

# Funcionalismo y ciencia cognitiva, lenguaje y pensamiento, modularidad y conexionismo

ENTREVISTA CON JERRY A. FODOR

JOSÉ E. GARCÍA-ALBEA

*Universidad Complutense de Madrid*

En el centro mismo de Manhattan, muy cerca de Times Square y de la afamada avenida Broadway, enfrente de la Biblioteca Pública, dando a la bulliciosa y sorprendente calle 42, se alza un edificio de la Universidad de la ciudad de Nueva York (CUNY) dedicado a los estudios de postgrado. En la planta 16 está el Departamento de Filosofía, donde en un despacho luminoso, y casi siempre con la puerta abierta, se puede uno encontrar de cara con el profesor Jerry Fodor. Se le puede ver sentado al ordenador, contestando al teléfono, comentando con algún colaborador la marcha de un experimento o simplemente mirando a las musarañas con las piernas encima de la mesa (actitud que, a pesar de su descrédito social, habría que reivindicar como postura intelectual por excelencia); en cualquier caso, siempre disponible para una discusión científica. Al entrar en el despacho, se tiene la impresión de que allí se está gestando algo importante que pronto trascenderá a los círculos especializados.

Jerry Fodor lleva cuatro años en esta Universidad, tras los más de veinte que había permanecido en el Instituto Tecnológico de Massachusetts (MIT) como director del Laboratorio de Psicolingüística desde su fundación y como profesor compartido por el Departamento de Lingüística y Filosofía y el Departamento de Psicología. Durante todo este tiempo ha publicado numerosos artículos en revistas especializadas y varios libros, algunos de los cuales han sido ya traducidos al español (*La explicación psicológica*, Ed. Cátedra; *El lenguaje del pensamiento*, Alianza Ed.; y *La modularidad de la mente*, Ed. Morata). La obra de Fodor ha tenido una gran influencia en el desarrollo de la llamada «psicología cognitiva» o, si se prefiere un contexto más interdisciplinar e integrador, en el establecimiento de lo que se conoce ya como Ciencia Cognitiva. Posiblemente, su libro *La modularidad de la mente* haya sido uno de los más citados en este área en los últimos siete años. Otros títulos de especial interés en su bibliografía son *The psychology of language* (escrito conjuntamente con T. Bever y M. Garrett), *Representaciones y Psychosemantics*.

*Agradecimientos:* Este trabajo ha sido financiado en parte por la DGICYT, Proyecto PB87-0531. Hay que agradecer la ayuda prestada por José M. Igoa y Christopher Davis en la puesta a punto de la transcripción de las cintas magnetofónicas y, al primero de ellos, su ayuda en la revisión final de la traducción.

*Dirección del autor:* Departamento de Psicología Básica (Procesos Básicos). Facultad de Psicología, UCM. Campus de Somosaguas. 28023 Madrid.

Mi colaboración con Jerry Fodor se inició en 1977, con motivo de una Beca Posdoctoral del Comité Conjunto Hispano-Norteamericano para investigación y ampliación de estudios en el MIT. Durante dos años tuve la oportunidad de seguir muy de cerca la labor de este intelectual, crítico e independiente, que tenía la costumbre de poner «patas arriba» cualquier propuesta explicativa y, al mismo tiempo, de llegar hasta las últimas consecuencias una vez que se hubieran asumido determinados postulados teóricos. Junto a él descubrí los límites y dificultades de la psicología científica, la necesidad de fijar la relevancia de la investigación empírica, y el atractivo por los procesos básicos que caracterizan el funcionamiento de la mente humana.

Desde entonces he mantenido una relación asidua con Jerry Fodor, a través de distintas colaboraciones científicas y de la divulgación de su obra en nuestro país. Aprovechando mi último encuentro con él, en un Congreso organizado por CUNY el pasado mes de marzo, y ante su anunciada visita a España para participar en varias sesiones científicas sobre su obra el próximo mes de septiembre, le sugerí la posibilidad de realizar una entrevista. De manera informal y relajada, me apetecía que expusiera sus puntos de vista sobre la psicología y sus problemas, tratando de establecer un nexo entre los distintos temas en que su contribución ha sido más destacada: funcionalismo, estatus de la ciencia cognitiva, teoría computacional de la mente, psicología del sentido común, modularidad, conexionismo, etc. Muy amablemente, aceptó mi sugerencia y se avino a que aquella conversación de dos horas que tuvimos en su despacho pudiera quedar registrada magnetofónicamente. Asimismo, se ofreció a revisar y pulir la transcripción de la entrevista antes de proceder a su traducción y publicación. Este es el resultado.

\* \* \*

**GA:** Podemos empezar con algunos apuntes biográficos. Primero te hiciste filósofo, después te formaste como psicólogo experimental y, finalmente, te viste implicado en la lingüística. ¿Podrías tú mismo describir la transición entre estas etapas y decirnos cómo has integrado tu interés por estos tres campos?

**JF:** Bueno, en realidad no he recibido otra formación más que la de filósofo. Estudié en la Facultad de Filosofía de la Universidad de Princeton y siempre estuve interesado por la filosofía de la mente. Mi atención profesional a la psicología proviene de cuando estuve como profesor visitante en la Universidad de Illinois. Allí fui invitado a colaborar en un programa de investigación en psicología, subvencionado por una ayuda concedida al profesor Charles Osgood, que por aquel entonces estaba haciendo investigación psicolingüística. Tuve la ocasión de conversar con muchos de sus estudiantes graduados, surgieron diversos proyectos que parecían ser interesantes desde un punto de vista experimental, y fue así como me introduje realmente en el trabajo experimental. Mi auténtico primer puesto de trabajo fue en el MIT, donde por cierto se estaba haciendo mucha lingüística. Esto ocurría en los primeros momentos de la revolución chomskiana. Ahora bien, mi conocimiento de la lingüística ha sido siempre como aficionado y, realmente, nunca me he dedicado en serio a ella. Hubo un tiempo, hace como unos quince años, en que era posible publicar en este campo con un conocimiento «amateur» del mismo, y yo, de hecho, publiqué un par de artículos. Pero, básicamente, mi conocimiento de la lingüística ha sido siempre desde fuera.

GA: De todas formas, creo que, a través de tu trabajo y de tus publicaciones, has realizado contribuciones importantes a estos tres campos. Ya sé que no eres partidario de las rígidas divisiones departamentales, pero en todo caso, y perdona que insista, ¿cómo te ves más a ti mismo, como filósofo, psicólogo, lingüista, o quizá como alguien dedicado a la ciencia cognitiva?

JF: Esa es una pregunta difícil. La verdad es que trato de no pensar mucho en ello. En lo que estoy realmente interesado es en dos tipos de cuestiones. Por una parte, estoy interesado en cierta clase de problemas técnicos de la filosofía, acerca de la filosofía de la mente y de la filosofía del lenguaje; los considero como problemas básicamente filosóficos, problemas de análisis conceptual o algo por el estilo. Y por otra parte, estoy interesado en toda una serie de cuestiones acerca de cómo funciona la mente, cómo es posible el pensamiento, cómo los estados mentales pueden tener el contenido que tienen, cómo es posible el aprendizaje, cómo es posible el aprendizaje de conceptos. Creo que todas ellas son cuestiones interdisciplinarias y mi aproximación a las mismas es a veces por la vía de la filosofía y a veces por la vía de la psicología; no creo que se pueda decir que pertenezcan a un único campo. Los avances que se hacen respecto a ellas dependen, en gran medida, de la utilización de todo un conjunto de técnicas extraídas de diferentes disciplinas. Así que cuando estoy trabajando en este tipo de materias, no sé muy bien cómo verme a mí mismo. Me considero una persona interesada en la mente más que, de modo particular, como un filósofo, un psicólogo o un lingüista. De lo que estoy bastante seguro es de que no soy un lingüista.

GA: Hablando ahora más en general de la ciencia cognitiva, yo te preguntaría qué perspectivas ves para el asentamiento de una auténtica ciencia de la cognición, cómo ves el futuro inmediato de la ciencia cognitiva.

JF: Bien, yo creo que la ciencia cognitiva se define, por así decirlo, por áreas de problemas. Hay un cierto tipo de problemas en lo que suelen estar interesados aquellos que se ocupan de la ciencia cognitiva, los problemas acerca del pensamiento y la representación mental. Asimismo, existe un amplio acuerdo —aunque de ningún modo universal— sobre las clases de modelos teóricos con que se puede contar para abordar dichos problemas. En general, son los modelos que provienen de la teoría de la computación, y que en último término están basados en las aportaciones de Turing, von Neumann y otros autores afines. En cierto sentido, la ciencia cognitiva comporta una apuesta. Apostar a favor de que los problemas centrales sobre la mente pueden ser resueltos —o, al menos, que se puede avanzar algo en su solución— mediante la utilización de modelos basados en las nociones de representación mental y proceso computacional. El que esta disciplina tenga éxito va a depender del resultado de la apuesta. Si esos modelos son productivos y esclarecedores, se avanzará en este campo. Si no lo son, entonces los problemas permanecerán y, en cambio, será poco probable que permanezcan las características concretas, la galaxia particular de talentos, técnicas e ideas que han definido la ciencia cognitiva y que están tan arraigadas en la noción de computación; a no ser, repito, que la visión computacional de la mente resulte provechosa.

Creo que la ciencia cognitiva no es tanto un campo de investigación como una estrategia de investigación. Y la cuestión a la que todavía no sabemos responder, y a la que probablemente seguiremos sin responder en otros cincuenta años, se refiere precisamente a lo provechosa que pueda ser esa

estrategia. Lo que está bastante claro es que el proseguirla requiere todo un complejo de sofisticación y de habilidades que atraviesa varias disciplinas tradicionales. De este modo, lo que sí creo que vamos a presenciar es un desarrollo de todo ese complejo de talentos, habilidades, conocimientos, etc., puestos al servicio de los problemas mencionados. Lo que no sabemos es si dichos problemas se prestarán a este tipo de tratamiento, y quizá tardemos todavía bastante en saberlo con certeza.

**GA:** Una de tus principales aportaciones a la psicología consiste en haber contribuido a sentar las bases de un nivel explicativo independiente para esta ciencia, en contra de las distintas formas de reduccionismo que la han amenazado con frecuencia. A este respecto, la noción clave es la de funcionalismo. Ahora bien, en la historia reciente de la psicología el funcionalismo ha sido objeto de toda una variada gama de interpretaciones. De forma más o menos oportuna, muchos psicólogos se han considerado a sí mismos funcionalistas, desde William James hasta los conductistas, e incluso autores tan dispares como Piaget o Gibson habrían gustado de tenerse por funcionalistas. No parece, pues, que todos ellos entiendan el funcionalismo de la misma manera. ¿Cuáles serían entonces las características específicas del funcionalismo que tú has mantenido, ese funcionalismo que nos permite estudiar la mente y que difiere de las otras posibles acepciones?

