificación de sonidos del habla (para un tratamiento más extenso de este asunto, véase el artículo de López Bascuas en este mismo volumen). Pero con independencia de esto, el problema reside en que si sólo reconocemos la existencia de un único sistema de entrada para el lenguaje, y su activación depende, entre otras cosas, de la presencia en la señal de entrada de cierta clase de rasgos (presumiblemente acústicos) a los que es sensible nuestro sistema perceptivo. ¿qué sucede con la percepción del lenguaje en la modalidad visual? Y aún más, ¿hasta qué punto la percepción del lenguaje escrito y del lenguaje de signos dependen de dos módulos diferenciados? ¿Hemos de pensar, entonces, que existen hasta tres módulos diferentes para la percepción del lenguaje?

El problema que acabo de enunciar no es comparable a los que Fodor ha puesto de manifiesto y ha analizado en otras arquitecturas modulares, en especial en la teoría de la modularidad masiva y en la teoría de la modularidad emergente de Karmiloff-Smith (para una crítica de esta última teoría, véase Fodor, 1999). Aun reconociendo esta situación, sería injusto condenar de antemano la propuesta de Fodor por el hecho de que no resuelva todos los problemas que suscita. A fin de cuentas, MM es más un esbozo de teoría de la arquitectura mental que una teoría acabada. Sin embargo, hay motivos para sospechar que la resolución de algunos de esos problemas puede requerir una revisión profunda de la propuesta original de Fodor, revisión que puede conducir a teorías considerablemente diferentes. Buen ejemplo de ello es la teoría de Jackendoff, con los resultados que he mencionado anteriormente en este artículo.

En los veinte años transcurridos desde su publicación, *La modularidad de la mente* ha tenido efectos paradójicos: ha sido un libro muy influyente, tanto para sus seguidores como para sus detractores; en el terreno teórico, ha producido unos frutos quizá distintos de los que se proponía, pero no por ello menos significativos o atractivos. Teniendo en cuenta la evolución metodológica (y también tecnológica) de las disciplinas de la ciencia cognitiva y la considerable acumulación de datos empíricos que esta evolución ha propiciado en este tiempo, podemos decir que MM ha envejecido considerablemente bien, habiendo resistido hasta la fecha los intentos de falsación global de que ha sido objeto y habiendo servido de espejo en el que se han mirado numerosas investigaciones en las disciplinas cognitivas, por más que este espejo haya dado, en ocasiones, una imagen en negativo. Todo esto, empero, no contradice un ápice el hecho de que la arquitectura modular de Fodor sea, a fecha de hoy, una propuesta relativamente minoritaria, excéntrica (en el sentido literal de la palabra) y poco aceptada, al menos en su estricta formulación original.