

En *Cognición social y lenguaje: La intersubjetividad en la evolución de la especie y en el desarrollo del niño*. Lima (Perú): Fondo Editorial de la Pontificia Universidad Católica del Perú..

Intersubjetividad y atribución psicológica.

Pérez, D. y Español, S.

Cita:

Pérez, D. y Español, S. (2014). *Intersubjetividad y atribución psicológica. En Cognición social y lenguaje: La intersubjetividad en la evolución de la especie y en el desarrollo del niño*. Lima (Perú): Fondo Editorial de la Pontificia Universidad Católica del Perú..

Dirección estable: <https://www.aacademica.org/silvia.espanol/129>

ARK: <https://n2t.net/ark:/13683/pH0V/vmq>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

COGNICIÓN SOCIAL Y LENGUAJE
LA INTERSUBJETIVIDAD EN LA EVOLUCIÓN DE LA ESPECIE Y EN EL DESARROLLO DEL NIÑO

Pablo Quintanilla, Carla Mantilla y Paola Cépeda
(editores)

COGNICIÓN SOCIAL Y LENGUAJE
LA INTERSUBJETIVIDAD EN LA EVOLUCIÓN DE LA ESPECIE
Y EN EL DESARROLLO DEL NIÑO



**FONDO
EDITORIAL**

PONTIFICIA **UNIVERSIDAD CATÓLICA** DEL PERÚ

Cognición social y lenguaje

Pablo Quintanilla, Carla Mantilla y Paola Céspedes (editores)

© Pablo Quintanilla, Carla Mantilla y Paola Céspedes, 2014

© Fondo Editorial de la Pontificia Universidad Católica del Perú, 2014

Av. Universitaria 1801, Lima 32, Perú

Teléfono: (51 1) 626-2650

Fax: (51 1) 626-2913

feditor@pucp.edu.pe

www.pucp.edu.pe/publicaciones

Diseño, diagramación, corrección de estilo
y cuidado de la edición: Fondo Editorial PUCP

Primera edición: abril de 2014

Tiraje: 000 ejemplares

Prohibida la reproducción de este libro por cualquier medio, total o parcialmente,
sin permiso expreso de los editores.

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2014-
ISBN:

Registro del Proyecto Editorial:

Impreso en Tarea Asociación Gráfica Educativa
Pasaje María Auxiliadora 156, Lima 5, Perú

ÍNDICE

Prólogo	9
Introducción	
<i>Mente y Lenguaje</i>	
<i>Grupo Interdisciplinario de Investigación</i>	13
PARTE I. MARCO CONCEPTUAL	
Evolución y desarrollo de la cognición social	
<i>Mente y Lenguaje</i>	
<i>Grupo Interdisciplinario de Investigación</i>	33
Evolución y desarrollo del lenguaje	
<i>Mente y Lenguaje</i>	
<i>Grupo Interdisciplinario de Investigación</i>	83
PARTE II. ENSAYOS SOBRE LA COGNICIÓN SOCIAL Y EL LENGUAJE	
Sección I. Ensayos sobre la evolución de la cognición social	167
The nonlinear evolution of human cognition	
<i>H. Clark Barrett</i>	169
The fragmentation of reasoning	
<i>Peter Carruthers</i>	181
La brecha en mente o por qué los humanos no son solo grandes simios	
<i>Robin Dunbar</i>	205
La naturalización de los conceptos morales. Sobre la tesis de la imaginación moral de Mark Johnson	
<i>Claudia Muñoz Tobar</i>	227
La evolución de la atribución psicológica: lectura de mentes y metacognición	
<i>Pablo Quintanilla</i>	241

Sección II. Ensayos sobre desarrollo y estructura de la cognición social	261
El emergentismo y la aparición de la conciencia <i>Ricardo Braun</i>	263
Cómo ser naturalista sin ser reduccionista en el estudio de la mente y el lenguaje: intersubjetividad, comprensión lingüística y atribución psicológica <i>César Escajadillo</i>	311
La sistematicidad de la cognición humana en cuestión <i>Antoni Gomila</i>	333
El vínculo de apego como escenario para el desarrollo de la cognición social temprana <i>Carla Mantilla</i>	351
Intersubjetividad y atribución psicológica <i>Diana Pérez y Silvia Español</i>	371
Sección III. Ensayos sobre la evolución del lenguaje	393
Perspectivas sobre el origen y la evolución del componente sintáctico <i>Paola Céspedes y Gabriel Martínez Vera</i>	395
<i>Homo combinans</i> . Explorando la Gramática Universal minimista <i>Paola Céspedes</i>	415
El «argumento del diseño» y la crítica chomskiana a la explicación adaptacionista de la evolución del lenguaje <i>Marcos Herrera Burstein</i>	435
Sección IV. Ensayos sobre desarrollo y estructura del lenguaje	463
¿Cómo se comunican los infantes en el segundo año de vida? El caso de las funciones declarativa e imperativa <i>María de los Ángeles Fernández Flecha</i>	465
El carácter mental de los sistemas fonológicos <i>Jorge Iván Pérez Silva</i>	501
La naturaleza de la facultad del lenguaje <i>Liza Skidelsky</i>	517
Sobre los autores	541

PRÓLOGO

Este libro es el resultado de un ingente número de horas de estudio y discusión realizado por los integrantes del Grupo Interdisciplinario de Investigación Mente y Lenguaje, entre los años 2010 y 2013. El Grupo nació a partir de informales reuniones interdisciplinarias entre psicólogos, lingüistas y filósofos en el año 2005. Lo que convocó a sus primeros miembros fue el hecho de que las tres disciplinas mencionadas tuvieran similares objetos de estudio (la mente y el lenguaje), pero que, con frecuencia, no dialogaran entre sí ni compartieran resultados ni enfoques metodológicos. A lo largo de muchas horas de presentaciones y confrontación de posiciones fueron apareciendo puntos de contacto, pero también de discrepancia, no solo entre los miembros del Grupo sino también, y quizá especialmente, entre las disciplinas que ellos representaban. No siempre hay consenso entre estas ciencias, ni tampoco al interior de cada una de ellas, respecto de qué se entiende por mente o por lenguaje, y, menos aun, sobre puntos de mayor detalle. Este hecho no es en sí mismo problemático, pues la discrepancia es consustancial al debate académico. Lo que sí es necesario es que se tenga claro en qué se disiente y por qué razones. Es necesario saber si las discrepancias son de fondo o de forma, si son de contenido o de formulación y, en última instancia, si lo que una disciplina (o un académico dentro de ella) reconoce como evidencia a favor de su posición lo sería también para otra disciplina (o para un académico que trabaje en ella). Así, pues, desde sus inicios, el objetivo del Grupo no residió en llegar a acuerdos temáticos, sino en llegar a acuerdos acerca de los desacuerdos, es decir, en tener claro en qué disintimos, por qué razones y sobre la base de qué evidencias.

Como era fácil de predecir, los intercambios de opiniones fueron cambiando a sus interlocutores. Los miembros del Grupo, progresivamente, fueron comprendiendo un poco mejor las posiciones de los otros y gradualmente estas les fueron pareciendo menos extrañas. Con el paso del tiempo, se fue constituyendo una verdadera comunidad académica que no necesariamente estaba unida por un conjunto

de convicciones compartidas, sino por la curiosidad intelectual acerca de los fenómenos que tienen que ver con la mente y el lenguaje, así como con el sincero deseo de aprender de los demás estando dispuestos a abandonar posiciones previamente sostenidas si el intercambio argumentativo lo mostraba necesario.

En el año 2011, *Mente y Lenguaje* postuló a un fondo económico otorgado por el Vicerrectorado de Investigación de la Pontificia Universidad Católica del Perú (PUCP) para investigaciones interdisciplinarias. El objetivo de la investigación era estudiar, tanto desde un punto de vista filogenético como ontogenético, la manera como los procesos lingüísticos y otros procesos mentales, como la cognición social, interactúan en el desarrollo del individuo e interactuaron en la evolución de la especie para generar los mecanismos que actualmente empleamos y estudiamos. Así mismo, el Grupo se propuso averiguar el rol de la intersubjetividad en la constitución de dichos procesos. Gracias a ese fondo, *Mente y Lenguaje* pudo organizar una serie de conferencias, talleres y congresos interdisciplinarios e internacionales. Uno de ellos fue un coloquio internacional realizado en julio de 2012 en la ciudad de Pisac, en el Valle Sagrado de los Incas. Muchos de los artículos ahora publicados fueron debatidos en ese coloquio.

El resultado de todo ese trabajo es este libro, que está conformado por dos partes principales. En la primera parte, los miembros de *Mente y Lenguaje* elaboraron una investigación conjunta para explicitar las intersecciones entre la psicología, la lingüística y la filosofía, entrecruzando dos ejes: el desarrollo y la evolución, tanto de los procesos de cognición social como del lenguaje. Esta primera parte tiene como finalidad generar un terreno transdisciplinario, de manera que los métodos, evidencias y hallazgos de las tres disciplinas puedan dialogar entre sí, y mostrar un panorama interconectado del estado de la cuestión en las tres disciplinas que nos convocan.

En la segunda parte del libro se publican artículos —en principio independientes— de los miembros del Grupo así como de los investigadores invitados. Cada artículo defiende una tesis en torno al tema del libro, pero los diversos artículos no pretenden formar un conjunto y pueden leerse de manera separada. Esta segunda parte se divide en cuatro secciones: la evolución de la cognición social, el desarrollo y estructura de la misma, la evolución del lenguaje, y el desarrollo y estructura del mismo. El objetivo de esta segunda parte no es ofrecer una teoría unificada, sino varias posturas, a veces complementarias y en ocasiones discrepantes, sobre el tema común.

Los miembros de *Mente y Lenguaje* deseamos agradecer a todos aquellos que han hecho posible la investigación que ha dado origen a este libro. Nuestra gratitud especial va dirigida al Vicerrectorado de Investigación de la PUCP, al Fondo Editorial de esta universidad y a los muchos académicos invitados que nos han visitado a lo largo de estos años, varios de los cuales han aceptado generosamente

publicar sus artículos en este libro. Entre muchos otros, agradecemos especialmente a Alberto Cordero Lecca, Alejandro Tomasini, Ángeles Eraña, Antoni Gomila, Axel Barceló, Catherine Wearing, Clark Barrett, Derek Bickerton, Diana Pérez, Eduardo García-Ramírez, Francisco Pereira, Isidora Stojanovic, Jesse Prinz, Julio Torres, Liza Skidelsky, Marcelo Sabatés, Manuel Rodríguez, Mario Montalbetti, Max Freund, Michael Devitt, Mikhail Kissine, Patricia Brunsteins, Peter Carruthers, Robert Stainton, Robin Dunbar y Silvia Español.

Los editores queremos agradecer muy especialmente a Andrés Abugattás, Carolina Romero y Luis Manuel Olguín, miembros de Mente y Lenguaje, quienes han tenido sobre sus hombros algunas de las partes más complicadas, delicadas y difíciles de la organización de las diversas actividades que han hecho posible este libro. También deseamos dar las gracias a los colegas, alumnos y amigos que han participado en nuestras actividades y que, con sus preguntas y objeciones, nos han ayudado a perfilar nuestras posiciones. Muy especialmente queremos agradecer a nuestras respectivas familias por las muchísimas noches en que estuvimos dedicados a la discusión, dando forma a las ideas que aquí publicamos. El resultado de todas esas horas de debate es este libro, pero también el hecho de que ahora cada uno de nosotros tiene una visión más completa y rica de su tema de investigación, así como nuevas interrogantes e inquietudes.