**JF:** Lo primero de todo es que nos encontramos aquí con una ambigüedad manifiesta y tediosa. Muy a menudo, lo que se ha entendido por funcionalismo, en disciplinas como la antropología o en la misma psicología, se basa en la idea de que todo aquello que se esté estudiando debería ser estudiado por referencia a la función que cumple. Así, por ejemplo, si lo que se estudia es la mitología popular, entonces la forma de comprenderla es recurriendo a su función en la cultura. El que esto sea o no una buena idea es algo que se me escapa. Pero eso no es lo que se ha entendido por funcionalismo cuando se asocia a éste con nociones tales como ciencia cognitiva, representaciones mentales y cosas por el estilo. En este contexto, hablar de funcionalismo es tanto como decir que hay un nivel de representación o nivel explicativo que, por una parte, está directamente implicado en la explicación de la conducta y demás fenómenos en los que tradicionalmente han estado interesados los psicólogos, y que por otra parte, es un nivel no-neurológico. Así pues, en este nivel de representación o nivel de explicación, sistemas biológicos de diferentes tipos podrían ser objeto de las mismas leyes y, por ello mismo, los sistemas biológicos podrían ser objeto de las mismas leyes que, por ejemplo, sistemas no biológicos, como los ordenadores inteligentes hechos con silicio. Por lo tanto, el sentido en el que el funcionalismo ha jugado un papel en el tipo de psicología que a mí me interesa es el de que puede darse un nivel de representación o de construcción teórica que se aplica a los organismos biológicos pero que no está especificado en términos biológicos. Es la idea de que los estados psicológicos pueden ser caracterizados de una forma que sea relevante para la explicación de la conducta, pero que, al mismo tiempo, no se vea confinada al vocabulario de la biología. ¿Cómo lograr esto? Pues individualizando dichos estados por referencia a sus causas y efectos, en vez de individualizarlos por referencia a aquello de lo que están hechos. Desde este punto de vista, aquellos procesos que produzcan los mismos efectos podrán ser considerados como idénticos, aun cuando se lleven a cabo en sistemas biológicos distintos o incluso en sistemas computacionales no biológicos. Ahora bien, no está claro lo extendido que pueda estar

este punto de vista entre muchos psicólogos. De hecho, y para el caso es lo mismo, muchos lingüistas, empezando por Chomsky, sostienen que la psicología es básicamente una ciencia biológica y que, cuando las aguas vuelvan a su cauce y sepamos realmente de qué va la cosa, el vocabulario de la psicología pasará a ser algo así como un cierto tipo de vocabulario neurológico o, si se quiere, como una subespecie del vocabulario de la biología. Sin embargo, si uno es funcionalista, no piensa así. Si eres funcionalista, pensarás que la psicología es, más bien, una subespecie de la teoría de la computación, de la teoría de los procesos computacionales, y sólo accidentalmente una subespecie de la biología.

Tomemos un ejemplo de un área menos tendenciosa. La aerodinámica se aplica, hablando en general, tanto a las alas de las aves como a las de los aviones, y, para el caso, se aplica también a los veleros y a ciertos tipos de quillas de barco. Abarca una serie de estructuras físicamente diferentes, en cuanto que obedecen el mismo conjunto de leyes. Me refiero a las leyes que te dicen por qué las velas tiran de los barcos, o por qué vuelan las aves o los aviones, las mismas leyes para todos los casos. Así pues, desde el punto de vista del experto en aerodinámica, están pasando las mismas cosas cuando un velero navega con el viento que cuando despega un avión del suelo. Por supuesto que las estructuras físicas implicadas son muy distintas. Y por supuesto, también, que existen ciencias que se ocupan de cada uno de estos tipos de estructuras. Lo único que ocurre es que el experto en aerodinámica asume que hay un nivel de explicación en el que objetos físicamente muy diferentes, como las velas de barco, las alas de avión o las alas de los pájaros, pueden caracterizarse de modo más abstracto como objeto de las mismas leyes. En este sentido, se podría decir que la aerodinámica es funcionalista. No plantea la teoría de las alas de avión en términos correspondientes al metal, o la teoría de las alas de ave en el vocabulario de las plumas, sino que se acoge a un nivel de abstracción en el que tanto unas como otras se caracterizan en términos de fuerzas elevadoras, movimiento a través del medio aéreo, etc. En esto consistiría un enfoque funcionalista de la aerodinámica. De modo análogo, lo que se plantea en el caso de la psicología es un enfoque funcional con respecto a la causación de la conducta, abstrayéndose del carácter biológico del organismo en cuestión o, si se quiere, de todo aquello que no sea su carácter computacional.

**GA:** Ya que te has referido a la postura de Chomsky respecto a este tema, me gustaría profundizar un poco más en vuestras discrepancias. Es cierto que Chomsky ha insistido bastante en la naturaleza biológica del lenguaje, en que éste puede ser abordado como un sistema biológico. Pero él siempre señala que se trataría de una especie de biología abstracta, o genética abstracta, algo que no tiene mucho que ver con la forma habitual de entender la biología.

**JF:** No quiero especular demasiado acerca de cuál es la postura de Chomsky, pero creo en efecto que él mismo no está muy seguro de qué línea tomar con respecto a este punto. Cuando habla de la mente como de un órgano, las imágenes y metáforas que usa son todas ellas biológicas. Llega a decir que es como estudiar el hígado, por ejemplo. Ahora bien, si estudias el hígado es muy probable que no pretendas adoptar el enfoque funcionalista. Lo que de verdad quieres saber es su historia biológica, cuáles son sus relaciones químicas con otras partes de la anatomía, cuál es su papel fisiológico. Es bastante razonable suponer que la teoría del hígado es realmente, en un sentido

estricto, parte de la biología. Si quieres saber cómo funciona el hígado, pregúntale a un biólogo. Asimismo, tiene que haber una ciencia del cerebro que es también biológica en cuanto que está interesada por su fisiología, por el modo de funcionar los transmisores neurales, por lo que ocurre en las sinapsis, etc. La cuestión está en si la psicología pertenece o no a ese rango de ciencias.

Podría ser que, en cierto sentido, Descartes tuviera razón. Chomsky dice que es cartesiano, pero creo que un cartesiano algo menos radical de lo que yo estoy inclinado a ser. Una manera de expresar el punto de vista cartesiano podría ser el siguiente: «Desde una perspectiva científica, no deja de ser un accidente el que seamos sistemas biológicos». Es cierto, desde luego, que nuestra psicología está determinada por nuestra neurología y que, a la larga, las funciones de las que hablamos son funciones del cerebro. Sin embargo, si piensas como un cartesiano actual no-dualista, lo que tendrías que decir a esto sería: «Eso es verdad pero, desde el punto de vista del nivel natural de explicación, es una especie de accidente el que los sistemas psicológicos resulten estar encarnados en sistemas biológicos». De hecho, la teoría biológica no te informa mucho acerca de lo que pasa; la que te informa es la teoría de las relaciones funcionales. No sé muy bien cómo vería esto Chomsky. Pienso que él cree que la psicología terminará realmente por ser biología de una manera que yo dudo enormemente que lo sea. Depende bastante de en qué pasajes de la obra de Chomsky te quieras fijar. Pero, ciertamente, él no está tan convencido como yo de que las afinidades naturales de las teorías psicológicas y lingüísticas lo sean con respecto a la lógica y la ciencia de la computación. Creo que para Chomsky las afinidades son mayores con respecto a la biología. En este sentido, él estaría más próximo a Piaget que al punto de vista del científico cognitivo. Ambos tienen la idea de que la historia final habrá de ser contada en un lenguaje neurológico más que en un lenguaje computacional. O eso es, al menos, lo que yo pienso que piensa Chomsky.

**GA:** El sentido que tú adoptas del funcionalismo te ha llevado a considerar a la psicología como una ciencia especial. A lo largo de varias publicaciones has prestado una gran atención al estatus de las ciencias especiales. ¿Te importaría ahora explicar el significado del término «especial» y por qué la psicología es una ciencia especial?

**JF:** En esos artículos a los que te refieres, «especial» simplemente significa «no-básica». Parto de que, en general, las ciencias están organizadas en una cierta jerarquía y de que, en la base de la jerarquía, la única ciencia que pretende aplicarse a todas las cosas es la física. De ahí se desprende una concepción metafísica en la que todo tiene una descripción física y todo cae bajo las leyes de la física en virtud de su carácter físico. Pero también nos encontramos con propiedades de niveles superiores. Existen, por así decirlo, leyes más restrictivas como las de la meteorología, la geología o la psicología, bajo las que caen no todos, sino algunos de los objetos físicos. Estas son las leyes que vienen enunciadas por lo que yo he llamado ciencias especiales. Así que «especial» sólo significa que no es básica.

Ahora bien, desde el punto de vista de la filosofía de la ciencia, creo que se dan diferencias importantes entre las leyes de la ciencia física básica y las leyes de las ciencias especiales. Por ejemplo, las leyes de una ciencia especial siempre tienen excepciones. Mientras que las leyes de la física sólo podrían tener excepciones cuando fueran no-deterministas, es decir, cuando los fenómenos

ocurrieran por azar. Cuando tratas, sin embargo, de establecer una ley psicológica —como pudiera serlo una ley de asociación tan simple como ésta: «si uno piensa en la sal, pensará en la pimienta»—, siempre te encuentras con excepciones. Lo cual incluye excepciones como la de que uno no se duerma, o no se distraiga, o no se muera, o cualquiera otra cosa que salga mal. Todas estas excepciones no tienen interés desde un punto de vista psicológico, pero, sin embargo, no indican que las leyes que tratas de formular tienen propiedades distintas de las que corresponden a las leyes de la ciencia física básica, que tienden a ser consideradas como sin excepciones. Así pues, el lugar que se ocupa en esta jerarquía de las ciencias tiene consecuencias metodológicas bastante claras. Además es plausible pensar que haya propiedades que comparten las leyes de la psicología con las de la geología, las de la economía o las de la meteorología, y que no son compartidas por las leyes de la ciencia básica simplemente porque, a diferencia de estas últimas, aquéllas no se aplican a todas las cosas.

**GA:** Dada esa diferencia, cabría preguntarse hasta qué punto las ciencias especiales, y en concreto la psicología, pueden hacer usos del concepto de causalidad en un sentido estricto.

**JF:** Depende mucho del concepto que se tenga de causalidad. Si se adopta un punto de vista kantiano acerca de la causalidad, cosa a la que yo me siento dispuesto, la causalidad es toda sucesión de eventos gobernada por una ley. Esta noción de causalidad podría aplicarse igualmente bien tanto a la ciencia básica como a la psicología. Aunque conviene observar que las leyes serán diferentes y no sólo en el sentido trivial de que, en el caso de la psicología, las leyes serán psicológicas y, en el de la física, serán físicas. Sino que también serán diferentes en sentidos más interesantes como, por ejemplo, que en el caso de las leyes físicas, éstas no tendrán excepciones, y en el de las leyes psicológicas, sí las tendrán. El que se aplique o no la misma noción de causalidad es una cuestión importante, y si yo creo que sí se aplica por igual, es precisamente porque creo que la causación no es más que sucesión sometida a leyes. Pero ten en cuenta una vez más que, al ser distintas en aspectos relevantes las leyes de las ciencias especiales y las de las ciencias básicas, la causación física y la causación psicológica también van a serlo en aquellos aspectos que se deriven de las diferencias entre sus correspondientes tipos de leyes.