Así pues, el objetivo del libro es integrar, de manera inter y transdisciplinaria, los estados de la cuestión de tres disciplinas (filosofía, lingüística y psicología) que abordan un mismo tema con vocabularios y metodologías diferentes: el rol de la intersubjetividad en la evolución y el desarrollo de la cognición social y el lenguaje. Este trabajo no se ha hecho en la actualidad o se ha hecho solo de manera muy limitada. Asimismo, la mayoría de los artículos que conforman la segunda parte del libro se proponen hacer contribuciones específicas en este terreno transdisciplinario, no solo con el objetivo de permitir a otros investigadores continuar la tarea de generar un espacio transdisciplinario sino también con el propósito de fortalecer y ampliar las fronteras de estas disciplinas.

Los editores

INTRODUCCIÓN

Mente y Lenguaje
Grupo Interdisciplinario de Investigación

Andrés Abugattás, Ricardo Braun, Paola Cépeda, César Escajadillo,
María de los Ángeles Fernández Flecha, Marcos Herrera Burstein,
Carla Mantilla, Luis Manuel Olguín, Jorge Pérez Silva,
Pablo Quintanilla y Carolina Romero

Las primeras discusiones de que se tiene noticia en la tradición occidental acerca de la naturaleza de lo mental comenzaron con el diálogo *Fedón*, de Platón, y el tratado *Peri psychés*, de Aristóteles. Es más que probable que estas primeras reflexiones que nos han llegado por la vía escrita estuvieran influidas por tradiciones orientales y pre-socráticas de las que solo tenemos fragmentos (Kirk & Raven, 1981; Rohde, 1948; Snell, 1982 y Bremmer, 1983). Es interesante, sin embargo, que ambos textos —el *Fedón* y el *Peri psychés*— ya establecieran dos tipos de acercamiento diferente, e inclusive podría decirse dos agendas distintas respecto de los estudios acerca de lo que hoy llamaríamos la vida de la mente. Por una parte, el *Fedón* defiende una posición dualista en la que lo mental es ontológicamente diferente y separable del cuerpo físico, objetando no solo todas las posiciones monistas y fisicalistas sino también las que después se llamarían epifenomenalistas y emergentistas. Aristóteles, por otra parte, en el *Peri psychés*, está más cerca de una posición monista en la que los conceptos de *hylé* y *morphé* (lo que traduciríamos problemáticamente por ‘materia’ y ‘forma’ dentro del modelo hylemórfico aristotélico) podrían ser considerados dos aspectos o dos descripciones de una misma sustancia ontológicamente inseparable. Es decir, la posición de Aristóteles podría ser vista como una versión de lo que hoy se llama monismo de aspecto dual, una posición compatible con el naturalismo.

Es factible decir que, a partir de esos dos textos fundacionales, la reflexión filosófica sobre lo mental ha oscilado en versiones y combinaciones de esas dos posiciones, por lo menos a lo largo de la Edad Media y la Modernidad. Mientras en el Medioevo hubo una tendencia a identificar vida psíquica con alma, durante la Modernidad se separaron esos dos conceptos, reservando los términos *alma* y *espíritu*

para una dimensión religiosa en la que existe un elemento inmortal en el ser humano, en tanto los conceptos de *mente*, *vida psíquica* y *subjetividad* aludirían a procesos de tipo cognitivo y afectivo que tendrían un sustrato físico. Es interesante notar que si uno pasa revista a los títulos de las obras más importantes de los filósofos modernos —desde Descartes pasando por todos los empiristas británicos y Kant hasta llegar a Hegel— la gran mayoría de ellos expresan un deseo de entender nuestra vida mental —individual y colectiva— ya sea tratando de aclarar el método, como podemos llegar a hacerlo; inquiriendo sobre la naturaleza de las ideas, el lenguaje, el entendimiento, la imaginación, los afectos, las pasiones, los sentimientos morales, la razón pura, la experiencia y el espíritu; o problematizando la manera en que conocemos los contenidos de las otras mentes y las nuestras. Todos esos temas y las preguntas que se formularon los filósofos modernos podrían ser englobados dentro del enorme campo de estudio que hoy llamaríamos ciencias cognitivas.

Sin embargo, recién a fines del siglo XIX y a comienzos del XX estas reflexiones dejaron de ser puramente especulativas y adquirieron bases empíricas con la aparición de la psicología experimental, la lingüística y los todavía incipientes estudios médicos en neurología. Estos fenómenos académicos fueron sugiriendo dos cosas de gran importancia.

Por una parte, empezó a quedar bastante claro que si bien los filósofos clásicos tuvieron que afrontar estas cuestiones de una manera estrictamente teórica y especulativa, porque no tenían ningún tipo de herramienta efectiva que pudiera serles de ayuda, con la aparición de los primeros resultados empíricamente verificables sobre el funcionamiento del cerebro, la cognición, los afectos, y la adquisición, el uso y funcionamiento del lenguaje, un enfoque únicamente conceptual sería sumamente limitado y reduccionista. Esto no significó que los filósofos debían abandonar sus prácticas y metodologías para cambiarlas por las pruebas clínicas y los laboratorios, sino que tenían que comenzar a prestar más atención a los resultados empíricos de las otras disciplinas para dar mayor fundamento y amplitud a sus posiciones. De la misma manera, las investigaciones empíricas hicieron notar que los experimentos, las pruebas clínicas y los resultados observacionales con frecuencia subdeterminan las hipótesis y teorías, es decir, pueden corroborar más de una teoría incompatible, pues requieren de un análisis conceptual de gran finura para poder establecer qué es lo que realmente están sugiriendo y con qué grado de confiabilidad. En otras palabras, así como un enfoque únicamente teórico y conceptual resultaría insuficiente, la observación empírica y la experimentación científica deben ir acompañadas de un complejo y detallado análisis conceptual, seguramente de factura filosófica, que permita dar claridad a los resultados que se van obteniendo e interprete los diversos logros integrándolos entre sí y mostrando sus limitaciones y potencialidades.

Esta mirada filosófica no solo debe ser metodológica, holística y conceptual sino también histórica, es decir, debe permitirnos ver los resultados a la luz del trayecto que venimos recorriendo desde hace varios miles de años. Así pues, para usar una vez más la repetida metáfora kantiana, el análisis conceptual sin fundamento empírico es vacío y los estudios empíricos sin claridad y norte conceptual son ciegos.

El otro punto que comenzó a quedar claro, ya desde principios del siglo XX, es que el universo de problemas y temas que iba surgiendo no podía limitarse a una sola ciencia o disciplina, sino que debía involucrar varias, no solo en un sentido multidisciplinar ni interdisciplinar sino, sobre todo, transdisciplinar. Es decir, no se trata solo de mirar el mismo objeto —por ejemplo, la cognición social o el lenguaje— desde varias disciplinas, ni tampoco de hacer que los diversos especialistas conversen entre ellos sino también de que vayan surgiendo espacios nuevos, transdisciplinarios, que ya no pertenezcan solo a una de las disciplinas clásicas sino, más bien, a ninguna en particular, porque la evidencia y metodología que resultan necesarias para afrontar exitosamente estos nuevos campos de estudio proceden de varias ciencias diferentes.

Así fue cómo, hacia mediados del siglo XX, comenzó a surgir lo que ahora se llama ciencia cognitiva. De manera institucional, a mediados de los años setenta se fundó la *Cognitive Science Society* y la revista *Cognitive Science*, cuyo primer número se publicó en 1976 y generó una revolución intelectual que ha impactado a instituciones académicas de todas partes del mundo. Actualmente son más de 70 las universidades del mundo que tienen programas académicos, de diversos niveles, sobre ciencia cognitiva y es muy probable que para cuando el lector tenga este libro en sus manos el número sea mucho mayor. Sin embargo, a pesar del auge que esta revolución tiene, la ciencia cognitiva todavía está en su infancia y es muy difícil lograr resultados conjuntos.

Aunque la mayor parte de centros académicos de estudio en ciencia cognitiva se encuentran en los países del primer mundo, hay, en Latinoamérica, investigaciones de gran calidad en disciplinas como filosofía, lingüística, psicología y neurociencias. Lo que aún no existe de manera suficientemente desarrollada en nuestro medio es una adecuada integración entre estas disciplinas. Eso se debe por lo menos a dos razones. De un lado, hay pocas universidades latinoamericanas con los recursos económicos y tecnológicos necesarios para la investigación en ciencia cognitiva. De otro lado, no es en absoluto fácil reunir a un grupo de académicos de diversas disciplinas, acostumbrados a sus propias metodologías, vocabularios y tradiciones, y hacerlos tratar de pensar juntos un tema, tomando en consideración las prácticas ajenas. Es solo en el permanente intercambio intelectual que, casi por ósmosis, los investigadores van aprendiendo de las prácticas de las otras disciplinas y los resultados del conjunto se van viendo muy poco a poco y después de periodos largos de trabajo conjunto.

En el Perú, prácticamente no hay una tradición de investigación en ciencia cognitiva y la única institución existente que se ha trazado como objetivo este tipo de estudio transdisciplinar es el Grupo Interdisciplinario de Investigación Mente y Lenguaje, que, con este libro, entrega a la comunidad académica su primer producto publicado; un libro sobre las interacciones entre el lenguaje y la cognición social, tanto desde el punto de vista ontogenético como filogenético.

Las investigaciones acerca de la evolución filogenética y el desarrollo ontogenético de la mente y del lenguaje en la especie y en el niño, respectivamente, han avanzado notablemente en extensión y profundidad en los últimos años, especialmente en las siguientes disciplinas: filosofía de la mente y del lenguaje, lingüística, psicología del desarrollo y psicoanálisis. En muchos casos, las investigaciones se han superpuesto entre sí, abordando los mismos o parecidos problemas con instrumentos conceptuales o vocabularios diferentes, lo que ha generado la necesidad de un mayor diálogo interdisciplinario. En otros casos han aparecido territorios inexplorados que ya no pertenecen a ninguna de las ciencias tradicionales. Surge, entonces, la necesidad de enfoques transdisciplinarios. Como ya se ha mencionado, mientras la interdisciplinarietà alude al hecho de que varias ciencias distintas estudian el mismo fenómeno, la transdisciplinarietà tiene que ver con un terreno nuevo —o un conjunto de preguntas originales— que surge en las fronteras de las ciencias clásicas, de manera que no pertenecen a una en particular, por lo que deviene necesario que distintas ciencias cooperen para estudiar algo que no ha sido estudiado por ninguna de ellas previamente.

Se observa, sin embargo, que las investigaciones en las disciplinas mencionadas acerca de la evolución y el desarrollo de la mente y del lenguaje han avanzado, por lo general, de manera independiente y no se han integrado ni fertilizado mutuamente. No lo suficiente, en cualquier caso. Tampoco ha habido acumulación de conocimiento. Integración, fertilización y acumulación son tres condiciones necesarias para el avance y el progreso de la ciencia.

La *integración* alude a la necesidad de que la información obtenida en distintos campos del conocimiento sea compartida y evaluada con los criterios de las diversas ciencias, de manera que, si hay incompatibilidad entre los datos provenientes de las distintas disciplinas, se pueda discriminar los que están mejor justificados de los que tienen débiles justificaciones. Adicionalmente, se espera una interpretación unificada de tales datos, de manera que se pueda construir una explicación más comprensiva que permita un avance combinado de las ciencias involucradas.

La *fertilización* tiene que ver con que los datos, interpretaciones y explicaciones de una disciplina en particular, al ser conocidos por investigadores de otras ciencias, generen en ellos intuiciones innovadoras que sean el producto de una reestructuración de la información.

La *acumulación* refiere al hecho de que la información y las explicaciones obtenidas por distintas ciencias son estructuradas, conformando una suerte de escalera explicativa en la que podemos notar que las explicaciones más amplias y comprensivas contienen a las más específicas, mientras que lo contrario no ocurre.

Tanto la primera parte de este libro como los artículos especializados de la segunda se han propuesto hacer un trabajo al mismo tiempo interdisciplinario y transdisciplinario, con el objetivo de lograr una mayor integración, fertilización y acumulación del conocimiento producido por las diversas ciencias involucradas, de manera que se puedan generar las condiciones apropiadas para el avance y el progreso del conocimiento en el tema que nos ocupa.