**GA:** ¿Crees que esta distinción entre ciencias básicas y especiales se puede parangonar, de algún modo, con la distinción tradicional (al menos, desde Max Weber) entre ciencias naturales y sociales, donde la psicología pertenecería a estas últimas?

**JF:** De nuevo, me parece ésta una pregunta muy acertada. Yo creo que la respuesta es que no. Pero si hablas con distintas personas, obtendrás otras tantas opiniones distintas sobre esta cuestión. La forma habitual de entender la diferencia entre las ciencias sociales y las ciencias físicas es que las primeras, incluyendo la psicología, versan sobre lo que los filósofos llaman estados intencionales, es decir, creencias, deseos y cosas similares; mientras que las ciencias físicas se supone que versan sobre estados no-intencionales, estados como el de ser una piedra o ser un protón, habiendo formas técnicas de caracterizar esta diferencia. De forma intuitiva, los estados de los que se habla en psicología son, en cierto sentido, informativos, semánticos, o algo de este tipo; es decir, tienen que ver con el contenido; mientras que no sucede lo

mismo con los estados de los que se habla en física. Pues bien, creo que esto es correcto, enteramente correcto. Pero la cuestión está en cuáles son las implicaciones de esta diferencia con respecto a la explicación que se pretende en cada caso. Y esta cuestión depende mucho de si se acepta que existen leyes intencionales o no. Supón que crees que existen leyes intencionales (esto es, por ejemplo, que hay leyes que relacionan las creencias con los deseos, o en el caso de la percepción, leyes que relacionan la estimulación con lo que uno finalmente ve u oye). Si crees que existen leyes intencionales, del mismo modo que existen leyes físicas, o como hay leyes meteorológicas o geológicas, entonces el carácter explicativo sería el mismo en las ciencias intencionales que en las no-intencionales, aunque, desde luego, las leyes a las que se apela en estas explicaciones serían diferentes. En un caso se trataría de leyes intencionales y en el otro, no. Por otra parte, los que piensan —y son muchos— que la ciencia de lo intencional tiene algo de particular, que la explicación intencional ha de ser diferente de la explicación física, tienden también generalmente a pensar que no puede haber leyes intencionales. Y es precisamente por esto por lo que el modelo de someter los hechos a leyes, que es el que se viene aplicando en las ciencias físicas, no se puede aplicar en las ciencias de lo intencional. Así pues, el auténtico debate creo que está en si hay o no leyes intencionales, y es un debate bastante complejo. Yo apuesto porque la respuesta va a ser que sí. De hecho, hay muchos ejemplos de candidatos razonables para esta clase de leyes. La respuesta va a ser que sí, y en tal caso, el tipo de explicación que se dé en las ciencias físicas y en las intencionales podría ser exactamente el mismo, es decir que, en ambos casos, la explicación va a consistir principalmente en mostrar que los eventos son instancias de leyes.

**GA:** Con respecto a otras de las así llamadas «ciencias sociales» (economía, sociología, etc.), ¿se las puede considerar como ciencias en toda la acepción del término? Por otra parte, ¿crees que en ellas se ha producido un desarrollo comparable al de la psicología, o que sería predecible para un futuro próximo?

**JF:** Una de las maneras de clasificar las ciencias es en intencionales y no intencionales. Otra manera de clasificarlas es en sociales e individuales. Quiero decir que, de entre las ciencias intencionales, hay unas que se ocupan principalmente de las interacciones sociales (la economía es un claro ejemplo) y hay otras que se ocupan sobre todo de los estados y los comportamientos individuales. Creo que la psicología es mayormente individualística, ya que se ocupa de procesos y relaciones causales que están instanciados en individuos más que en grupos de individuos. Admito, por supuesto, que, a costa de un considerable nivel de abstracción, se puede tener una ciencia social auténtica y encomiable. Una vez más, la economía es el mejor ejemplo. En su caso, el precio que hay que pagar por que sea una ciencia social es el de que tenga que ser, asimismo, enormemente abstracta, en cuanto que su objeto viene a ser una sociedad de agentes puramente racionales. El que se pueda conseguir una ciencia que sea social y, al mismo tiempo, no tan abstracta —como podría serlo la psicología social— es una cuestión completamente abierta. Mi sospecha es que no se va a conseguir, aun cuando reconozco que tal sospecha puede ser sólo un prejuicio. Hasta el momento, me parece que los avances principales de la psicología se han realizado en la psicología individual; quizá mañana alguien pueda tener una idea maravillosa de cómo hacer realmente una psicología social sería. Por ahora no parece que sea éste un campo en el que se



esté yendo muy lejos o que se mueva a gran velocidad. Lo que indica mi sospecha es que ello no es por casualidad. Sospecho que la unidad natural de explicación en psicología es el individuo. Pero, como vengo diciendo, no es más que una sospecha. Algo que está bastante claro es que no se puede sentenciar a priori qué es lo que va a constituir una ciencia. Qué cosas puedan ser estudiadas científicamente es una cuestión enteramente empírica. En este momento, parece que lo mejor es apostar por el estudio de una gama bastante limitada de fenómenos individuales, como la percepción, el aprendizaje o la memoria. No veo una forma clara de ampliar los resultados de este tipo de estudios a los fenómenos sociales.

**GA:** Sin embargo, es curioso comprobar que, en el campo de la psicología social, se han desarrollado teorías cognitivas muy desde el principio, incluso coincidiendo con el auge del conductismo. La psicología social vio pronto la necesidad de introducir nociones cognitivas para dar cuenta de las relaciones interpersonales. De esta manera, si se parte de sistemas individuales que interactúan y se dispone de una cierta gramática que permita describir esas interacciones, parecería posible hacer de ello una ciencia.

**JF:** Es posible y, de hecho, existen áreas en las que se ve que eso está dando resultado. Spongo que la teoría de la comunicación, a diferencia de la teoría gramatical, es parte de la psicología social. Es, desde luego, una parte muy restringida de la psicología social, pero hay algunas otras teorías en áreas como la pragmática o similares. Teorías de cómo, hablando en general, la gente se las arregla para inferir las intenciones comunicativas del otro a partir del significado literal de lo que éste dice. Por lo tanto, existe al menos un caso en que se da realmente una proyección seria de la teoría de la gramática o del procesamiento gramatical, propia de la psicología individual, al ámbito de la psicología social. Lo que se sigue de ese caso es que la psicología individual ha de estar bastante bien elaborada antes de que se pueda ascender un escalón para hacer psicología social. A mí me parece que la parte de la psicología social que quizá se comprende mejor que ninguna otra, o que quizá disponga de teorías más elaboradas, está constituida por la teoría de la comunicación y la teoría de la decisión; y en ambos casos ya disponíamos de formalismos potentes para describir la conducta individual. No debe ser una casualidad que aquellos trozos y partes en los que se ha dado una psicología social rigurosa sean también los trozos y partes en que la psicología individual ha trabajado de forma más minuciosa.

**GA:** Echemos un último vistazo al carácter científico de la psicología. En tu libro *Psychosemantics*, te ocupas de temas relativos a la psicología del sentido común (también llamada «folk psychology»), compartiendo sus principales postulados. Sin embargo, tratas de ir más allá al introducir la hipótesis del lenguaje del pensamiento. Con ello sugieres que dicha hipótesis constituiría uno de los aspectos principales en que la psicología científica se separaría de la psicología del sentido común. ¿Es esto así?

**JF:** En realidad, ha habido muchas separaciones. Entiendo por psicología del sentido común lo mismo que antes llamaba psicología intencional, es decir, la idea de que la conducta ha de ser explicada por referencia a los deseos y creencias. Pues bien, lo primero que debo decir es que, tal y como están las cosas, no me parece que haya serias alternativas a la idea de que la explicación de la conducta se efectúe en términos de creencias, deseos y demás estados semánticos e intencionales. Sin embargo, creo que ya es notorio que existen

varias áreas en las que la psicología se aparta de los postulados del sentido común. La más espectacular es la que se basa en la propuesta de Freud de que casi todos estos estados resultan ser inconscientes. Eso mismo es lo que pasa ahora. Si observas lo que hacen los psicólogos dedicados a la psicología cognitiva, puedes ver que dan por hecho que prácticamente todo aquello de lo que hablan es intencional y también inconsciente, lo cual no deja de ser una palpable desviación de las nociones de creencia y deseo del sentido común.

Con respecto a lo del lenguaje del pensamiento o, si quieres, a la historia computacional, no se trata tanto de que se aparte del sentido común como de que contribuya a explicarlo, a darle una fundamentación teórica. La idea del sentido común es que nuestros deseos y creencias causan nuestra conducta, y también que unas creencias causan otras. Cuando se piensa, se piensa en tal y tal, y eso hace que pienses en esto o aquello, lo cual a su vez hace que pienses en lo siguiente, y así, por un proceso encadenado de pensamientos, se llega a solucionar problemas y a plantear lo que hay que hacer. Ahora bien, lo que no nos proporciona el sentido común —de hecho, ni siquiera está interesado en ello— es el mecanismo de estas relaciones causales entre estados intencionales. Un mecanismo que explique, por ejemplo, cómo unas creencias son causadas por otras, cómo los deseos interactúan con las creencias para causar las acciones. El sentido común asume que se dan algo así como generalizaciones intencionales, a veces realmente simplonas, como, por ejemplo, que si uno tiene sed, trata de conseguir algo de beber; asume, pues, que se dan leyes intencionales tan triviales como ésa. Pero el sentido común no tiene la más mínima idea sobre cuáles son los mecanismos que permiten que las creencias y deseos interactúen para producir acciones. Cuando se trata de ver cómo funciona esto con cierto detalle, la cuestión resulta bastante espinosa y compleja. Fijémonos, por ejemplo, en este fragmento de psicología del sentido común, dentro de toda la variedad que existe. Juan tiene sed, así que va a la pila, abre el grifo, coge un vaso, lo llena de agua y se lo bebe. Toda va bien hasta que empiezas a preguntarte en qué consiste, por así decirlo, la microestructura de todo esto. Después de todo, ¿por qué va Juan a la pila cuando tiene sed? Bueno, pues porque tiene ciertas creencias. ¿Qué creencias? Por ejemplo, que si hace tal cosa entonces saldrá agua en la pila. Hay que apelar a un conjunto muy complejo de creencias para justificar por qué se va a la pila para conseguir algo de beber y no se va a la bodega o a otro sitio parecido. De nuevo, habría que ver las creencias que le hacen ir a buscar agua cuando está sediento, en vez de hacerle buscar algo que él no cree que le quitará la sed, y así sucesivamente. Así pues, cuando tratas de bajar a los detalles de por qué Juan se comporta como lo hace cuando desea algo, la cosa se vuelve bastante complicada. Pero supongamos que se dan todos estos deseos y creencias, finamente estructurados, que entran en interacción y que son como uña y carne respecto a la causación mental de la conducta. Todavía queda por preguntarse cómo será el mecanismo que lleve a cabo eso, ya que, después de todo, tendrá que haber algún tipo de estructura causal que sustente dicha interacción entre creencias y deseos. Pues bien, lo que hace la historia computacional, lo que intenta hacer la historia del lenguaje del pensamiento, es proveernos, al nivel funcional del que hablábamos antes, de un mecanismo caracterizado en términos de procesos mecánicos definidos sobre representaciones. Un mecanismo efectivo en que se sustenten esos deseos y creencias

que interactúan y que, de acuerdo con el sentido común, determinan la forma en que uno se comporta.