Este libro no es el producto de un trabajo de campo ni de la aplicación de un diseño metodológico empírico, sino de la sistematización, análisis, integración y discusión de la información y la evidencia empírica reciente en las disciplinas relevantes.

El libro está dividido en dos partes. La primera, que hemos denominado «Marco conceptual», ofrece un recuento de las investigaciones realizadas por académicos que, con diversa metodología, han tratado la evolución y el desarrollo de la cognición social y el lenguaje desde diferentes disciplinas científicas. En este sentido, hemos dividido esta parte del libro en dos capítulos que se ocupan, respectivamente, de cada una de estas capacidades.

El primer capítulo aborda la evolución en la especie humana y el desarrollo en el niño de la cognición social. De esta manera, se inicia con una revisión de la teoría darwiniana de la selección natural y la evolución de las facultades sociales humanas para encontrar en ella las bases de la teoría de la evolución actual y acercarnos, con más detalle, al desarrollo filogenético del cerebro social y la atribución psicológica en el desarrollo social primate. Sobre esta base, que hace explícita la complejidad de la organización social humana, el capítulo examina la hipótesis de la denominada *inteligencia social* y las propuestas consecuentes en torno de la evolución de la lectura de mentes, la autoconciencia y la agencia. A continuación, se aborda el desarrollo de las capacidades sociales en el niño desde el punto de vista de la psicología evolucionista. De esta manera, se presentan los hitos ontogenéticos relevantes para la cognición social. La presentación se centra, hacia el final, en la capacidad de *mentalización*, entendida como la capacidad humana para poder representarse los estados mentales ajenos, en tanto cumple una función crucial para el desarrollo óptimo del individuo en sociedad.

El segundo capítulo de esta primera parte se dedica a la evolución en la especie y el desarrollo, en el niño, del lenguaje. El capítulo comienza buscando definir el lenguaje a partir de la revisión del concepto de lengua utilizado en la lingüística moderna

y la descripción de la diversificación de las lenguas en el mundo a raíz del denominado *cambio lingüístico*. Para tratar la evolución del lenguaje y el desarrollo lingüístico en el niño se define una lengua como un tipo de saber estructurado en cinco componentes: la fonología, el léxico, la gramática morfológica y sintáctica, la semántica y la pragmática. A continuación, representa el panorama de debate actual en relación con la evolución del lenguaje. Se distingue, así, entre las principales propuestas en torno del origen y la evolución del lenguaje como sistema formal de representación y aquellas en torno de sus ventajas adaptativas para la comunicación y la cooperación. La presentación sobre el desarrollo del lenguaje parte por señalar los hitos principales en la adquisición de una lengua por parte del niño. Esto nos ofrece un panorama del desarrollo lingüístico infantil que es explicado desde diferentes modelos del proceso de adquisición que divergen entre sí por su descripción de la relación entre el lenguaje y otros sistemas de cognición humana. El capítulo pasa revista de tres modelos de adquisición del lenguaje, actualmente considerados como los más importantes: el innatista, el sociopragmático y el conexionista. Tanto las referencias de este capítulo como las del anterior se presentan unificadas al final de esta parte del libro.

La segunda parte de este estudio se propone profundizar en aspectos específicos del tema tratado, por lo que está conformada por un conjunto de investigaciones puntuales acerca del rol que cumple la intersubjetividad en la evolución y el desarrollo de la mente y del lenguaje. Las diversas investigaciones abordan aspectos puntuales de este tema, pero todas lo hacen con una vocación y un punto de partida inter y transdisciplinario. Cada una de estas investigaciones revisan y proponen tesis que tienen por objetivo ampliar las fronteras disciplinarias e interdisciplinarias del conocimiento acerca del tema planteado, de manera que, en su conjunto, se proponen presentar un modelo más completo del rol de la intersubjetividad en la evolución y el desarrollo de la mente y el lenguaje.

Dada la complejidad de los temas estudiados, es natural que los diversos autores tengan discrepancias conceptuales y que, en ocasiones, también usen los términos de maneras diferentes. Sin embargo, en lo que sigue definiremos de forma general algunos conceptos clave que son usados por la mayoría de los autores y que son el foco de atención conjunta en las investigaciones que conforman este libro: intersubjetividad, mente, lenguaje y cognición social.

Por *intersubjetividad* se suele entender la interacción entre individuos en relación con un mundo real compartido, que es tanto físico como social y cultural. Tal como se suele concebir, entonces, la intersubjetividad es un fenómeno básicamente triangular, que involucra a un yo, a un otro y a una realidad objetiva que ambos asumen compartir constituyendo sus diversas subjetividades mutuamente (Davidson, 2001 y Cavell, 2006).

Por *mente* se entiende normalmente la totalidad de procesos psíquicos, también llamados *estados mentales*, que tienen, por lo menos, una de estas tres características: a) son estados conscientes (actual o potencialmente), b) son estados intencionales y c) son estados conformados, o causados, por un conjunto de reglas cognitivas que constituyen un tipo de conocimiento implícito o tácito como, por ejemplo, las reglas sintácticas del lenguaje. Aunque no todos los autores estarían de acuerdo en incluir estos estados como mentales, dado que la extensión del concepto de *mente* es fundamentalmente estipulativo, incluir también esta propiedad permite reconocer como estados mentales a los procesos cognitivos no conscientes que, sin embargo, constituyen formas de conocimiento implícito. Todo estado mental es también un estado físico, aunque no todo estado físico es siempre un estado mental, de suerte que se cumple, por lo menos, una de estas tres condiciones:

- a) *Conciencia*. Para que un estado físico también pueda ser descrito como mental es necesario que su portador o bien sea consciente de él o pueda llegar a ser consciente de él, en el sentido de que debe tener las experiencias fenoménicas (o *qualia*) de lo que significa estar en ese estado mental. Para ello basta con la existencia de conciencia nuclear y no es necesaria la conciencia extendida o autobiográfica, es decir, basta con la pura experiencia fenoménica y no es indispensable la autoconciencia, también llamada conciencia reflexiva o metacognición.
- b) *Intencionalidad*. Un estado mental es intencional cuando está dirigido a un objeto diferente de sí mismo. Este fenómeno también es denominado *aboutness*, es decir, la propiedad que tiene algo de ser, versar o referir a un objeto diferente de él mismo.
- c) *Conocimiento implícito*. Para que un estado físico pueda, también, ser descrito como mental es necesario que su portador posea la habilidad de emplear un conjunto de reglas cognitivas de transformación que permitan producir nuevo conocimiento a partir de información previa.

Así, por ejemplo, una sensación o un dolor tienen la propiedad a pero no la propiedad b. Una creencia inconsciente tiene la propiedad b pero no la propiedad a. Una creencia consciente y una emoción tienen las propiedades a y b. Las operaciones lógicas (por ejemplo, el *modus ponens*) y las reglas sintácticas (por ejemplo, el *parsing*) tienen la propiedad c, aunque no necesariamente las dos anteriores.

En suma, para que un estado físico también pueda ser descrito como un estado mental debe tener al menos una de las tres propiedades mencionadas más arriba.

Así, la palabra *mente* referirá al conjunto de funciones conscientes (o potencialmente conscientes) de un cerebro y/o al conjunto de sus funciones intencionales y/o al conjunto de sus funciones de procesamiento cognitivo. De esta manera, si bien toda mente requiere de un cerebro, no todo cerebro genera una mente. Aunque es lógicamente posible que pudiera surgir una mente, en el sentido descrito, en un soporte no biológico, este no es un tema discutido en este libro. Conciencia e intencionalidad son propiedades emergentes del cerebro o, incluso más ampliamente, de un organismo biológico. Esto significa que, a partir de la complejidad de un conjunto de propiedades físicas que tiene el cerebro (o el organismo en su conjunto) se generan propiedades de segundo orden que no son físicas ni se pueden reducir a propiedades físicas. La conciencia y la intencionalidad serían ejemplos de estas propiedades emergentes de segundo orden.

Por *lenguaje* suele entenderse una característica/propiedad específicamente humana y universal que consiste en la capacidad de desarrollar un *saber hacer* o un *saber representacional*, parcialmente compartido por una comunidad de hablantes, que permite producir e interpretar enunciados. Este saber está compuesto, básicamente, por signos lingüísticos (unidades) y una gramática (reglas) para su combinación, lo que hace posible la producción e interpretación de infinitos enunciados y, de este modo, la transmisión de infinitos sentidos o mensajes distintos. El lenguaje sería tanto una herramienta para la comunicación (interacción entre individuos) como para el pensamiento (representación); diversas posturas teóricas enfatizan uno y otro de los dos puntos extremos en este espectro. Asimismo, distintas teorías caracterizan de forma particular esta capacidad que tenemos los seres humanos para desarrollar saberes lingüísticos de forma natural en la infancia. De este modo, desde ciertas posturas, esta facultad del lenguaje es concebida como innata, de dominio específico y, en esa medida, con un diseño específico para el procesamiento lingüístico. Desde otras posturas, en cambio, esta capacidad es entendida como producto de la evolución cultural humana, más bien de dominio general.

La *cognición social* alude a la capacidad del cerebro o de la mente humana para procesar la información social y, consecuentemente, poder interactuar de manera exitosa y sobrevivir en entornos sociales. La información social que la cognición social debe procesar tiene que ver con las relaciones de cooperación y competencia, los grupos de poder y las relaciones jerárquicas, las alianzas estratégicas y los diversos ordenamientos sociales. Estas habilidades de procesamiento de información social permiten que el individuo atribuya estados mentales a otros individuos, o a grupos de individuos, con la finalidad de predecir sus acciones, cooperar, generar alianzas ventajosas, pero también para adelantarse a posibles formas de comportamiento agresivo, intuir engaños, manipular, engañar y contraengañar. Asimismo, estas habilidades

requieren de varios niveles de intencionalidad: desde uno en primer grado, en que el individuo tiene estados mentales acerca del mundo; pasando por un segundo grado, en que el individuo tiene estados mentales acerca de los estados mentales de otro; hasta el cuarto o quinto grado de intencionalidad, que caracterizan el nivel estándar en nuestra especie.

En los humanos las creencias en primer grado suelen aparecer hacia los dos años, las creencias en segundo grado hacia los tres años (cuando los niños están en condiciones de aprobar el test de la falsa creencia, esto es, cuando pueden atribuir creencias falsas a los demás) y, hacia los cinco años tienen creencias en tercer grado o más. Los seres humanos adultos normales suelen llegar a tener creencias hasta de cuarto o quinto grado de intencionalidad. Aparentemente hay una correlación entre los grados de intencionalidad que pueden alcanzar los primates, el aumento del volumen del lóbulo frontal y el número de individuos con los cuales uno tiene que interactuar exitosamente, ya sea para cooperar o para competir. Todo indica, entonces, que la aparición de atribución psicológica en varios niveles de intencionalidad, así como el surgimiento de complejos estados mentales en los primates, sería consecuencia de la necesidad de interactuar en grupos numerosos de individuos con complejas relaciones sociales (Dunbar, 1998).

Así, la cognición social, esto es, las habilidades de procesamiento de información social, es fundamental para la supervivencia de los animales sociales y, de hecho, puede encontrarse en todas las especies sociales. Es, sin embargo, en los primates y, específicamente, en los homínidos, en los que alcanza niveles de mayor sofisticación. Los diversos elementos que conforman la cognición social deben haber sido escogidos por la selección natural por su alto valor adaptativo en un entorno que no solo era físico sino también social. Así, a partir del reconocimiento del comportamiento ajeno (por ejemplo, competitivo o cooperativo) fue surgiendo progresivamente la capacidad de la atribución psicológica en sus dos formas: lectura de mentes (atribución de estados mentales a otros individuos) y metacognición (atribución de estados mentales a uno mismo). Lo más probable es que *lectura de mentes*, *metacognición* y *concepción de realidad objetiva*, es decir, el conocimiento de los estados mentales ajenos, el conocimiento de los propios estados mentales y el conocimiento de los eventos físicos del entorno compartido, surgieran simultáneamente y fuesen seleccionados por las mismas razones. Parece difícil suponer que se diera alguna de estas tres formas de conocimiento sin las otras dos. Es importante subrayar, sin embargo, que al hablar de estas formas de conocimiento no solo estamos asumiendo la capacidad de reconocer y atribuir contenidos proposicionales sino, además, la capacidad de distinguir entre lo real y lo aparente, lo verdadero y lo falso, lo objetivo y lo subjetivo. Es evidente que muchas especies de animales pueden detectar formas

de comportamiento en otros individuos y «predecir» su comportamiento futuro, pero eso no contaría como atribución psicológica ni como lectura de mentes, dado que no habría una teoría de la mente en juego, es decir, una representación acerca de la mente del otro, y que el animal en cuestión estaría reaccionando a formas de comportamiento, no al reconocimiento de estados mentales. Así, tiene sentido decir que un individuo atribuye estados mentales a otro, o tiene la capacidad de lectura de mentes, si ese individuo hace una atribución de estado mental que, considera, puede ser diferente del suyo y que, además, asume que podría ser verdadera o falsa, es decir, si el individuo que interpreta reconoce que su interpretación podría ser errada. Análogamente tiene sentido hablar de autoconocimiento solo si el individuo que se conoce tiene la habilidad de atribuirse estados mentales a sí mismo, admitiendo que podría errar en esa autoadscripción. En ambos casos, en la lectura de mentes y en el autoconocimiento, debe haber por lo menos dos niveles de intencionalidad en juego. Finalmente, solo tiene sentido hablar de conocimiento acerca del mundo exterior si el individuo que conoce está en capacidad de distinguir entre realidad y apariencia y, por tanto, admite la posibilidad de errar en sus creencias acerca de la realidad objetiva. Así, la intersubjetividad, esto es, la interacción con otros individuos en contextos sociales complejos, ha sido el elemento fundamental que ha determinado que la selección natural haya ido escogiendo, progresivamente, los diversos elementos que llegaron a conformar la cognición social, así como los diversos grados de intencionalidad que caracterizan a la cognición social humana.

Como ya se ha mencionado, la segunda parte del libro reúne dieciséis artículos de investigadores provenientes de la psicología, el psicoanálisis, la lingüística, la filosofía de la mente y el lenguaje, y la antropología evolucionista. Estas contribuciones se encuentran ordenadas en cuatro secciones que se desprenden del tema general de la investigación: a) evolución de la cognición social, b) desarrollo y estructura de la cognición social, c) evolución del lenguaje y d) desarrollo y estructura del lenguaje.

La primera sección inicia con el artículo de Clark Barrett «The nonlinear evolution of human cognition». En él, el autor aborda uno de los problemas teóricos cruciales de la evolución de la cognición, a saber, la pregunta sobre cómo han evolucionado las adaptaciones psicológicas en una mente que está compuesta por múltiples adaptaciones que interactúan entre sí. El artículo se propone, en primer lugar, describir la naturaleza del problema; posteriormente, pretende explicar cómo este debe ser abordado para poder comprender la evolución de la mente, y finalmente, concluye con algunas ideas sobre cómo se puede solucionar dicho problema. El autor sugiere que dejemos de pensar en las adaptaciones psicológicas en términos de propiedades, tales como rigidez, autonomía e innatismo, y propone concebirlas en términos de diseño evolucionado, es decir, en términos de cómo el proceso evolutivo

ha moldeado la adaptación para que interactúe con el resto del sistema en el cual se encuentra, incluyendo el cerebro y el mundo.

En el artículo «The fragmentation of reasoning», Peter Carruthers evalúa la validez científica de una distinción frecuentemente trazada entre los científicos cognitivos, a saber, entre *Sistema 1* (S1) y *Sistema 2* (S2). Como es conocido, S1 alude a los mecanismos cognitivos no conscientes y automatizados, mientras que S2 refiere a mecanismos cognitivos conscientes y voluntarios. Si bien el autor sostiene que existen aspectos válidos en esta distinción, ya que existe una diferencia real entre procesos cognitivos intuitivos y reflexivos, termina por argumentar que la distinción tal cual no es consistente y que no está alineada con muchas otras propiedades comúnmente atribuidas a S1 y S2. En consecuencia, afirma que esta distinción no es real y debe ser abandonada.

En «La brecha en mente o por qué los humanos no son solo grandes simios», Robin Dunbar señala que, aunque compartimos muchos aspectos de nuestro comportamiento y nuestra biología con nuestros primos primates, los humanos somos diferentes en un sentido crucial: nuestra capacidad para vivir en el mundo de la imaginación. Esto se refleja en dos aspectos centrales de nuestro comportamiento que son, en muchos sentidos, arquetípicos de lo que es ser un humano: la dimensión religiosa y la capacidad para narrar historias. El autor se propone mostrar cómo estos rasgos notables han emergido del desarrollo natural del cerebro social. También muestra cómo la naturaleza subyacente de nuestra cognición social, como primates y en las sociedades humanas, forzaron a nuestro cerebro a expandirse en tamaño durante el curso de la evolución, desde los últimos cinco millones de años.

Claudia Muñoz Tobar, en «La naturalización de los conceptos morales. Sobre la tesis de la imaginación moral de Mark Johnson», elabora una síntesis de la epistemología de la moral de Johnson, quien propone la tesis de la corporeización del significado, que integra, entre otros, aportes de la filosofía del lenguaje, la filosofía de la mente, la fenomenología, la lingüística cognitiva y la psicología cognitiva. La autora basa su exposición en dos libros, *El cuerpo en la mente* (1991) y *Moral Imagination* (1993). En el primero, Johnson presenta su proyecto de una teoría de la imaginación que complete la consideración kantiana de las funciones imaginativas por medio de la extensión de estas al ámbito de la razón práctica; además, en él se introducen las nociones de «esquema de imagen» y «proyección figurativa». En el segundo, este autor desarrolla más exhaustivamente su proyecto de una teoría de la imaginación en la forma de una epistemología de la moral.

Finaliza esta sección el artículo de Pablo Quintanilla, que aborda la pregunta sobre la conexión lógica, causal y genealógica entre el conocimiento de los estados mentales ajenos (lectura de mentes) y el de los estados mentales propios (metacognición).

El artículo argumenta, desde el punto de vista filogenético, en contra de las propuestas que sostienen que hay anterioridad lógica y causal de uno de estos tipos de conocimiento, como en el caso de Descartes y Goldman (2006) que privilegian la autoconciencia o de Carruthers (2009) que sostiene la anterioridad lógica y causal, pero no temporal, de la lectura de mentes. La tesis que el texto defiende es que estos tipos de conocimiento son lógicos, causales y genealógicamente inseparables, de manera que no se puede dar ni concebir uno de ellos sin los otros. También sostiene que la evidencia empírica prueba que estos tipos de conocimiento se dan juntos en la evolución de la especie y en el desarrollo del individuo, potenciándose mutuamente. De esta manera, la presión del entorno habría seleccionado, al mismo tiempo y por las mismas causas, a los individuos con más habilidades metarrepresentacionales y metacognitivas.

La segunda sección de los artículos, sobre el desarrollo y estructura de la cognición social, se inicia con un texto sobre la conciencia humana, en tanto componente esencial para la comprensión de los actos mentales y la formación de la llamada teoría de la mente. En «El emergentismo y la aparición de la conciencia», Ricardo Braun analiza la problemática del mismo concepto de conciencia, admitiendo que se pueden establecer varios sentidos de él. Luego sostiene que desde un punto de vista evolutivo la aparición de la conciencia se debería a un proceso emergente, que determina cualidades irreducibles, tal como se da en otros procesos de esa naturaleza. Finalmente, discute dos propuestas recientes acerca del emergentismo de la conciencia: la tesis de David Chalmers (1995, 1996 y 2006) y la de John Searle (1992 y 1997). El emergentismo, concluye Braun, como teoría filosófica de largo recorrido histórico, puede proporcionar los elementos teóricos para conceptualizar mejor el fenómeno de la conciencia humana.

El artículo de César Escajadillo, en «Cómo ser naturalista sin ser reduccionista en el estudio de la mente y el lenguaje: intersubjetividad, comprensión lingüística y atribución psicológica», aborda el papel de la intersubjetividad, específicamente, el papel que desempeña la comunicación interpersonal en el estudio de la comprensión lingüística y la atribución psicológica. La discusión tiene como marco el problema de cómo entender la relación entre el nivel personal y subpersonal de explicación —aquel que corresponde a la explicación psicológica de sentido común, basada en la atribución de creencias y deseos, y aquel que corresponde a la explicación en términos de estados, mecanismos y procesos de carácter no consciente—. Se consideran dos respuestas a este problema, cada una de las cuales representa un modo particular de entender el *naturalismo*: el modelo reduccionista y el no reduccionista. Su propuesta es que, de estos dos, solo el segundo hace justicia al hecho de que la comprensión que tenemos de otros agentes conlleva situar sus acciones —lo que hacen y dicen—

sobre el trasfondo de una visión de la realidad que ha de ser manifiesta para dos o más individuos, condición que se cumple en la comunicación, al interactuar simultáneamente con otros individuos y la realidad compartida. La conclusión que extrae es que el intento de explicar algunas propiedades de la mente y el lenguaje en términos de relaciones verticales interniveles resulta inviable, lo que no conlleva renunciar al naturalismo en el estudio de la mente y el lenguaje.

En su artículo «La sistematicidad de la cognición humana en cuestión», Antoni Gomila discute la idea predominante en el paradigma cognitivista (tanto clásico como conexionista) de que la sistematicidad del pensamiento es anterior al lenguaje. El artículo pasa revista a un número de experimentos recientes en el estudio de la cognición que usa como evidencia para mostrar que no hay sistematicidad cognitiva antes de que aparezca el lenguaje en el desarrollo del niño: la adquisición del léxico facilita la percepción, la explosión sintáctica (que ocurre durante el tercer año de vida) permite un sinnúmero de tareas cognitivas sistemáticas que no están disponibles para el infante sino en forma fragmentaria. Igualmente, se aduce que la evidencia empírica (por ejemplo, con los chimpancés) confirma que los seres no verbales carecen de una sistematicidad cognitiva. El autor entiende que para poder hablar de sistematicidad diferentes habilidades cognitivas deben estar nómicamente conectadas, es decir, debe existir una conexión necesaria entre los diferentes contenidos que un sistema puede manejar. En ese sentido, los seres no verbales tendrían habilidades cognitivas que pueden ser vistas como independientes (es decir, no conectadas, no sistemáticas). Dado que los seres humanos son diferentes en la medida en que tienen lenguaje, Gomila concluye, la fuente de la sistematicidad este lenguaje y no el pensamiento. Así, los animales tendrían una cognición especializada, con habilidades no conectadas, mientras los seres humanos tenemos una cognición de propósito general, con habilidades sistematizadas gracias al lenguaje.

En «El vínculo de apego como escenario para el desarrollo de la cognición social temprana» Carla Mantilla integra aportes recientes de la psicología evolutiva, el psicoanálisis y la filosofía de la mente, con el fin de subrayar el contexto relacional en que se desarrolla y articula el autoconocimiento, el conocimiento de las otras mentes y del entorno, lo que se conoce como *triangulación*. La autora propone una línea de desarrollo que contempla una primera forma de triangulación preverbal, implícita y no reflexiva, que se hace visible entre los nueve meses y el primer año de vida, momento en el cual el infante es capaz de atender sostenidamente a un foco atencional junto a otra persona. Una segunda forma de triangulación —verbal, explícita y reflexiva— comenzaría a desplegarse hacia los tres años y se cristalizaría hacia los cinco con la adquisición de la mentalización o capacidad para percibir, imaginar e interpretar las acciones propias y las de los demás a partir de atribuir estados mentales (Fonagy, 2008).