He aquí una manera de ver esto. El sentido común dice que la conducta es típicamente racional, ajustándose a creencias y deseos. La psicología del sentido común piensa realmente que, para entender lo que hace alguien, es preciso saber cuáles son sus deseos y creencias. Así su conducta será inteligible, es decir, será racional a la luz de estos mismos deseos y creencias. Esto parece enteramente plausible, aunque suscita una especie de cuestión metafísica. Si se admite que somos mecanismos, es decir, que, en último término, somos sistemas causales de uno u otro tipo, entonces no hay más remedio que enfrentarse con la pregunta de cómo un mecanismo puede tener creencias y deseos, y cómo éstos pueden interactuar para producir su conducta de la manera en que el sentido común piensa que nuestros deseos y creencias interactúan causalmente para producir nuestra conducta. La historia computacional es la que da solución, aunque sea parcial, a la cuestión de cómo es el mecanismo de las interacciones causales de deseos y creencias a las que recurre el sentido común para dar cuenta de la conducta.

**GA:** La hipótesis del lenguaje del pensamiento parece, a primera vista, como una formulación muy genérica sobre la naturaleza del funcionamiento mental. ¿Has considerado la posibilidad de ir algo más allá, con el fin de determinar la gramática del pensamiento? ¿Cómo se podría dar con esa gramática? ¿Es algo abierto a la investigación empírica, a un desarrollo en el futuro?

**JF:** Sí, creo que sí. Creo realmente que es una cuestión empírica y, de hecho, constituye un tema central de la ciencia cognitiva, al menos tal como ésta se entiende en la actualidad. Lo que resulta bastante sorprendente, sin que por ello deje de ser cierto, es que no sólo se puede hacer trabajo empírico sobre el tema de cómo son las representaciones mentales —suponiendo que el pensamiento consista en relaciones causales entre representaciones mentales—, sino que existe además una cantidad enorme de fuentes de datos que pueden constreñir la solución a ese tipo de problemas. He aquí, por ejemplo, una fuente de datos. Supongamos que se piensa en un lenguaje en el que se representan los contenidos del pensamiento, consistiendo éste en procesos computacionales definidos sobre esas representaciones. Y supongamos que te planteas cómo podría ser la sintaxis de ese lenguaje. Pues bien, tendría que ser lo bastante rica como para expresar los contenidos del pensamiento. Lo cual no es, desde luego, un asunto trivial, formalmente hablando. Por referirnos a un caso muy notorio, si se quieren expresar las relaciones entre cuantificadores y variables, se requiere un lenguaje bastante complejo. Podemos pensar cosas como «Algunas chicas quieren a todos los chicos que están en la sala». La representación mental que expresa el contenido de ese pensamiento ha de tener algún significado, tiene que expresar la relación apropiada que pueda darse entre las chicas y los chicos. Pues bien, si disponemos de un lenguaje que pueda expresar ese tipo de pensamientos con cuantificadores, a dicho lenguaje le corresponderá una sintaxis igualmente rica. Al admitir que las representaciones mentales son los vehículos en los que se expresan los pensamientos, se establece inmediatamente un determinado límite inferior de cuál deba ser la riqueza sintáctica y semántica de ese lenguaje. Los hechos empíricos acerca de qué clases de pensamientos podemos pensar tienen implicaciones con respecto a cómo deben ser las representaciones mentales. Esto podría ser ya una fuente

de datos empíricos. Y he aquí otra: hay unos pensamientos que son más complejos que otros. Presumiblemente, la complejidad de los pensamientos y la complejidad de los procesos mentales en que éstos intervienen son un reflejo de la complejidad de los vehículos lingüísticos en que dichos pensamientos se expresan. Con lo cual tenemos una forma de comprobar nuestra teoría del lenguaje del pensamiento mediante datos acerca de las clases de pensamientos que son difíciles de concebir, los tipos de argumentos que son difíciles de seguir, los conceptos que son difíciles de aprender, etcétera. Así que, al menos en principio, existe una gran variedad de consideraciones procedentes de la lógica y la semántica, de la lingüística, la psicología, la teoría de la computación, que podrían servir para esclarecer las clases de teorías de la representación mental que estamos dispuestos a aceptar. De forma bastante sorprendente, una noción de apariencia filosófica, como la de que se piensa en una especie de lenguaje, lleva a consecuencias empíricas inmediatas de amplio espectro. En realidad, no es que no podamos imaginar alguna manera de constreñir las teorías sobre cómo es ese lenguaje; sino que podemos imaginar tantos tipos de constricciones diferentes que el problema está más bien en lo contrario. Los hechos que sabemos que ese sistema debe satisfacer son más que los sistemas que podemos concebir con tales propiedades.

Déjame ponerte otro ejemplo, ya que todo esto puede sonar demasiado abstracto. Supongamos que alguien pensara muy en serio que existen las representaciones mentales y que éstas no son más que imágenes. De forma inmediata, ya podemos ver que esto no puede ser verdad, y la razón principal es empírica en un sentido amplio. De entre todas las cosas en que podemos pensar, o de entre los pensamientos hipotéticos, pensamientos cuantificados, y otros por el estilo. Ahora bien, considera cómo sería una imagen de «Juan ha decidido no ponerse hoy el sombrero». Quizá podrías mostrar a Juan sin sombrero, pero ¿sería esa la imagen de que ha decidido no ponerse el sombrero, o sería la de que él no posee un sombrero, o la de que se le ha caído el sombrero? Todas esas distinciones que podemos efectuar en nuestro pensamiento no se corresponden, desde luego, con las que se podrían hacer si pensáramos solamente en imágenes. Así pues, de modo inmediato, contamos con una forma de argumento que diría algo así: no se puede admitir hipótesis alguna acerca de la representación mental a menos que las representaciones que se postulen sean lo suficientemente ricas como para poder expresar los pensamientos que de hecho podemos tener.

GA: ¿Se podría aplicar al estudio del pensamiento la famosa distinción chomskiana entre competencia y actuación? Pregunto esto porque algunos consideran que tu postura es una forma de reduccionismo lógico, reduciendo la psicología a lógica. ¿Qué tienes que decir de esta objeción? ¿Crees que la distinción entre competencia y actuación es relevante para dilucidar este asunto?

JF: Lo que yo pienso es que uno de los momentos principales en el desarrollo de una ciencia es aquel en que se decide qué hay que considerar como efectos interactivos y qué considerar como las variables de estudio que en realidad interesan. La distinción competencia-actuación no es sino una manera de expresar lo que cualquiera que hace ciencia reconoce de inmediato. No se trata de algo peculiar que se haya inventado Chomsky para importunar a los psicólogos. La distinción competencia-actuación viene a indicar que una gran parte de la varianza que puedes observar cuando atiendes a la conducta

de las personas ha de ser considerada como un efecto interactivo. Por ejemplo, la gente se muere. Así que, después de un cierto período de tiempo, las personas dejan de comportarse. Este hecho no pertenece a la psicología, sino que tiene que ver con la interacción que se da entre los sistemas psicológicos y hechos como el de que una persona puede recibir un ladrillazo en la cabeza, o puede contraer una enfermedad fatal, o cualquier otra cosa. De acuerdo con aquello en lo que está interesada la psicología, ésta no debe predecir los efectos conductuales que puedan derivarse del hecho de que uno se muera; y no es sólo que no pueda, sino que no debe. Hay ejemplos menos burdos que éste. Algunos hechos relacionados con nuestra actividad intelectual tienen claramente que ver con interacciones entre sistemas psicológicos, por ejemplo, con aquello que sabemos —por usar un ejemplo típicamente chomskiano— y otros hechos tienen que ver con la interacción entre lo que sabemos y nuestra capacidad de atención, los límites de nuestra memoria —y no sólo con los límites, sino también con la manera en que nuestra memoria está organizada—, con nuestros intereses y todo lo demás. Ahora bien, lo que se pretende hacer en cualquier ciencia —no sólo en psicología— es descubrir cuáles son los efectos interactivos, aunque sólo sea para asegurarnos de que dicha ciencia los trata como interacciones. Supón que te digo: «Mira, te doy un duro si cuentas en voz alta por números pares». Y tú empiezas con 2, 4, 6, 8, 10, 12, 14, 16, hasta que en un momento dado te paras. Habría dos hipótesis diferentes. Una es que paraste porque se te acabaron los números pares. Aunque te hubiera ofrecido más dinero, o tuvieras más memoria, realmente no sabes el número que va después de 132. Esa sería una posible hipótesis. La otra hipótesis sería que te aburraste de contar. Si te hubiera pagado suficiente, o no te hubieras cansado, o no te hubieras dormido, podrías perfectamente haber sabido que el número par que va después de 132 es 134; lo que pasó es que dejó de interesarte la tarea. Así pues, estamos ante dos hipótesis distintas sobre aquello que explica tu forma de comportamiento. En unos casos, una de estas clases de hipótesis es la correcta, y en otros casos, lo es la otra. Pero lo importante es ver que se trata de una cuestión empírica, no de algo que se pueda establecer de antemano. Cualesquiera que sean tus preconcepciones, hay fenómenos que pueden parecer psicológicamente interesantes y que, en ocasiones, acabarán resultando ser efectos interactivos. Entonces tendrás que decir que te ocupas de la actuación y no de la competencia. Pero no porque Chomsky acuñara esta distinción, sino porque ésta es consustancial con la ciencia. Lo que en realidad pretende la ciencia es explicar la varianza observada en términos de la interacción entre variables inobservadas. Suponte que alguien dijera: «Bueno, a mí no me gusta la teoría de la trayectoria que está vigente —la que mantiene que la trayectoria tiene que ver con la interacción que se produce de la mecánica con la fricción y resistencia en el medio aéreo, y cosas por el estilo— y voy a coger una teoría de la trayectoria que sea función de todas esas variables al mismo tiempo». Lo que ocurriría entonces es que se perderían las generalizaciones subyacentes. Pues bien, lo mismo se puede aplicar a la psicología. Algunos fenómenos interesantes —quizá los que son más fáciles de medir, o quizá los que son más obvios para el ser humano— van a acabar siendo efectos interactivos. Tendrás que admitir entonces que la psicología se ocupa de ellos solamente de forma secundaria, en el sentido de que, suponiendo que sepamos cuáles son las variables subyacentes, podremos predecir los fenómenos observables como interacciones. Según esto, la ciencia

tiene como algo inherente el moverse hacia una mayor abstracción, trata de abstraerse de la riqueza de las variaciones superficiales para dar con las variables subyacentes que entran en interacción. Por eso, siempre que se intenta seriamente hacer ciencia, parece como si uno se moviera en la dirección de lo abstracto, de la lógica y las matemáticas. Porque lo que estás intentando hacer es descubrir lo inobservable, abstraerte de la interacción observable y encontrar aquello que no es observable pero que está produciendo o causando los fenómenos que puedes observar. Por lo tanto, no creo que la psicología tenga nada de especial a este respecto. Creo que la distinción competencia-actuación no es más que una forma de referirse a la estrategia principal de la ciencia, que es la de abstraer a partir de los efectos de interacción.