El vínculo afectivo más importante de la infancia o *vínculo de apego*, sería, de acuerdo con la evidencia empírica revisada, el escenario relacional encargado de viabilizar esta secuencia de desarrollo, central para la cognición social.

El artículo que finaliza esta sección, «Intersubjetividad y atribución psicológica», que publican Diana Pérez y Silvia Español, explora el tránsito desde los modos no proposicionales de atribución psicológica, presentes en las formas básicas de la intersubjetividad —la primaria y secundaria—, hacia los modos proposicionales de atribución psicológica que aparecen con la intersubjetividad terciaria, aproximadamente hacia el cuarto año de vida. Adoptando el marco explicativo que ofrece la *perspectiva de segunda persona*— la comprensión mediada por la atribución recíproca de estados mentales en contextos interactivos cara a cara—, las autoras defienden la tesis de que no todas las atribuciones de creencias y deseos involucran contenidos proposicionales, esto a contracorriente de la tradición que las entiende como atribuciones de estados con contenidos semánticamente evaluables. Para ello, analizan el modo en que el juego funcional y el ficticio dan cuenta de la aparición de metarrepresentaciones y, por consiguiente, de la intersubjetividad terciaria.

La sección sobre evolución del lenguaje inicia con el artículo de Paola Céspedes y Gabriel Martínez Vera, titulado «Perspectivas sobre el origen y evolución del componente sintáctico». En él los autores intentan, a partir de la revisión de evidencia lingüística y paleoantropológica, establecer los límites temporales en relación con dos momentos en los que se originaría y evolucionaría la sintaxis. El artículo se enmarca, principalmente, en las propuestas de Derek Bickerton (2009), así como de Marc D. Hauser, Noam Chomsky y W. Tecumseh Fitch (2002), las que concuerdan con los 200 mil y 50 mil años que los autores proponen como los límites más lejano y más cercano, respectivamente, del origen del lenguaje.

En el artículo de Paola Céspedes, «*Homo combinans*. Explorando la Gramática Universal minimista», la autora se propone indagar en aquellas propiedades que se han considerado especiales y específicas de la Gramática Universal (GU) postulada por el paradigma de investigación minimista en la lingüística formal: la operación *Merge* y los rasgos combinatorios responsables de la derivación sintáctica. Sobre la base de una exposición versada de la literatura generativa, la autora destaca la centralidad de los rasgos formales en la computación lingüística, su carácter determinante en tanto desencadenan mecanismos de corrección de imperfecciones en la derivación, camino a sostener que serían la propiedad exclusiva del lenguaje. *Merge*, por otra parte, no podría ser exclusivo de la capacidad lingüística, ya que constituiría un mecanismo combinatorio presente también en otros sistemas físicos. La recursividad propiamente lingüística sería de carácter eminentemente estructural, pues necesita incorporar el anidamiento o incrustación de elementos del mismo tipo en otros.

Se sostiene, por último, que, si bien el proceso de externalización ha sido ventajoso para la adaptación humana, sería secundario en el devenir evolutivo de los componentes del lenguaje explorados.

La sección culmina con el artículo de Marcos Herrera Burstein, titulado «El “argumento del diseño” y la crítica chomskiana a la explicación adaptacionista de la evolución del lenguaje». En él, el autor sostiene que el tajante rechazo de Massimo Piattelli-Palmarini (1989) y de Hauser, Chomsky y Fitch (2002) a la explicación adaptacionista de la evolución del lenguaje defendida por Steven Pinker y Paul Bloom (1990), así como por Pinker y Ray Jackendoff (2005), a partir del argumento del diseño de Richard Dawkins (1986), es menos un resultado de un examen neutral y objetivo de los datos empíricos y más la consecuencia inevitable de uno de los supuestos fundamentales del programa de investigación de la gramática generativa desde sus inicios. El postulado de la autonomía de la sintaxis (Chomsky, 1957) implica que la forma del lenguaje (al menos en su núcleo más central) es independiente de cualquier función. La explicación adaptacionista asume, por el contrario, que la forma del lenguaje debe cumplir alguna función adaptativa, pues, sino, no habría evolucionado mediante la selección natural, que sería el único mecanismo capaz de producir diseños complejos en los seres vivos. El autor muestra, adicionalmente, cómo la propuesta de Chomsky (2010) acerca de la evolución del lenguaje otorga solo un rol secundario a la intersubjetividad.

La última sección de esta segunda parte, sobre el desarrollo y estructura del lenguaje, se inicia con un artículo de María de los Ángeles Fernández Flecha titulado «¿Cómo se comunican los infantes en el segundo año de vida? El caso de las funciones declarativa e imperativa». Sobre la base de evidencia empírica, la autora muestra que en la comunicación temprana se observan ya asociaciones de forma y función bastante estables y, además, multidimensionales, en tanto cada conducta comunicativa infantil supone un ajuste de las diversas habilidades involucradas en su expresión, según cuál sea su función o intención. El periodo analizado estaría marcado por procesos generales de desarrollo: la reducción progresiva de las conductas gestuales en favor de las vocales, el aumento en la producción de protopalabras y el incremento en la producción de respuestas por parte del niño. De este modo, a través de un patrón continuo de desarrollo, las expresiones más próximas al modelo adulto se irían imponiendo progresivamente a las más inmaduras, con el resultado de que el infante será un usuario cada vez más competente de su lengua y, también, un comunicador más eficaz.

En «El carácter mental de los sistemas fonológicos», Jorge Iván Pérez Silva muestra, a partir de evidencia empírica, que los seres humanos aprendemos a procesar auditivamente el habla en función de los estímulos acústicos propios de nuestro

medio ambiente lingüístico, lo que determina que categoricemos perceptivamente como una misma unidad fonológica señales que son físicamente diferentes. El autor propone que estas *unidades perceptuales* tienen un carácter mental que no se puede reducir a las propiedades físicas de las ondas sonoras ni a la activación del sistema nervioso causadas por estas a través de las operaciones anatómico-fisiológicas del oído. A la vez, sostiene que las unidades de percepción se distinguen de las otras unidades mentales que emergen de la organización de la información que realiza nuestro cerebro cuando categorizamos diferentes individuos en clases: las *unidades conceptuales*. Si bien ambos tipos de categorías se construyen a partir de la experiencia sensible y de la intersubjetividad, las últimas se ubican en un nivel de conciencia compartido por los miembros de una comunidad —el pensamiento o conocimiento del mundo—, mientras que las primeras se ubican en un nivel «operativo» o «procesual», por lo general, no consciente —el conocimiento lingüístico—.

La sección finaliza con el artículo «La naturaleza de la facultad del lenguaje», de Liza Skidelsky, en el que se discute la versión arquitectural del lenguaje propuesta por John Collins (2004, 2006 y 2008). Según dicho autor, muestra Skidelsky, la facultad del lenguaje sería un sistema computacional de la mente/cerebro especificado de manera abstracta, sin que dicha especificación pueda concebirse como un mecanismo, es decir, como parte de unnexo causal. La autora plantea, en contraposición a Collins, que la facultad del lenguaje es un mecanismo causal (aunque idealizado) de la misma naturaleza que los mecanismos de los cuales se pretende dar cuenta en cualquier teoría científica de la arquitectura cognitiva. Así, la autora defiende la posibilidad de que ciertas teorías o modelos ofrezcan una explicación que aluda solo a las propiedades funcionales de la entidad responsable de los fenómenos cognitivos por explicar —lo que no niega que el objeto de estudio sea un mecanismo que posee tanto propiedades funcionales como estructurales que manifiestan disposiciones causales que permiten insertarlo en una red causal—. Finalmente, propone que la versión arquitectural causal de la facultad del lenguaje rescata la concepción chomskiana de la lingüística como el estudio abstracto de los mecanismos de la mente/cerebro y que el modelo de la facultad del lenguaje es una hipótesis empírica acerca de procesos psicológicos.

Aunque los artículos que conforman la segunda parte del libro son, en principio, independientes entre sí, abordan aspectos diferentes del tema genérico del libro y defienden tesis que pueden o no ser compatibles con las de los otros artículos. Todos ellos son parte de la discusión actual acerca de la evolución y el desarrollo de las relaciones entre cognición social y lenguaje, de manera que este libro ofrece un panorama bastante completo y, al mismo tiempo, especializado del estado de la cuestión.

REFERENCIAS BIBLIOGRÁFICAS

- Aristóteles (1988). *Acerca del alma*. Traducción de Tomás Calvo Martínez. Madrid: Gredos.
- Bickerton, Derek (1990). *Language & Species*. Chicago: University of Chicago Press.
- Bickerton, Derek (2009). *Adam's Tongue: How Humans Made Language, How Language Made Humans*. Nueva York: Hill and Wang.
- Bremmer, Jan (1983). *The Early Greek Concept of the Soul*. Nueva Jersey: Princeton University Press.
- Carruthers, Peter (2009). How We Know Our Own Mind. The Relationship between Mindreading and Metacognition. *Behavioral and Brain Sciences*, 32, 121-182.
- Cavell, Marcia (2006). *Becoming a Subject. Reflections in Philosophy and Psychoanalysis*. Oxford: Oxford University Press.
- Chalmers, David (1995). Facing up the Problem of Consciousness. *Journal of Consciousness Studies*, 2, 200-219.
- Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Nueva York: Oxford University Press.
- Chalmers, David (2006). Strong and Weak Emergence. En Philip Clayton y Paul Davies (eds.). *The Reemergence of Emergence: The Emergentist Hypothesis from Science to Religion* (pp. 244-256). Nueva York: Oxford University Press.
- Chomsky, Noam (1957). *Syntactic Structures*. Berlín: Mouton de Gruyter.
- Chomsky, Noam (2010). Some Simple Evo Devo Theses: How True Might They Be for Language? En Richard Larson, Viviane Déprez y Hiroko Yamakido (eds.), *The Evolution of Human Language. Biolinguistic Perspectives* (pp. 45-62). Cambridge: Cambridge University Press.
- Collins, John (2004). Faculty Disputes. *Mind & Language*, 19(5), 503-533.
- Collins, John (2006). Between a Rock and a Hard Place: A Dialogue on the Philosophy and Methodology of Linguistics. *Croatian Journal of Philosophy*, 6, 471-505.
- Collins, John (2008). Knowledge of Language Redux. *Croatian Journal of Philosophy*, 7, 3-42.
- Davidson, Donald (2001). *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- Dawkins, Richard (1986). *The Blind Watchmaker: What the Evidence of Evolution Reveals a Universe without Design*. Nueva York: Norton.
- Dunbar, Robin (1998). The Social Brain Hypothesis. *Evolutionary Anthropology*, 6, 178-190.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky (2005). The Evolution of the Language Faculty: Clarifications and Implications. *Cognition*, 97, 179-210.

- Fonagy, Peter (2008). The mentalization Approach to Social Development. En Fedric N. Busch (ed.), *Mentalization. Theoretical Considerations, Clinical Findings and Research Implications* (pp. 3-56). East Sussex: The Analytic Press.
- Goldman, Alvin (2006). *Simulating Minds: The philosophy, Psychology and Neuroscience of Mindreading*. Oxford: Oxford University Press.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598), 1569-1579.
- Johnson, Mark (1991). *El cuerpo en la mente. Fundamentos corporales del significado, la imaginación y la razón*. Madrid: Debate.
- Johnson, Mark (1993). *Moral Imagination. Implications of Cognitive Science for Ethics*. Chicago-Londres: The University of Chicago Press.
- Kirk, Geoffrey Stephen & Raven, John Earle (1981). *Los filósofos presocráticos. Historia crítica con selección de textos*. Madrid: Gredos.
- Piattelli-Palmarini, Massimo (1989). Evolution, Selection and Cognition: From «Learning» to Parameter Setting in Biology and in the Study of Language. *Cognition*, 31(1), 1-44.
- Pinker, Steven & Paul Bloom (1990). Natural Language and Natural Selection. *Behavioral and Brain Sciences*, 13(4), 707-784.
- Pinker, Steven & Ray Jackendoff (2005). The Faculty of Language: What's Special about It? *Cognition*, 95, 201-236.
- Platón (1983). *Fedón*. Traducción de Conrado Eggers Lan. Buenos Aires: Eudeba.
- Rohde, Erwin (1948). *La idea del alma y la inmortalidad entre los griegos*. México DF: Fondo de Cultura Económica (FCE).
- Searle, John (1992). *The Rediscovery of the Mind*. Cambridge: The MIT Press.
- Searle, John (1997). *The Mystery of Consciousness*. Nueva York: The New York Review of Books.
- Snell, Bruno (1982). *The Discovery of the Mind. In Greek Philosophy and Literature*. Nueva York: Dover.