**GA:** La psicología del sentido común también versa sobre deseos, emociones, afectos, motivos, etc. Sin embargo, las propuestas que se hacen para fundamentar las nociones del sentido común están relacionadas principalmente con la cognición. Tú mismo has señalado muchas veces que aquello de lo que te vas a ocupar en tal libro o en tal artículo se limita principalmente a los problemas relacionados con la psicología de los procesos cognitivos. A este propósito, y aun cuando no sea tu principal tema de interés, ¿podrías dar tu opinión sobre las posibilidades que se le ofrecen a la psicología al tratar de estudiar todo ese conjunto de las así llamadas variables orécticas (deseos, emociones, afectos, etc.)? Como sabes, para muchos psicólogos de corte tradicional, estos son los temas verdaderamente interesantes —los temas «calientes»— de los que debe ocuparse la psicología.

**JF:** Pues les deseo toda la suerte del mundo, pero la verdad es que yo no sabría cómo empezar. Existen, desde luego, ciertas relaciones entre las emociones y los estados cognitivos, por ejemplo, que ambos son típicamente intencionales. Del mismo modo que crees que está lloviendo, puedes estar enfadado porque esté lloviendo. Ambos estados tienen la propiedad de referirse a cosas y de tener objetos intencionales. Al menos desde el enfoque del lenguaje del pensamiento, la intencionalidad de la cognición es el factor clave, por el que conecta con las nociones de representación y demás. Pienso que esto podría ser también compartido por las teorías de la afectividad. Ahora bien, si es que hay más ciencia que hacer en este área, yo no tengo ni idea de cómo hacerla.

Me temo que podamos encontrarnos en psicología con un caso particular del principio general que se cumple de forma bastante extendida en las ciencias, el principio de que cuanto más interesante es un fenómeno para el ser humano, tanto más resulta ser algo que no puedes estudiar científicamente. Si te fijas en las ciencias físicas, casi todo aquello que nos interesa humanamente del mundo que nos rodea —los colores, los paisajes, los objetos de tamaño medio, etc.— ha de quedar al margen tan pronto como te pones a hacer física. La clase de las cosas que son verdes es tan heterogénea, desde el punto de vista físico, que aún no disponemos de una buena teoría física acerca de qué es lo que hace que algo sea verde. Y, de hecho, ni siquiera estaríamos interesados en ella si no fuera porque lo verde viene a ser una propiedad a la que responde la gente. En el desarrollo de la ciencia, se puede comprobar una y otra vez que las propiedades del mundo que son más accesibles al ser humano, aquellas que nos preocupan, o sobre las que se escriben novelas, no son las mismas que las que pueden ser estudiadas científicamente. Lo cual no significa, por supuesto, que no sean reales, o que no sea razonable interesarse por ellas. Después de

todo, somos seres humanos y nos interesa todo aquello que le interesa al ser humano. Sin embargo, otra cosa es que constituyan un dominio científico, un dominio en el que se considere la mayor parte de la varianza como efectos de interacción y se descubran profundas generalizaciones sobre las causas subyacentes. Lo que hace la ciencia de forma prácticamente invariable es coger aquellos fenómenos, o aspectos de los fenómenos, que son interesantes no porque se lo parezcan al ser humano en general, sino precisamente porque constituyen las causas subyacentes de los procesos superficiales. Es como si aquello que nos interesa a nosotros del mundo no coincidiera con lo que a Dios le interesa. Y sospecho que esto mismo pasa en la psicología. Sospecho, por ejemplo, que cosas tan interesantes para nosotros como las diferencias individuales, la teoría de la afectividad, y demás, puedan, sin embargo, no ser científicamente interesantes. Al considerar otras especies, sí que nos abstraemos de las diferencias individuales. Es verdad que, entre conejos, no hay mucho para elegir y, por razones obvias, no queremos considerar a nuestra especie del mismo modo; pero podría ser que, cuando se hace ciencia de nuestra especie, cuando se hace la psicología de nuestra especie, eso fuera lo que hubiese que hacer. Ciertamente en la lingüística eso es lo que parece ocurrir: las diferencias entre individuos no se consideran de especial importancia. Desde el punto de vista humano, lo que más nos puede interesar sobre las personas es, por ejemplo, cómo influyen las emociones en la cognición, o cómo aquellas se ven influidas por ésta. Puede ocurrir que eso sea sobre lo que haya que escribir novelas y, hasta ahora, ha tenido bastante éxito escribirlas. Pero puede ocurrir también que, en cuanto a poder de generalización, escribir novelas sea lo mejor que se pueda hacer. Quiero decir simplemente que no veo cómo se puede hacer ciencia sobre estos temas. Y si es que hay que hacerla, me parece que todavía nadie ha tenido éxito en el empeño.

**GA:** En otras disciplinas suele ocurrir que, aun cuando los temas que tratan no sean intuitivamente interesantes para el hombre de la calle, al menos tienen un amplio impacto social a través de la ciencia aplicada. No sé muy bien si esto mismo se puede decir de la psicología. A veces parece como si la psicología estuviera demasiado sometida a las demandas sociales de la gente que busca consejo, apoyo y curación, habiéndose comprobado que hay muchos efectos placebo en todo ello. En cualquier caso, ¿cuáles crees que son las perspectivas de una psicología aplicada efectiva?

**JF:** Mi sospecha es que de existir una psicología aplicada, la habrá en el campo de la inteligencia artificial. Esto significa que las aplicaciones de la psicología se van a centrar básicamente en la construcción de máquinas inteligentes. Me sorprendería mucho si resultara que, por ejemplo, pudiéramos encontrar alguna forma de ayudar a los pacientes afásicos. Podremos, desde luego, descubrir muchas cosas sobre la afasia o sobre la esquizofrenia, pero me sorprendería mucho que descubriéramos algo que nos permitiera curar a la gente que las padece. Me parece que si hay algo que se puede hacer en estas áreas, será la biología quien lo haga. Si se descubre alguna vez un tratamiento efectivo de la afasia, o de la enfermedad de Alzheimer o de la esquizofrenia, creo que provendrá de la biología, ya que se trata en dichos casos de perturbaciones fundamentales biológicas. Esto no quiere decir que no se pueda hacer terapia psicológica, aunque en modo alguno a base de aplicar elaboradas teorías psicológicas. Se hará de la forma en que siempre se ha hecho, es decir, entrenando a los interesados a desarrollar intuiciones terapéuticas y dejándoles

tocar de oído. Lo que está bastante claro es que cuando la terapia tiene éxito, éste no depende tanto de las teorías que el terapeuta sostiene como de todo un conjunto de saberes prácticos. Así que no creo que la psicología vaya a sacar provecho de esta forma, si es que saca algún provecho en absoluto. Creo que la forma de sacar provecho consiste en aprender lo que haga falta acerca de la inteligencia para que podamos plantearnos seriamente la posibilidad de construir artefactos inteligentes. Una cuestión interesante es si, ante el fracaso reconocido que ha experimentado la Inteligencia Artificial, ésta podrá tener éxito en algún momento. Pues bien, una manera en la que podría tener éxito, y, asimismo, una manera en la que podría resultar provechosa la psicología cognitiva rigurosa, o la psicología del pensamiento rigurosa, consistiría en llegar a entender suficientemente cómo funciona la inteligencia humana con el fin de ser capaces de construir artefactos inteligentes.

**GA:** Con respecto a los trastornos psicológicos, aun suponiendo que el tratamiento tuviera que ser biológico, como dijiste antes, tendría que efectuarse bajo una cierta inspiración psicológica, ¿no? Por ejemplo, si se trata de curar a los afásicos, tendrías que saber primero lo que implica la afasia; y siendo la afasia un trastorno del lenguaje, habrá que saber algo de la conducta lingüística normal. Por lo tanto, y cuando menos para la evaluación de la afasia, habrá que echar mano de la psicología o, en este caso, de la psicolingüística, ¿no crees?

**JF:** Bueno, yo no creo que sea algo tan obvio. Podrías querer estudiar a los afásicos porque quisieras entender el lenguaje y eso podría tener éxito o no. Pero si quieres estudiar a los afásicos de cara a encontrar un tratamiento, no veo claro que para ello tengas que entender la afasia como una patología del lenguaje. No sé lo que tendrías que hacer, pues no sé el tipo de tratamiento que podría valer. Dicho de otra manera: la mayoría de las enfermedades que podemos curar no las comprendemos. Alguien da con una forma de elaborar la píldora adecuada, y muchas veces es por ensayo y error, explorando todo un conjunto de productos químicos. Al final, llegar a ser capaz de aplicar una terapia aun cuando no entiendes realmente por qué funciona. Y supongo que podríamos incluir ahí hasta la misma aspirina. Por lo tanto, dudo de que vayamos a ser capaces de conseguir algo para la curación de la afasia, pero si es que lo somos, no me parece nada obvio que debamos suponer que lo conseguiremos gracias a una comprensión teórica de la enfermedad. No sé si los afasiólogos, realmente en el fondo, estudian la afasia porque creen que es la manera de descubrir su curación o por la razón, perfectamente respetable, de que piensan que ello les enseñará algo acerca del lenguaje, o de la cognición, o algo por el estilo.

**GA:** ¿Y qué ocurre, por ejemplo, con los trastornos de aprendizaje? ¿Sería posible avanzar algo en este área y otras relacionadas?

**JF:** Si es una cuestión de lo que es posible, seguro que es posible. Pero si es una cuestión sobre la que tuviera que apostar, yo no apostaría por ello. Si lo que me ofreces es un provecho práctico, mi sospecha es que, de hecho, los trastornos de aprendizaje no van a resultar ser trastornos psicológicos. Sí que van a ser, sin duda, alteraciones de procesos psicológicos, pero el comprender cuál es el carácter de la alteración va a tener muy poco que ver con comprender su etiología, y aún menos con curarla. Todo esto es una conjetura, y como tal podría estar completamente equivocada. Así que ten por seguro que no voy a impedir a nadie que se dedique a este tipo de investigación.



GA: Podemos pasar ahora a otros temas más específicos. En concreto, hay tres temas sobre los que me gustaría preguntarte. El primero es el de la modularidad, a continuación el de la investigación sobre el lenguaje, y finalmente el conexionismo. Así que empecemos por el tema de la modularidad. A primera vista, se puede producir algún desconcierto cuando se tratan de compaginar los postulados asociados a la hipótesis del lenguaje del pensamiento, considerada como teoría general de la cognición, y los asociados a la supuesta clase de los sistemas modulares, donde el énfasis principal recae en la especificidad de dominio y el encapsulamiento informativo. Parece ser que el lenguaje del pensamiento refleja principalmente el funcionamiento de los procesos centrales (pensar, razonar, etc.), pero harían falta otros sublenguajes para hacerse cargo de procesos más periféricos (la percepción visual, la integración motora, etc.). Es cierto que, al final de tu libro *El lenguaje del pensamiento*, dejas claro que el «mentales» tiene la suficiente flexibilidad como para ser compatible con una amplia variedad de códigos. Pero, al mismo tiempo, en *La modularidad de la mente* insistes en que cada módulo está cerrado a las influencias procedentes de otros módulos y de los procesos centrales. ¿Cómo se integran entonces las dos partes de esta historia? No sé si, para ti, será la misma historia, pero parece como si, al menos, tuviera dos partes bien diferentes.