PARTE I
MARCO CONCEPTUAL

EVOLUCIÓN Y DESARROLLO DE LA COGNICIÓN SOCIAL

Mente y Lenguaje

Grupo Interdisciplinario de Investigación

Andrés Abugattás, Ricardo Braun, Paola Cépeda, César Escajadillo,
María de los Ángeles Fernández Flecha, Marcos Herrera Burstein,
Carla Mantilla, Luis Manuel Olguín, Jorge Pérez Silva,
Pablo Quintanilla y Carolina Romero

La cognición social, en un sentido amplio, es un concepto que se refiere a todas las funciones mentales implicadas en la comprensión del mundo social; de este modo, es un campo que interesa de manera especial a la psicología social. Esta rama de la psicología ha sido, por algún tiempo, una ciencia cognitiva; sin embargo, luego se vio profundamente influida por la revolución cognitiva que tuvo lugar entre los años 1960 y 1970. A partir de ese momento, la cognición social tomó prestadas las teorías y técnicas de la nueva psicología cognitiva para aplicarlas a los cuestionamientos tradicionales de la psicología social, como las actitudes, la atribución causal, la inferencia social, entre otros. La revolución mencionada se caracterizó, principalmente, por la metáfora del procesamiento de la información, en la cual se plantea la existencia de una serie de procesos mentales por los que circularía la información y daría lugar a la cognición. El circuito sería transitado ante cualquier tipo de cognición, siendo una de ellas la cognición social. Es decir, mientras que los contenidos varían, los procesos serían siempre los mismos. Asimismo, el énfasis fue puesto en las representaciones mentales, con lo cual se dejaba de lado las motivaciones, las emociones, el comportamiento y la interacción social, temas de interés central para los psicólogos sociales. Además, la idea de que todo tipo de cognición tuviera el mismo proceso, sin variación alguna, resultaba un poco radical. Es así como, más tarde, la cognición social se fue ubicando en un punto más intermedio; es decir, si bien incluye la metáfora de procesamiento de información y considera que existen procesos cognitivos generales, también plantea que hay procesos específicos para ciertas áreas. Las teorías modulares se ubicarían aquí (Gilbert, 1999).

A continuación explicaremos la cognición social a partir de la metáfora del procesamiento de la información —la cual es ampliamente aceptada entre los psicólogos sociales—. Esto es, cómo la información es incorporada, elaborada, representada y guardada en la memoria, para luego ser utilizada en la percepción e interacción con el mundo social (Hamilton, 2005). La idea general es que el individuo adquiere nueva información del contexto social, la procesa de distintas maneras, la integra con conocimientos previos y la utiliza para entender a los otros, tomar decisiones y guiar su comportamiento.

Lo primero necesario para que se lleve a cabo este proceso es la atención, la cual permitirá luego la codificación. Usualmente el entorno contiene más información de la que un individuo es capaz de atender; de este modo, la persona se ve forzada a seleccionar ciertos elementos del contexto mientras que otros quedan sin registrar. La atención selectiva es funcional en tanto permite dirigir la atención hacia los elementos más relevantes; sin embargo, al mismo tiempo deja fuera información que podría ser importante para una mayor comprensión de una situación social determinada.

Después de seleccionar y registrar la información, el individuo realiza una serie de elaboraciones sobre aquel material —antes de almacenarlo en la memoria— para comprenderlo mejor y darle un sentido, lo cual a su vez le permite adaptarse a su medio social. Dos tipos de elaboración importantes son la interpretación y la inferencia. Dentro de los tipos de inferencia posibles, el proceso de atribución es uno esencial en tanto nos proporciona la explicación del comportamiento del otro, es decir, sus causas. Ello no solo nos ayuda a entender el porqué de los eventos sino que también nos ayuda a anticipar eventos futuros similares. Un tercer tipo, también importante, es la evaluación, que permite ubicar los eventos en categorías como bueno o malo, bonito o feo, admirable o despreciable, etcétera.

El siguiente paso después de la elaboración consiste en representar la información en la memoria y almacenarla para hacer uso de ella en el futuro. Existen distintas teorías sobre cómo se guarda la información. Algunas sugieren que cada elemento de la información es tomado como un ejemplar y almacenado de manera individual (teorías de ejemplares); otras señalan que los elementos se almacenan interconectados, formando circuitos (teorías de circuitos); y otras indican que de los elementos de la información se obtiene un significado, un concepto más general —esquema o prototipo— bajo el cual se los agrupa (teorías basadas en la abstracción). Independientemente de cuál sea la teoría más adecuada para entender el almacenamiento de la información, lo cierto es que aquella que se almacena no consiste en una representación objetiva de la realidad. Primero, se han seleccionado únicamente ciertos elementos y, en segundo lugar, ha habido una serie de procesos elaborativos previos al almacenamiento, lo que implica una transformación de la información.

Y lo importante es que esa información transformada es la que va a guiar la acción futura del individuo.

Luego de almacenada la información, esta puede ser recuperada (evocada) para formar juicios y guiar el comportamiento. Sin embargo, así como el almacenamiento no es objetivo, tampoco lo es la recuperación de la información. Este proceso va a estar influido por las metas o propósitos inmediatos de la persona, seleccionando ciertos elementos y no otros. Se pierde la mirada global. Algunos elementos serán recordados más fácilmente que otros por el impacto que tuvieron en el individuo y podrán evocar otros recuerdos que tengan alguna conexión con los primeros. Es más, nuestros objetivos actuales pueden reconstruir nuestros recuerdos de tal modo que encajen con nuestras creencias y metas actuales.

Finalmente, como resultado de todo este proceso se da la respuesta, que puede incluir juicios, reacciones afectivas, decisiones y comportamientos. Como hemos visto, una serie de factores influyen en el procesamiento de la información y lo convierte en un proceso único y capaz de variar de acuerdo al momento en que se realiza. De este modo, una misma persona puede parecer inconsistente en tanto frente a estímulos similares puede realizar procesamientos distintos y, por tanto, mostrar respuestas diferentes.

Algo sumamente interesante que se ha observado en varias investigaciones es que, a menudo, llevamos a cabo el procesamiento de información social de manera automática, no deliberada. Es decir, muchas veces no nos damos cuenta de aquello que estamos registrando, elaborando y almacenando en la mente, pues ni siquiera interrumpe las tareas cognitivas que se estén realizando de manera paralela y consciente. No obstante, se ha encontrado que este procesamiento automático tiene gran influencia en nuestras respuestas y, por tanto, en la adaptación a nuestro entorno social.

Hasta aquí se ha descrito, fundamentalmente, los pasos del procesamiento de la información social, mas no se ha dicho nada sobre los contenidos. De manera general, nos manejamos con estructuras cognitivas que son conceptos que se han ido construyendo a partir de las propias experiencias y que representan nuestro conocimiento, nuestras creencias y expectativas en distintas áreas de nuestra vida. Estas estructuras, como las teorías implícitas de la personalidad o los estereotipos, pueden influir en cada paso del procesamiento de la información descrito más arriba. Ayudan a mantener el *statu quo* y nos permiten anticipar el futuro. Esto debido a que, en la mayoría de los casos, la nueva información es integrada a las estructuras cognitivas existentes (Hamilton, 2005).

Varias de estas estructuras cognitivas funcionan como atajos mentales que nos ayudan a realizar un menor gasto de energía al procesar la información nueva. Sin embargo, utilizarlos conlleva el riesgo de caer en error. Estos se denominan *heurísticos*.

Por ejemplo, en lugar de hacer un análisis detallado de una persona la ubicamos en una categoría por la similitud que le encontramos con algún grupo particular (persona de barba y pelo largo, con una vestimenta desgarbada y sandalias nos podría remitir a un *hippie*). Otro heurístico común es el que nos permite pensar en la probabilidad de un atributo a partir de la facilidad con que recordamos ese atributo al pensar, por ejemplo, en un grupo (si cuando pensamos en gente negra lo primero que se nos viene a la mente es que cantan muy bien, cuando estemos frente a una persona negra le vamos a atribuir esa cualidad casi de inmediato).

De otro lado, es interesante mencionar que nuestras propias metas y objetivos pueden influir de manera importante en el procesamiento de información social y en la respuesta que a partir de él se genere. Esto es, no solo puede determinar qué se registra sino cómo se registra. Por ejemplo, se ha encontrado que las personas tienden a ser menos exigentes en cuanto a las evidencias cuando una propuesta es cercana a nuestro propio pensamiento. En contraste, si se aleja de él, es más probable que seamos más meticulosos al momento de evaluar la información para constatar su validez.

A partir de lo expuesto anteriormente vemos cómo la cognición social, aunque presentada desde un modelo que la describe y explica de una manera relativamente sencilla, es en realidad un proceso sumamente complejo en el que intervienen innumerables variables de distinto tipo. A nosotros nos interesa centrarnos especialmente en una sección del procesamiento de información que se encuentra dentro de las actividades de elaboración: la atribución.

Los procesos de atribución no solo han sido de interés para la psicología social sino también para otras áreas de estudio, utilizando sus propios métodos de investigación y conceptos particulares. En las ciencias cognitivas, y por aporte de filósofos, por ejemplo, un concepto particularmente relevante en este contexto es el de metarrepresentación. La atribución de representaciones mentales es, pues, una forma de representación de segundo grado: es una representación de una representación. Cuando intentamos explicar o predecir el comportamiento de algún agente solemos emplear ciertos verbos que describen la actividad mental que, asumimos, acompaña al comportamiento. Así, decimos que el agente actuó de cierta manera porque quería o deseaba algo, porque pensaba en algo, etcétera. Querer, desear o pensar en algo implica poder dirigir la conciencia hacia aquello que nos ocupa, lo cual implica, a su vez, alguna forma de representación.

La metarrepresentación adopta dos formas básicas. Cuando la empleamos en agentes distintos de nosotros mismos recibe el nombre de 'lectura de mentes' (*mindreading*). Otra manera de referirse a ella es como atribución de representaciones en tercera persona. Pero, también podemos realizar atribuciones de representaciones en primera persona, es decir, vernos a nosotros mismos como agentes. En este caso, se habla de *metacognición*.

La metarrepresentación es una poderosa herramienta de la sociabilidad humana. Debido a su poder predictivo es comparable con la postulación de hipótesis empíricas de la actividad científica. En ciencias cognitivas se habla de ‘teoría de la mente’ (*theory of mind*) o de *psicología folk* para hacer referencia al conocimiento común, prácticamente intuitivo, que poseemos sobre la relación entre los contenidos mentales y el comportamiento de los agentes, incluyéndonos a nosotros mismos. El término *teorías de la teoría de la mente* (TT) se refiere a los posibles modelos teóricos que explican el conocimiento psicológico folk.

Muchas veces quienes comparan nuestras labores cotidianas de intérpretes psicológicos con el empleo de teorías científicas, sostienen que nuestras predicciones respecto del comportamiento de los agentes son el producto de deducciones hechas a partir de principios psicológicos generales, ya sean implícitos o explícitos. Sin embargo, también hay quienes consideran que para explicar la interpretación psicológica no es necesario postular una suerte de conocimiento teórico folk. La alternativa más conocida a las TT son las llamadas *teorías de la simulación* (TS). Las teorías de la simulación se limitan a postular que todo lo que necesitamos para predecir el comportamiento de otros agentes es intentar proyectarnos a su situación y, desde ahí, deducir qué podrían estar pensando. En otras palabras, en tanto intérpretes del comportamiento, simulamos ser los agentes cuyas acciones queremos explicar o predecir.