JF: Sí, yo creo que la historia tiene dos partes, pero que una de ellas viene a ser una especialización de la otra. Para mí, la noción de que existe un lenguaje del pensamiento y la noción de que los procesos mentales son computacionales responden exactamente a la misma idea; las computaciones no son más que operaciones sobre representaciones. Ahora bien, los supuestos procesos modulares —el que los haya de hecho es una cuestión enteramente empírica— son considerados en sí mismos como computacionales. Un caso paradigmático de proceso modular (según la teoría estándar) es el del análisis sintáctico («parsing»), que consiste en una operación que se efectúa sobre la representación de oraciones. Y prácticamente para todo el mundo, el «parsing» es asimismo un caso paradigmático de operación computacional. Por lo tanto, no existe incompatibilidad entre la idea de que algunos procesos mentales puedan ser modulares y la idea de que, a la vez, puedan ser procesos que se efectúen, o se definan, sobre representaciones. En este sentido, la teoría del lenguaje del pensamiento —es decir, la teoría computacional de la mente— es una teoría que se aplica a lo largo de todo el espectro de procesos que compartan de alguna forma las propiedades del pensamiento. En contraste con esto, el tema de la modularidad se refiere al modo en que los recursos informativos se distribuyen entre los procesos computacionales. Concretamente, se refiere a que algunos procesos computacionales puedan estar relativamente encapsulados con respecto al trasfondo general de opciones que tiene el organismo; encapsulados con respecto al registro general de información acerca del mundo y con respecto al sistema general de valores y metas. Ahora bien, la hipótesis de que, por ejemplo, pueda darse un aislamiento informativo entre la percepción y la cognición, es perfectamente compatible con que ambas sean procesos computacionales, es decir, procesos definidos sobre representaciones.

Por otra parte, me parece a mí que, de entre las distintas clases de computaciones que se han estudiado, las mejor comprendidas han sido aquellas que están relativamente encapsuladas y que son, por ello, relativamente

modulares, no-globales, y limitadas en cuanto a sus recursos informativos. Hablando en general, y por lo que se ha visto hasta ahora, cuanto más cosas sabemos que pueden influir en un proceso mental, menos posibilidades tenemos de entenderlo. Los procesos que parece que entendemos mejor —y eso que queda mucho para que los entendamos bien del todo— son procesos como los del «parsing» y la comprensión del lenguaje, ciertos aspectos del reconocimiento visual, es decir, áreas de la percepción y quizá algo de la integración de la conducta motora. Lo que llama la atención de estos procesos es que la información que se utiliza al ejecutarlos parece ser muy local y constreñida. Por el contrario, se nos da bastante mal entender aquellos procesos que tienen acceso a bases de datos muy grandes y que se utilizan de forma muy eficiente; es decir, aquellos procesos que parecen ser no-modulares.

Así pues, la idea es que el modelo computacional es aplicable a todos los procesos, tanto a los relativamente modulares como a los relativamente no-modulares; ambos son computaciones, operaciones sobre representaciones. El problema está en que, solamente en el caso de los relativamente modulares, tenemos hipótesis aceptables sobre qué clase de computaciones son esas.

**GA:** La primera ley de Fodor sobre la inexistencia de la ciencia cognitiva («Cuanto más global es un proceso cognitivo, tanto menos se entiende», cfr. Fodor, 1983-86, pág. 151) ha provocado un gran escepticismo sobre la posibilidad de traspasar el estudio de procesos tan básicos como la percepción visual o el análisis de oraciones. Sin embargo, se podría pensar en proyectar la modularidad a los pretendidos sistemas centrales de tal modo que, conforme se avanzara más en el conocimiento de los procesos superiores y se descubrieran relaciones precisas entre ellos, pudieran acabar siendo considerados también como modulares. En cualquier caso, la cuestión es la de cómo superar las dificultades que provienen del manejo de una base de datos tan grande. En tu opinión, ¿hay alguna expectativa de que se puedan estudiar bajo un enfoque modular temas tan tradicionales en la psicología cognitiva como el razonamiento, la resolución de problemas, etc.? ¿Y qué será del estudio de la acción y otros sistemas de salida como, por ejemplo, la producción del lenguaje?

**JF:** Creo que esa es exactamente la forma de plantear la cuestión. Y hay dos maneras de afrontarla. Tomemos el caso del razonamiento, que de forma paradigmática es no-modular, al menos cuando se trata del razonamiento no-demonstrativo, inductivo o analógico. Una manera de abordarlo es pensando que si lo miras muy de cerca, acabarás considerándolo como modular; disponemos de diversos tipos de estrategias que nos permiten tratar esos problemas globales como si tuvieran soluciones locales. Esta ha sido la línea normalmente adoptada por la Inteligencia Artificial, la cual me parece una estrategia de investigación perfectamente razonable. Sin embargo, no creo que haya funcionado demasiado bien, ya que lo que pienso que ocurre es que, en todas estas áreas no-modulares hay algo que no entendemos. Se trata, más o menos, del problema de la relevancia. En los sistemas modulares, aquello que es relevante para la solución de los problemas está predeterminado, está definido arquitectónicamente. Quiero decir, por ejemplo, que si aceptas la historia estándar de la modularidad acerca del «parsing» —cosa que, por supuesto, no todo el mundo hace—, entonces el que haya una base de datos disponible para analizar una oración constituye un rasgo de la arquitectura cognitiva. No es algo que esté a disposición de cualquiera. Los únicos datos disponibles para el «parsing» son los que incluye una gramática o algo de ese tipo. Con lo cual, el

problema de la relevancia en un sistema modular viene a estar resuelto de una vez por todas. Queda resuelto por la arquitectura del sistema y no hay nada que podamos hacer sobre ello. En un sistema modular, la relevancia no es un parámetro que fluctúe libremente. Por otro lado, creo que la razón por la que tenemos tanta dificultad para entender el pensamiento, en el amplio sentido en el que nos interesa, es la de que no entendemos cómo se maneja la relevancia. Aquí la relevancia no viene dada con la arquitectura. Es un parámetro que está bajo el control del sistema, sin que sepamos cómo éste lo controla. Lo que, de forma inmediata e intuitiva, es obvio acerca del razonamiento es que las personas están increíblemente dotadas para captar lo que es relevante de cara a solucionar un problema, para captar al vuelo nuevas relaciones de relevancia conforme se va realizando la tarea de resolver el problema. Se nos da tan bien esto —el sortear obstáculos inesperados y llegar a la otra orilla— que ni siquiera nos damos cuenta de ello. Parece que estamos extraordinariamente capacitados para manejar la relevancia, para captar instantáneamente que, si el problema tiene tal o cual propiedad, entonces es aquello que hay ahí lo que es relevante para solucionarlo. No hubiéramos pensado en ello antes de advertir la estructura del problema, pero tan pronto la advertimos, nos damos cuenta. Te das cuenta, por ejemplo, de que el asa de la perola está caliente porque ha estado en el fuego y, por ello, sin siquiera pensarlo, sacas el pañuelo porque sabes que, si coges la perola con el pañuelo, éste te aislará del calor. No tienes por qué haberlo previsto con anticipación, sino que, simplemente, ves que ese es el tipo de solución que requiere el problema y llegas sin más a ella; ¿quién puede pensar que uno esté ahí sentado diciéndose a sí mismo: «Ahora cuando vaya a la cocina debo tener en cuenta que si pongo el pañuelo en el asa de la perola no me quemaré al cogerla»? Estas cosas no se planifican con esa anticipación, sino que descubrimos las relaciones de relevancia como si fueran surgiendo mientras se va haciendo la tarea. Pienso que, desde luego, la tarea es computacional, ya que la mente es, en cierto sentido, una computadora, o, si se quiere, una máquina de lenguaje del pensamiento. Pero me parece que no tenemos ni una pista acerca de cómo funciona eso. Es un profundo misterio y, hasta que no lo resolvamos, no vamos a tener máquinas inteligentes. Podremos tener máquinas perceptivas, máquinas que analicen oraciones o que puedan recuperar la forma visual de un objeto, pero no creo que vayamos a tener máquinas que piensen. La razón por la que no lo creo es que, para abordar este problema, no van a valer estrategias fragmentarias por las que se enfoque de una manera más local de como aparente; la razón es que hay algo que no entendemos.

Llama mucho la atención que esto no ocurra sólo en la psicología, sino también en la filosofía de la ciencia, donde uno de los problemas vigentes es que no acabamos de entender el hecho de la confirmación científica. Y una de las razones consiste, de nuevo, en que la misma no depende de propiedades localizables —como podría ser la aplicación de determinados enunciados a determinados conjuntos de datos—, sino de propiedades muy globales de las teorías científicas, como su naturalidad, simplicidad, parsimonia y demás; las cuales se asemejan a la propiedad de la relevancia en que parecen depender, al mismo tiempo, de la estructura total de la base de datos y de la estructura total de la teoría del conocimiento. Y no tenemos ni idea de cómo medir eso y de cómo hacer que el sistema responda a las propiedades globales de sus representaciones. Todas las buenas teorías computacionales de que disponemos

son muy locales. Así que yo creo que hay algo que, simplemente, se nos escapa, quizá todo un conjunto de cosas que no entendemos. Aunque, como dije antes, siempre queda la estrategia alternativa de decir: «Mira, todo esto es una ilusión; con lo único que realmente contamos es con una gran cantidad de microprocesos, más o menos modulares, que se ponen en marcha mediante estrategias más o menos locales, y lo que deberíamos hacer es dejar ya de enredar con los problemas globales y reducirlos a problemas locales». A lo mejor esta estrategia funciona, en cuyo caso, estupendo. Pero aquí ocurre como con las cosas que hablábamos antes. Se trata de una conjetura, y si yo tuviera que apostar, apostaría contra ella. Yo apostaría que, cuando finalmente entendamos de qué va la cosa, nos encontraremos con algo que es verdaderamente distinto en cuanto a su estructura. Será computacional, pero con computaciones diferentes de las que caracterizan a los procesos modulares.

**GA:** A veces da la impresión de que estamos ante una especie de paradoja. Si el pensamiento se resiste tanto a ser comprendido, ¿hasta qué punto se puede utilizar el enfoque del lenguaje del pensamiento para estudiar el pensamiento? Por otra parte, tampoco estoy seguro de que, con esa clase de reduccionismo modular al que te acabas de referir, podamos ir muy lejos. En cualquier caso, has señalado algo que parece muy interesante: que las áreas en las que se ha progresado, y que se corresponden con los procesos modulares, están predefinidas por la arquitectura del sistema. Quizá por eso es por lo que su existencia no es predecible desde fuera. En último término, parece como si el conjunto de sistemas modulares incluyera toda una serie de facetas bastante asistemáticas y heterogéneas (capacidades musicales, reconocimiento de caras, relaciones espaciales, análisis de oraciones, etc.).