Más allá de las evidentes discrepancias generales, un punto importante en el debate entre defensores de las TT y defensores de las TS es en torno de la atribución de representaciones en primera persona. Para el segundo grupo de autores la interpretación de comportamiento de los agentes parte de la aprehensión directa de los propios estados mentales del intérprete. Como dijimos arriba, todo el conocimiento que podemos aspirar a tener de las mentes de los demás es solo una proyección de lo que hemos simulado, en nosotros mismos, como la situación del otro. Una consecuencia de esta tesis es que la metacognición es considerablemente menos falible que la lectura de mentes, pues solo el conocimiento de nuestras propias mentes sería directo. Se sostiene, además, que la seguridad superior de la metacognición, también conocida como la *autoridad de la primera persona*, se debe a que poseemos un mecanismo de introspección mental. Uno de los defensores más conocidos de esta posición es Alvin Goldman (2006). Algunos simpatizantes de las TT niegan la necesidad de postular un mecanismo semejante y, por ello, admiten que la metacognición puede ser casi tan falible como la lectura de mentes. Como veremos más adelante, autores como Peter Carruthers (2009), conocido defensor de las TT, piensan que el debate debe solucionarse mediante la comparación de las explicaciones filogenéticas de los mecanismos en cuestión.

Como podemos apreciar, las perspectivas hasta ahora expuestas se centran, esencialmente, en los procesos cognitivos implicados en la construcción de representaciones mentales sobre los otros y sobre nosotros mismos, lo cual es fundamental para la cognición social. Sin embargo, existe otro grupo de teorías que más bien se concentra en los procesos sociales que permitirían aflorar aquellas capacidades mentales necesarias para la cognición social. Aclaremos que ninguno de los grupos niega la importancia social o la importancia del factor biológico, respectivamente; se trata de distintos enfoques en la investigación de la cognición social.

Dentro de las teorías relacionales encontramos la de la mentalización que, a su vez, se apoya en la teoría del apego. La capacidad para la mentalización involucra, en realidad, una serie de habilidades que permiten al individuo interpretar su propia conducta y la de los demás a partir de la atribución de estados mentales —como sentimientos, deseos, creencias, entre otros (Fonagy, 2008)—. Para que la mentalización se consolide de manera adecuada, es necesario que se genere un apego seguro en el niño y, para que eso ocurra, los cuidadores deben ser capaces de leer las señales del bebé y responder a ellas de manera adecuada. Algunas de las labores más importantes de los adultos son amplificar las expresiones emocionales positivas del bebé y modular las negativas. De manera general, es a través del intercambio con los padres o cuidadores que el infante hace uso de las funciones ya maduras del cerebro adulto para organizar sus propios estados mentales. Es gracias a esos encuentros interpersonales que el bebé logra desarrollar una serie de funciones mentales muy importantes como la memoria, la narración, las emociones, las representaciones y los estados mentales (Siegel, 1999). Todas ellas esenciales para un adecuado desarrollo y desenvolvimiento en el mundo social.

Este primer capítulo del trabajo está constituido por dos subtítulos. El primero se centrará en algunas teorías sobre la historia filogenética de las facultades cognitivas humanas implicadas en la cognición social y, el segundo, abordará el tema desde una perspectiva del desarrollo, es decir, desde una mirada ontogenética, e incluirá algunos planteamientos de las teorías relacionales.

1. LA EVOLUCIÓN DE LA COGNICIÓN SOCIAL

Este apartado cuenta con cinco partes principales. La primera es una breve introducción histórica a la investigación filogenética de la cognición social que, como en muchas otras formas de estudios filogenéticos, tiene sus raíces en Darwin. La segunda parte está dedicada a algunas teorías sobre la evolución de los elementos del cerebro humano que influyen en la interacción social. En la tercera parte, veremos discusiones actuales sobre conceptos centrales para las teorías filogenéticas de la atribución psicológica. La cuarta parte trata sobre la hipótesis de la inteligencia social, la explicación

más aceptada de la evolución de la atribución psicológica. En la última parte del apartado, veremos algunos problemas filosóficos que hoy se abordan empleando argumentos extraídos de las teorías sobre la evolución de la cognición social.

1.1. La teoría darwiniana de la selección natural de las facultades sociales humanas

Si bien la idea de que las especies evolucionan o han evolucionado se remonta hasta prácticamente los orígenes de la filosofía griega, con Anaximandro, fue a partir de la postulación de la teoría de la selección natural que las investigaciones evolutivas han sido aprovechadas como una importante herramienta para la discusión en diversas ciencias. Como se sabe, Darwin sostuvo que los rasgos biológicos aleatoriamente heredados son seleccionados de acuerdo con un principio simple, a saber, si son útiles para la supervivencia de los individuos. Naturalmente, los rasgos más exitosos son los que se transmiten sistemáticamente de generación en generación, pues gracias a ellos los organismos portadores tienden a sobrevivir y reproducirse con eficacia.

Uno de los méritos de Darwin es el haber ofrecido un argumento, probablemente definitivo, contra el esencialismo platónico y aristotélico. Para Aristóteles, las especies animales son eternas e inmutables. Los individuos participan de una esencia común que se puede rastrear hasta Platón, de hecho, la palabra *especie* fue una traducción latina para *forma* o *eidos*. Darwin rechaza la inmutabilidad de las especies, para él las especies están siempre evolucionando, en el sentido de transformándose (de hecho, antes que se usara la palabra *evolución* se empleaba *transformismo*), y si bien los individuos de la misma especie comparten muchos rasgos, no hay entre ellos rasgos esenciales y accidentales, sino que todos los rasgos son accidentales y llamamos *esenciales* a aquellos rasgos accidentales compartidos por muchos individuos en un momento dado.

Según Darwin, entonces, las especies no son entidades fijas, como sostenía el llamado fijismo de su época que, siguiendo a Aristóteles, defendía que los individuos que pertenecen a una especie comparten una esencia común, que siendo estable permite que un individuo sea lo que es y no otro. Asimismo, habría una diferencia cualitativa y no meramente cuantitativa entre especies diferentes. Para Darwin, por el contrario, no solo no hay esencias estables compartidas por los individuos que pertenecen a una especie sino que, además, las diferencias entre especies son meramente cuantitativas. Es verdad que los individuos de una especie pueden reproducirse entre sí y no con miembros de otras especies (la posibilidad de la reproducción suele considerarse el criterio principal para distinguir entre especies), pero eso solo se produce cuando dos grupos de individuos se han apartado genéticamente lo suficiente entre sí para que sea imposible que se reproduzcan, que es precisamente lo que se llama el proceso de *especiación* o formación de las especies.

Fue el mismo Darwin quien no tardó en explotar la potencia explicativa de su propia teoría, empleándola no solo en la investigación de los orígenes de rasgos meramente fisiológicos de nuestra especie. No es extraño, entonces, que los capítulos III, IV y V de *El origen del hombre* (1969) estén dedicados a defender la tesis de que las diferencias entre las facultades mentales en general, y morales en particular, del ser humano y de los animales no son esenciales. Por el contrario, en muchos casos se trataría de distinciones de grado determinadas, evidentemente, por la selección natural.

El primer problema que debe enfrentar Darwin para defender su tesis consiste en la distinción entre lo perteneciente al mero instinto y la inteligencia, entendida esta, a veces, como el libre albedrío y, quizás más importante aún, también como lo aprendido y no fijo (1969, p. 87). Como se verá más adelante, a medida que la investigación avanza, la distinción se hace más opaca, lo cual era de esperarse. Como advierte el autor desde un principio, una vez asumida la distinción no es fácil determinar qué facultades pertenecen a una y otra categoría. Podría ser, por ejemplo, que aquello que tuvo origen gracias al uso de la inteligencia se convierta, a través de las generaciones, en actos instintivos. Además, no siempre se cumple la idea clásica de que la relación entre instinto e inteligencia es inversamente proporcional (pp. 86-87).

No obstante lo dicho arriba y aunque sus argumentos no carezcan de tensiones, Darwin parece hacer recaer el peso de la distinción en cuestión en una mayor *automaticidad* de los actos animales en comparación con aquellos de los humanos, provistos de conciencia. Afirma que el fundamento del comportamiento humano radica en «la memoria de los acontecimientos pasados, en la previsión, en la razón y en la imaginación» (1969, p. 89). Y aunque en algunos animales podemos ver acciones que se asemejan en muchos aspectos de su ejecución y fines a las humanas, en aquellas «la variabilidad de los órganos mentales, unida a la selección natural, es la que adquiere paso a paso capacidad para la ejecución de tales acciones sin que contribuyese a ello la inteligencia ni la conciencia del animal en cada una de las distintas generaciones» (p. 89). El mismo punto es reforzado cuando el autor hace referencia a la adquisición de ciertas habilidades en los animales: en muchos casos, para estos bastaría con la sencilla imitación, o a veces ni siquiera eso, para que aquellas alcancen su máximo nivel de desarrollo, mientras que tratándose de los individuos humanos la experiencia juega un papel preponderante.

Frente a la carencia de información sobre la evolución y el desarrollo de las diversas facultades mentales tanto en animales como en humanos, el principal recurso argumentativo de Darwin consiste en citar muchos ejemplos de investigaciones de naturalistas de la época, incluidas algunas suyas, e intentar mostrar cómo ciertos comportamientos animales registrados son indicadores de emociones y procesos mentales semejantes a los nuestros. Resultan particularmente notables las observaciones que

el autor hace acerca de la capacidad que algunos animales poseen para resolver tareas complejas o que distan considerablemente de los procesos que usualmente resuelven mediante hábitos. En muchos casos, ante el error (generalmente repetido, aunque en los mamíferos superiores esto no es necesario), aquellos pueden corregir el curso de sus acciones y realizarlas exitosamente. Sin mayores problemas, Darwin compara esta facultad animal con lo que normalmente llamamos *reflexión*, *deliberación* o simplemente *raciocinio* en los seres humanos, salvando las distancias respecto de nuestra muchísimo mayor capacidad de observación y deducción.

De la misma manera, Darwin considera al fenómeno del lenguaje. La facultad de asociar sonidos a conceptos más o menos bien determinados no es un problema tratándose de animales; bastaría con observar cómo estos responden a ciertas voces nuestras. Inclusive, el autor llega a poner en cuestión que el encadenamiento de ideas requiera necesariamente del empleo de signos lingüísticos (1969, p. 117), lo cual también echa algunas luces sobre la teoría lingüística que aquel aceptaba. Tampoco sería un problema la capacidad de articular signos, aunque en torno de este punto el naturalista emplea el cuestionable ejemplo de las aves capaces de imitarnos acústicamente (p. 113). Nuevamente, la distinción debe ser una cuestión de grado, y en este caso particular ocurre que los seres humanos simplemente poseemos una muchísima mayor capacidad de enlazar libremente signos y significados (p. 114).

Darwin también se pone de lado de quienes consideran que, en el fondo, es el comportamiento moral lo que marca la principal diferencia entre seres humanos y animales. No obstante, según pensaba, la moral también puede explicarse genealógicamente empleando la teoría de la selección natural. Es más, si los antepasados de los hombres (y otras especies no tan cercanas, como muchos mamíferos) ya contaban con formas básicas de simpatía¹ y de ciertos instintos de conservación grupal, bastaría con añadir a la ecuación una capacidad de raciocinio semejante a la nuestra y obtendríamos, de forma natural, el sentido moral humano tal y como lo entendemos hoy en día.