**JF:** Quiero pensar que los sistemas modulares, en general, están altamente especializados desde el punto de vista biológico y quizá por eso forman un conjunto tan poco natural. Una vez más, no creo que se pueda decir lo mismo del pensamiento; podemos pensar en tantas clases de cosas diferentes que, claramente, no pueden estar determinadas por lo que la biología quiera pretender que pensemos, como es el caso de la teoría de los números —por usar el ejemplo preferido por Chomsky—, o la física nuclear, o cualquier otra cosa. Por el contrario, es muy probable que muchos de los sistemas modulares —como el reconocimiento del lenguaje, el reconocimiento de caras, o el reconocimiento de formas en tres dimensiones— se correspondan con sistemas antiguos y altamente especializados desde el punto de vista biológico.

**GA:** Podemos conectar aquí con el tema de la investigación sobre el lenguaje. Si te parece, sería bueno empezar aclarando la diferencia entre tu postura y la de Chomsky con respecto al módulo del lenguaje. ¿Se podría considerar como un módulo el famoso «órgano del lenguaje» del que habla Chomsky tan a menudo?

**JF:** Como pasaba con el funcionalismo, también aquí nos encontramos con un problema de terminología bastante lamentable. Se trata de dos acepciones distintas de la modularidad. En gran parte, es culpa mía; quizá no debería haber utilizado este término, pues creo que Chomsky lo usó primero. De todas formas, lo que yo entiendo por sistema modular es una clase determinada de sistema computacional, cuya modularidad consiste en todo un conjunto de restricciones sobre la transmisión de información hacia dentro y hacia fuera del sistema. Así que, para mí, la noción de modularidad es básicamente una noción computacional. Para Chomsky, la noción de modula-

ridad viene a ser una combinación de innatismo y especificidad de tarea. De ahí que cuando dice que el lenguaje es modular, lo que quiere decir, según creo, son dos cosas. En primer lugar, que los principios que rigen la percepción del lenguaje no son una mera especialización de los principios cognitivos generales, por lo que los principios que se usan al hablar un idioma no tendrían mucho que ver con los que se usan para hacer, por ejemplo, la declaración de la renta o cualquier otra cosa. Esta sería la noción de especificidad de dominio o de tarea. Y en segundo lugar, está la idea de que estos sistemas son innatos. Estos dos aspectos vienen a constituir condiciones sobre los cuerpos de información que se utilizan al realizar esas tareas. Pero no son nociones computacionales, no implican restricciones sobre las computaciones que se llevan a cabo al realizar las tareas. Lo que diría a esto Chomsky sería: «Mira, yo no sé cómo se efectúa la percepción del lenguaje». (Yo creo que Chomsky no está demasiado interesado en el procesamiento psicológico, al menos no tanto como lo está en los cuerpos de información utilizados por dichos procesos). Así pues, Chomsky dice: «No me interesa la cuestión de qué procesos psicológicos, qué computaciones, se dan en la percepción y producción del lenguaje; lo que me interesa es el cuerpo de información que utilizan esos procesos. Y he aquí lo que yo pienso acerca de ese cuerpo de información, llamémosle gramática. Creo que ésta es probablemente innata y muy diferente de los cuerpos de información que se usan para otras cosas». La de Chomsky es, pues, una noción de modularidad perfectamente justificada. Pero existe esa otra noción de modularidad que vendría a decir: «A mí lo que me interesa es, por ejemplo, el análisis de oraciones considerado como un proceso computacional. Los procesos computacionales son esos que, entre otras cosas, intercambian información unos con otros. Los procesos computacionales modulares son aquellos cuyos intercambios informativos satisfacen ciertas constricciones generales que se especificarán a continuación...». Por lo tanto, hay dos nociones realmente distintas, aun cuando estén obviamente interconectadas de muchas maneras. Por ejemplo, si se piensa que los sistemas modulares están, en gran parte, especializados biológicamente, entonces su carácter innato podría ser un rasgo de la modularidad en ambos sentidos. Pero básicamente son dos nociones distintas.

GA: Así pues, hay algunos rasgos que, al menos en principio, son comparados por tu idea de la modularidad y la de Chomsky. Por consiguiente, no son incompatibles.

JF: Efectivamente, no son incompatibles.

GA: Con respecto al lenguaje, me gustaría hacerte un par de preguntas más. En primer lugar, ¿cómo ves en este momento, casi treinta años después de los primeros pasos de la psicolingüística, las relaciones entre lingüística y psicología de cara a hacer psicolingüística? La segunda cuestión es de tipo metodológico y guarda relación con la sensibilidad de las tareas que los psicolingüistas usamos para afrontar empíricamente los problemas de interés teórico; me refiero al uso de los tiempos de reacción, los potenciales evocados, los registros de movimientos oculares, etc.

JF: Bien, yendo a lo primero, no creo que esas sean dos ciencias diferentes en algún sentido particular. Creo que son simplemente dos bases de datos diferentes a las que se han agarrado unos u otros. El conjunto de fenómenos que han utilizado los lingüistas incluye juicios intuitivos, ciertos rasgos de la adquisición, algunos aspectos de los universales, y cosas así. Con este tipo de

datos, tratan de construir una teoría de, por ejemplo, las estructuras informativas que se manejan en la producción de oraciones, en el aprendizaje de la lengua, en la percepción de oraciones y demás. Los psicólogos se han agarrado a otro tipo de datos: los tiempos de reacción, el desarrollo ontogenético, los errores del habla y demás técnicas que se impongan en un momento dado. De forma ideal, estos dos cuerpos de datos deberían dar lugar a resultados convergentes; debería ocurrir que la teoría que proporcionara la mejor explicación de cómo se comprenden las oraciones con cláusulas de relativo también diera la mejor explicación de los tiempos de reacción que se producen cuando se comprenden dichas oraciones. Obviamente, ninguno de estos cuerpos de datos son infalibles, ni, por así decirlo, epistemológicamente primitivos, o privilegiados, o cosas por el estilo. Son simplemente cuerpos de datos a los que se podría recurrir cuando se intenta construir una teoría razonable de lo que pasa en la cabeza de la gente. Cuando hay teorías en conflicto o cuando las teorías dejan de producir resultados convergentes, ello se podría deber a todo tipo de razones. Una razón puede ser la de que el cuerpo de datos en cuestión no es lo bastante sensible, o no está suficientemente articulado en algún área. Otra razón puede ser la de que el cuerpo de datos tiene mucho ruido. Los lingüistas han tendido a pensar, aunque creo que sin mucha justificación, que los datos experimentales son más ruidosos, groseros, y difíciles de tratar que los datos intuitivos o los datos acerca de los universales. Pero creo que esto no es más que un prejuicio y no hay que darle la menor importancia.

Otro asunto diferente es que las dos disciplinas puedan tener distintos puntos de vista sobre la manera de establecer la distinción entre actuación y competencia; es decir, sobre qué cosas deben ser consideradas como efectos interactivos y cuáles deben ser los aspectos básicos subyacentes. Creo que ésta es una cuestión empírica y no metodológica. Uno tiene que dar con la explicación correcta y, si es la correcta, debería aparecer así desde el punto de vista de todos los datos, esto es, tendría que poder dar cuenta de todos los datos fiables, bien sea en términos de los efectos interactivos o bien en términos de las variables subyacentes. En este punto, decir cuáles son unos u otros no deja de ser una simple conjetura. Es cierto que mucho de lo que hemos estado haciendo en psicología, y una buena parte de lo que hemos estado haciendo en psicolingüística durante los últimos treinta años, han sido experimentos sobre experimentos, tratando de comprobar lo que miden de hecho nuestras medidas o instrumentos. Creo que una forma de hacer tales experimentos es calibrando nuestros procedimientos experimentales frente a los casos más claros en que pueda parecer que los lingüistas han acertado. Cuando, por razones lingüísticas, tienes claro cuáles son las propiedades de la descripción estructural de una oración, entonces parece razonable que te puedas decir a ti mismo: «De acuerdo, pero ahora voy a poner a prueba mis procedimientos experimentales para ver si detectan o no esas propiedades». Ahora bien, lo que no me parece evidente es que, cuando la respuesta sea que no, ello se deba a algún fallo en los procedimientos experimentales. Después de treinta años, contamos ya con bastantes casos en que ha resultado que eran los lingüistas los que tenían las estructuras y las reglas equivocadas; no parece, pues, que se pueda hablar aquí de privilegios epistemológicos o metodológicos. Uno tiene que seguir trabajando, por un lado en lingüística y por otro en psicología, y ver si puede conseguir que converjan. Pienso que hay áreas en

que convergen y otras áreas en que no, siendo esto lo que determina los problemas de investigación en el campo.

Por lo demás, creo particularmente que el interés de la psicolingüística no tiene por qué estar orientado exclusivamente a la obtención de una teoría del procesamiento lingüístico como tal. También puedes recurrir a la psicolingüística si estás interesado en las cuestiones de la modularidad, en cuestiones acerca de las relaciones entre sistemas perceptivos y cognitivos, y en cosas así. Precisamente porque contamos con una noción relativamente compendiosa y avanzada de cuál es la estructura de una oración, la teoría del «parsing» es un caso magnífico de prueba, quizá mejor que cualquier otro.

**GA:** Bien, por último, me gustaría que hicieras algunas consideraciones sobre algo tan de moda como el conexionismo, sus teorías y su enfoque general de la psicología y de la ciencia cognitiva. La primera pregunta sería si realmente se trata de una moda. En segundo lugar, ¿qué aportaciones crees que ha hecho el conexionismo al desarrollo de la investigación psicológica actual? Y finalmente, ¿cómo ves el futuro del programa conexionista?

**JF:** Bueno, en cierto sentido, no soy la persona más adecuada para responder, ya que tengo bastantes prejuicios con respecto a todo esto. Lo que sí creo es que no se puede subestimar la diferencia que existe entre los modelos conexionistas y los modelos clásicos de la ciencia cognitiva. La base del conexionismo consiste en rechazar por completo la idea de computación que subyace a la ciencia cognitiva clásica. Dicha idea es que los procesos mentales son computaciones, es decir, transformaciones de representaciones. Nada de esto aparece en las teorías conexionistas, lo cual tiene toda una serie de consecuencias con respecto a la capacidad representacional de estas teorías. Puesto que no se dispone de un lenguaje del pensamiento, hay muchas cosas verdaderamente difíciles de explicar desde un modelo conexionista. Hace ya como treinta años que Chomsky replicó a Skinner —que, por cierto, no era un conexionista, aunque se pueda aplicar el mismo punto— en términos parecidos a estos: «Mira, no puede ser verdad que aprender una lengua equivalga a aprender un conjunto de hábitos, porque las lenguas son productivas. Lo que llega a dominar el que domina un idioma —suponiendo, de nuevo, la distinción competencia/actuación y que nos abstraemos de las limitaciones de la memoria y demás— es todo un conjunto infinito de capacidades. Lo que adquiere el que adquiere un idioma es un conjunto infinito de capacidades, a saber, las capacidades para comprender cualquier oración de entre un número arbitrariamente grande». Pues bien, el carácter infinito del lenguaje depende de que la descripción estructural de las oraciones tenga una cierta clase de propiedades combinatorias. Y ésta es precisamente la clase de propiedades que no tienes en los modelos conexionistas. Como consecuencia de ello, los modelos conexionistas son realmente, y en lo más profundo, modelos de carácter finito. Representan la mente como un sistema finito, y no sólo en el sentido vulgar de que la memoria sea finita o de que se acabe el tiempo, los motivos y demás, sino en el sentido de que las capacidades están representadas como algo intrínsecamente finito. Por eso resultan tan difíciles de entender todas aquellas actividades productivas que son características de los seres humanos y, por lo que sabemos, de muchas mentes animales. Así pues, creo que ésta es la razón para no tomar en serio la concepción conexionista de la mente, al menos en cuanto concepción general, independientemente de los modelos de procesos específicos que se puedan proponer.

Hay, por ejemplo, procesos mentales como la búsqueda en el léxico que son esencialmente finitos. Las capacidades lingüísticas de cara a las oraciones son presumiblemente productivas, pero el conjunto de palabras del idioma no es más que una lista. Así que se podría pensar que el modelo de capacidades como la del reconocimiento de palabras podría ser muy bien un modelo esencialmente finito. Para estos casos, el sistema conexionista podría resultar una idea útil; pero la consideración básica de la mente como no-productiva me resulta francamente inadmisibile.

Creo que se podrían decir cosas parecidas acerca de los dispositivos conexionistas para el aprendizaje. Estos vienen a ser como máquinas de inferencia estadística, por lo que guardan relación con los modelos clásicos de la estadística, con modelos computacionales clásicos de la inferencia estadística, de forma análoga a como las calculadoras de bolsillo guardan relación con los modelos clásicos de la teoría de los números. En otras palabras, se trata de dispositivos analógicos para efectuar un tipo de inferencias que pueden ser efectuadas igualmente, y de formas bastante bien comprendidas, por dispositivos computacionales clásicos no-analógicos. Ahora bien, creo que hay numerosas razones para pensar que las partes difíciles del aprendizaje, aquellas que no llegamos a entender (incluyendo el aprendizaje del lenguaje, el aprendizaje de habilidades, y demás), no son simplemente procesos de inferencia estadística. La concepción del aprendizaje basada en la construcción de modelos internos de las redundancias estadísticas del medio —que, por cierto, ha sido la concepción característica de los modelos empiristas desde Hume— ha mostrado ser un completo fracaso. Al ser ésta la explicación del aprendizaje que sostiene el conexionismo, creo que la probabilidad de que produzca efectivamente resultados interesantes es cada vez menor. Con esto no se niega que, bien por razones pobres (como la de que el sistema se asemeja presuntamente al cerebro) o bien por razones válidas (como que necesitemos entender mejor las arquitecturas en paralelo, y la conexionista lo es), estos modelos vayan a seguir explorados por mucho tiempo. Después de todo, ya han sido explorados durante bastante tiempo. Básicamente, suponen una prolongación de la tradición empirista y asociacionista a la que ha pertenecido la mayor parte de la psicología anglo-americana. Esto quiere decir que el conexionismo está en continuidad con Locke, Hume y Bekeley de una forma en que no lo están los modelos de la ciencia cognitiva, los modelos del lenguaje del pensamiento. Los modelos del lenguaje del pensamiento tienen su origen en las teorías de la computación, mientras que los conexionistas representan una vuelta al asociacionismo clásico.

**GA:** Y también están vinculados al conductismo, ¿no?

**JF:** Tienen una vinculación con el conductismo en cuanto que, como éste, son también asociacionistas. Sin embargo, no son antimentalistas como lo fueron los conductistas. Esta es la diferencia. El conexionismo está más relacionado con el empirismo británico de corte clásico y, dentro del conductismo, más con gente como Hull que con gente como Thorndike o Skinner.

**GA:** ¿No te recuerdan mucho al conductismo las leyes del aprendizaje que postulan los conexionistas, con reminiscencias claras de la Ley del Efecto de Thorndike o la Ley de Hebb? En cualquier caso, hay algo que me ha llamado bastante la atención de lo que has comentado antes. ¿En qué sentido dices que los modelos conexionistas no son computacionales? Al menos, a primera vista, hay un gran despliegue de aparato computacional, ¿no es así?



**JF:** Esto es largo de contar, pero desde luego no son computacionales en el sentido en que lo son los procesos de una máquina de Turing. Lo que ocurre con una máquina de Turing, o con cualquiera de los sistemas que se sitúan en la tradición de Turing —considerada como la tradición de los modelos computacionales clásicos—, es que las operaciones que efectúa la máquina son, en realidad, procesos que deforman las representaciones. Escribes algo en la cinta, la máquina opera sobre lo que has escrito y produce algo distinto que también escribe en la cinta. Los modelos conexionistas no son así. Transmiten, como si dijéramos, excitación entre los estados representacionales, pero ninguno de los procesos del modelo consiste en una operación sobre representaciones. Esa es una gran diferencia y tiene que ver con el hecho de que las representaciones clásicas tienen estructura sintáctica; las máquinas de Turing operan sobre las representaciones en virtud de su sintaxis. Distinguen secuencias de ceros y unos, determinando si cada una de estas secuencias se transforma de una manera o de otra. La máquina responde, pues, a la estructura interna de los símbolos que está presente. En lugar de eso, las máquinas conexionistas transmiten excitación entre los estados representacionales de que disponen. Pero dichos estados no llegan a ser manipulados. En una máquina conexionista, lo único que, en cierto modo, se modifica es el nivel de excitación y resistencia de las conexiones entre los estados. En una máquina de Turing, la población de símbolos se modifica realmente. Unas representaciones se transforman en otras, cosa que no ocurre en una máquina conexionista. La noción de que la computación es el proceso de convertir unos símbolos en otros es realmente la idea básica que subyace a la tesis de que pensar consiste en la transformación de representaciones mentales.

**GA:** Se me ocurre que, si tomáramos en sentido literal la noción de excitación/inhibición, sería bien fácil mostrar que, como muchas veces se ha dicho, el conexionismo no es más que una teoría de la implementación neurológica. Pero quizá se quieran referir a algo más abstracto cuando hablan del mecanismo de excitación/inhibición. ¿Qué piensas tú de esto?

**JF:** Desde luego, hay dos puntos de vista muy distintos acerca de la inferencia. (Una inferencia es una relación entre premisas y conclusiones). Según el punto de vista clásico, lo que ocurre en una inferencia es que las premisas se transforman realmente en la conclusión. Me refiero a que esto ocurre literalmente así. Suponte que tienes una inferencia de la forma «modus ponens». Así que escribes la premisa « $p$  implica  $q$ », escribes la premisa « $p$ », y entonces lo que hace la máquina es mirar las premisas, quitar « $q$ » de la primera premisa y ponerla como conclusión, o si prefieres, mira las premisas y copia la « $q$ » de la primera en la conclusión. Por lo tanto, la relación que se da entre las premisas y la conclusión es la de una deformación de las primeras para llegar a la segunda. Esto es computación y es lo que viene a mantener el punto de vista clásico, a la manera de Turing, sobre la inferencia. Según los modelos interactivos conexionistas, la relación que se da entre premisas y conclusión es una relación de excitación, las premisas excitan la conclusión. Unas y otra están ya presentes en la red, de tal manera que no hay que construir o crear la conclusión; unas y otra están presentes en la red todo el tiempo. De hecho, todos los estados representacionales están presentes desde el principio y lo único que puede cambiar es la probabilidad de que la excitación de un estado produzca la excitación de otro. Así que en algún lugar de la red tienes « $p$  implica  $q$ », en otro sitio tienes « $p$ », y la inferencia para llegar a « $q$ » se

representa diciendo: «Bien, excita esas dos y se excitará el nodo de 'q'». Esta es la razón por la que esto es tan distinto de la concepción clásica. La idea de la inferencia que se obtiene con estos dispositivos asociacionistas no es la de que las premisas se transformen en conclusiones, sino la de que sólo cambia la forma en que fluye la excitación en un espacio donde las premisas, las conclusiones, y todo lo demás que quieras representar, están dados de antemano. Estos modelos no son realmente productivos. Tú no creas la conclusión a partir de las premisas, sino que simplemente dispones las cosas de tal forma que si aceptas las premisas, la conclusión queda aceptada en cualquier otra parte de la red.

Es importante señalar que todo lo anterior vale, suponiendo que nos tomamos en serio el conexionismo, es decir, considerándolo como un modelo de hacer inferencias, y no simplemente como un modelo del mecanismo en el que se implementan las inferencias.

**GA:** Por otra parte, y ya para terminar, ¿crees que valdría la pena proseguir el estudio de algunas de las propuestas del conexionismo, aunque sólo sea como estrategia de investigación que nos permita descubrir nuevas cosas en ciertos dominios, por ejemplo, en el del procesamiento léxico?

**JF:** No me sorprendería. Fíjate, todo el mundo admite, independientemente de sus prejuicios sobre esto, que, entre las cosas que hace la mente, está la de buscar en una lista. Para algunos, la importancia que tenga esta actividad será mayor que para otros, pero nadie duda de que se produzca. Como nadie duda de que la búsqueda en listas sea muy rápida y eficiente. Precisamente, el interés que tienen los modelos finitos en psicología está en que algunos procesos psicológicos, como el de buscar en una lista, son intrínsecamente finitos. Sería un gran avance el que pudiéramos dar con buenos modelos finitos para cosas como la búsqueda en listas. Y quién sabe si la búsqueda en listas no se llevará a cabo mediante procesos que funcionan en paralelo. Sería absurdo no considerar esa posibilidad. Ya tenemos los mecanismos y tenemos las matemáticas. Muchas de las ecuaciones puede resultar que sean ecuaciones de la termodinámica y se podrían, por lo tanto, adaptar a partir de partes de las matemáticas que ya se conocen bien. Sería estupendo, maravilloso. Lo único que pasa es que la mente, en general, no puede ser una máquina conexionista.

Esto es lo que yo pienso del cerebro. Si eres lo suficientemente listo como para descubrir un mecanismo que le pueda ser útil al cerebro, seguro que el cerebro tiene ese y otros veinte más como él. Quiero decir que si a alguien se le ocurre una idea, sería absurdo que no quisiéramos atender a ella por razones meramente ideológicas. Ahora bien, si se trata de algo más que una idea, si lo que se propone es toda una arquitectura general de un sistema que hace inferencias, uno se podría entonces preguntar, ¿puede ser la mente en general un sistema del tipo que proponen los conexionistas? Mi respuesta es que no, que ni hablar. ¿Y podría haber en la mente algunos subsistemas que funcionaran de esa manera? Pues seguro que sí, y sería insensato no tomarse en serio tal posibilidad.

## Referencias

- FODOR, J. A. (1968). *Psychological Explanation*. Nueva York: Random House. (Traducción española en Ed. Cátedra, Madrid, 1980).
- FODOR, J. A., BEVER, T. G. y GARRETT, M. F. (1974). *The Psychology of Language*. Nueva York: McGraw-Hill.
- FODOR, J. A. (1975). *The Language of Thought*. Nueva York: Harper & Row. (Traducción española en Alianza Ed., Madrid, 1985).
- FODOR, J. A. (1981). *Representations*. Cambridge, MA.: MIT Press.
- FODOR, J. A. (1983). *The Modularity of Mind*. Cambridge, MA.: MIT Press. (Traducción española en Ed. Morata, Madrid, 1986).
- FODOR, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.