A pesar de las dificultades que el propio Darwin reconoce cuando se intenta atribuir a los animales estados mentales y afectivos que sin dificultad alguna atribuimos a otras personas, no duda en responsabilizar a la simpatía por muchos actos altruistas que se observan en el mundo natural². Igualmente, piensa que es difícil explicar

¹ Darwin no distingue aún entre «simpatía» y «empatía», aunque podemos atrevernos a afirmar, a juzgar por los ejemplos que utiliza, que lo que él llama «simpatía» se acerca mucho más a lo que los estudiosos contemporáneos llamarían por el mismo nombre.

² Como veremos más adelante, a Darwin no le caben dudas sobre la preponderancia de la simpatía en el mundo humano. Basta con leer la siguiente cita: «Cuanto más importancia ponga el hombre en el aprecio de los demás, mayor fuerza tendrá en él el sentimiento innato o adquirido de simpatía, y asimismo su propia capacidad para deducir las consecuencias de sus actos» (1969, p. 153).

comportamientos tales como las rondas de vigilancia y otro tipo de precauciones en diversas especies animales sin apelar a un instinto común de conservación de grupo. Resulta interesante señalar, a propósito de la simpatía y otros instintos morales, que tomarlos como punto de partida para la evolución de la moral permite evitar tesis simplistas que apelan a relaciones de placer y dolor como la fuente original del comportamiento social. La simpatía y el instinto de conservación de grupo conducen a los individuos a acciones que incluso podrían poner en riesgo su propia seguridad, y esta tendencia parece haberse mantenido y alimentado en el caso de los seres humanos, lo cual confirmaría la hipótesis darwiniana de que la selección natural ha favorecido el comportamiento altruista a través de las eras y las especies.

Antes dijimos que las facultades mentales superiores de los seres humanos jugaban un rol determinante en el distanciamiento entre estos y sus antepasados animales, a pesar de la relación de parentesco. Siguiendo algunas pistas que deja Darwin sobre este punto, podemos inferir lo que sigue. Tratándose de los animales, los mecanismos que configuran su sociabilidad están más fuertemente determinados por los instintos. Si se quiere, podemos decir, utilizando lenguaje informático, que están *preprogramados* en un porcentaje mucho mayor que los mecanismos humanos equivalentes. En este último caso, la carencia de instintos es subsanada con la mayor potencia del raciocinio (1969, p. 146) (podemos agregar, también, las mayores posibilidades de comunicación que nos brinda nuestro lenguaje). Una de las consecuencias de ello es la importancia enorme que poseen las opiniones de la comunidad cuando nos disponemos a actuar. De hecho, Darwin piensa que nuestra facultad de considerar los juicios que los demás pueden hacer sobre nuestros actos sigue evolucionando a medida que nos volvemos más apegados a formas cada vez más abstractas de valoración de los primeros y se atenúa la relevancia de las impresiones subsecuentes, ya sean de placer o de dolor (p. 147).

Si bien, como vimos, las facultades racionales parecen ser un elemento fundamental de la distinción entre moral humana y animal, no por ello se debe disminuir el valor del instinto en los asuntos morales, ni siquiera tratándose del ser humano. Ya hemos visto que Darwin parte en su investigación de la diferenciación entre razón e instinto, mas esta tampoco es clara tratándose de lo moral. De una parte, sabemos que la acción de lo que podemos llamar «instinto» en los asuntos morales humanos no es poca. De otra, también en este ámbito encontramos cierta deliberación entre los animales, a quienes no pocas veces puede vérselos indecisos ante más de una opción para resolver una situación difícil. Tanto en los seres humanos como en los animales que viven en sociedad, la reflexión o deliberación moral proviene de la constante presencia en los individuos de «impresiones o imágenes del pasado» (1969, p. 150). Pero, dadas las mayores facultades mentales humanas, dichas impresiones,

que incluyen los juicios aprobatorios o desaprobatorios de nuestros compañeros en la sociedad, son muchísimo más persistentes (persistencia que, sobre todo en lo que respecta a los juicios de los demás, se apoya en la simpatía). El resultado de su persistencia son los sentimientos como el arrepentimiento, la vergüenza o el remordimiento. A su vez, dichos sentimientos son distintas caras de lo que usualmente llamamos *conciencia moral*, y es su peso lo que nos obliga a revisar nuestras acciones y querer corregirlas en las siguientes ocasiones (p. 152).

La simpatía, entonces, se ha visto favorecida por la selección natural debido a su capacidad de fomentar la sociabilidad de humanos y ciertos animales. De ahí que Darwin afirme que es la consecución del *bien general*³ lo que ha servido como principio regulador de la evolución de la moral humana, en la cual la preponderancia de la simpatía es mucho mayor que en otras especies sociales. Es más, el autor no duda en afirmar que, entre nosotros, la búsqueda de la felicidad queda constantemente relegada a un segundo plano con relación al bien de todos. En virtud del éxito que este principio demuestra en la formación de sociedades cada vez más grandes y complejas a lo largo de la historia humana, fuerza la expansión de la simpatía entre los individuos más allá de los límites de la propia comunidad, de manera que los seres humanos poseemos, hoy en día, la capacidad de sistemáticamente sentir simpatía hacia cualquier miembro de nuestra especie e, inclusive, con miembros de otras especies⁴.

A pesar de las tensiones manifiestas del texto, de que en la época recién se desarrollaban las primeras investigaciones al respecto y de que, además, muchos de los ejemplos y referencias citadas por el autor hoy en día serían altamente cuestionables, encontramos entre sus páginas las raíces de muchas teorías actuales o, en algunos casos, intuiciones que recientemente han tomado vigor entre algunos investigadores. Desde ya, el intento de Darwin por encontrar en formas de vida no humanas facultades mentales similares, en alguna medida, a las humanas lo conduce a poner en tela de juicio (aunque algunas veces contra sus propias premisas) ideas como la antes mencionada separación clara y distinta entre la razón y el instinto. Ciertamente, la distinción común en la filosofía de la mente contemporánea entre *Sistema 1* y *Sistema 2* ha ampliado el rango de funciones que estamos dispuestos

³ Darwin define el término de la siguiente manera: «[...] es el medio de producir, dentro de las condiciones existentes, el mayor número de individuos en pleno vigor y plena salud, dotados de facultades todo lo perfectas posible» (1969, p. 162). En vista de ello, no es de extrañar que el propio Darwin considere que su tesis constituye una refutación de cualquier explicación moral que sitúe en el punto de partida principios derivados del egoísmo.

⁴ Ya antes David Hume había descrito, en el tercer libro de su obra *A Treatise of Human Nature*, un movimiento similar de los lazos morales entre los individuos humanos, desde niveles más básicos y locales, hasta niveles más generales o abstractos (cabe recordar la cercanía entre Hume y Adam Smith, quien es citado por Darwin en la obra que aquí tratamos).

a aceptar bajo la categoría de *razonamiento*. Muchas veces, se define lo perteneciente al Sistema 1 de una manera similar a la que Darwin pensaba lo instintivo, como procesos, dijimos, más automáticos y con una mayor carga hereditaria, mientras que el Sistema 2 se acerca a lo que anteriormente se pensaba como lo propiamente racional (procesos más complejos, conscientes y culturalmente variables). Esta nueva distinción, a su vez, se vuelve a poner en cuestión hoy en día (Carruthers, 2009).

Respecto de muchas investigaciones contemporáneas en torno de los orígenes de la moral, también podemos señalar huellas darwinianas en algunas ideas importantes. Ya que Darwin considera que la sociabilidad es el principal factor que ha determinado y determina la evolución de nuestra especie, se afirma que él es también el padre de la teoría de la selección de grupos. Esto se debe a su tesis de que, finalmente, la simpatía, el altruismo y otras facultades de los seres altamente sociables, producen sociedades que prevalecen sobre otras donde las mismas cualidades se encuentran menos desarrolladas entre sus miembros (2009, pp. 427-429)⁵.

1.2. El desarrollo filogenético del cerebro social

Como vimos, Darwin propuso una historia evolutiva naturalista de las facultades sociales humanas. Hoy en día, gracias a los avances en la tecnología, las neurociencias constituyen una nueva entrada para la exploración de dichas facultades a partir del estudio del cerebro social, es decir, investigar aquellas partes del cerebro vinculadas al área social tanto a nivel estructural como funcional. Un estudio profundo de ello permite, además, vislumbrar aquellos sistemas que han evolucionado a lo largo del tiempo debido a las presiones de selección.

Barton (2007) presenta el recorrido por el cual se intentó analizar las diferencias cognitivas entre especies. El punto de inicio fue el estudio del tamaño cerebral total. En efecto, algunas investigaciones encontraron que la evolución del tamaño del cerebro correlaciona con ciertas especializaciones que probablemente responden a una selección de capacidades cognitivas. Sin embargo, otros estudios han obtenido resultados dispares entre sí y existe la posibilidad de que las diferencias en el tamaño de los cerebros se deban a diversas variables.

⁵ Como afirma Quintanilla (2008, pp. 5-6), Darwin no resuelve el problema acerca de cómo se desarrolló la competencia entre los individuos más altruistas y los más egoístas, interrogante que ha dado lugar al desarrollo de teorías como la de la selección de parentesco o del altruismo recíproco. A favor de Darwin podemos agregar que, en el capítulo V de *El origen del hombre* (p. 176), el autor adelanta una hipótesis acerca del motivo del desarrollo de las cualidades altruistas en los individuos: «En primer lugar, aumentándose de día en día las facultades de razonar y prever los sucesos, cada hombre se convenció bien pronto de que, ayudando a sus compañeros, los obligaba a pagarle después en igual moneda sus servicios». Esta afirmación se acerca mucho a la idea del altruismo recíproco.

Debido a ello se propuso otro método, un poco más minucioso, que consiste en investigar las variaciones de ciertos sectores del cerebro de manera independiente al tamaño total del mismo. A esta variación de regiones se le conoce con el nombre de evolución en mosaico y es producto de ciertas presiones de selección. Sin embargo, si bien ciertas subdivisiones se han agrupado y han evolucionado de un modo particular e independiente para cumplir una serie de funciones —como el cerebelo, la amígdala o el hipotálamo—, dentro de ellas existen grandes diferencias. Esto es, en cada subdivisión existen áreas que se encargan de funciones diferentes —por ejemplo, la amígdala se divide en el grupo basolateral, el núcleo centromedial y el núcleo cortical—, lo cual dificulta una comprensión detallada de la evolución de funciones específicas estudiando regiones cerebrales como un todo. Además, y esto es quizá lo más importante, las áreas de las subdivisiones presentan amplias conexiones con muchas otras regiones del cerebro. No se trata, entonces, de circuitos cerrados en cada subdivisión sino más bien de circuitos abiertos que interactúan continuamente con otros sectores.

En consecuencia, para determinar cuáles han sido los procesos cognitivos seleccionados por la evolución es necesario indagar sobre los sistemas neuronales funcionalmente integrados, que incluyen distintas áreas del cerebro como las que procesan información sensorial, motora y emocional. En este sentido, los procesos cognitivos no deben ser tratados como ajenos a estos últimos. Lo que se espera es que los componentes de cada sistema muestren una evolución correlacionada, es decir, que hayan cambiado de tamaño conjuntamente pero independientemente del cambio en otros sistemas.

Un ejemplo de lo anterior es el circuito ventral-prefrontal, en el que se aprecian grandes conexiones entre ciertas áreas del lóbulo temporal, la amígdala y la corteza prefrontal. La región temporal es esencial para la interacción social en tanto responde a caras de manera selectiva; la amígdala, por su parte, también responde selectivamente a caras y a señales emocionales como el temor y la rabia, además de a las expresiones faciales y movimientos corporales; y el área prefrontal se encarga de procesar la información obtenida de ambas fuentes. Así, este circuito procesa señales sociales en un contexto emocional.

Siguiendo con esta línea de pensamiento, Panksepp (2007) sostiene que el cerebro social humano surge de una serie de tendencias emocionales innatas que interactúan con procesos cognitivos de nivel superior. Lo interesante de su propuesta es que estas tendencias constituyen la base sobre la cual los procesos cognitivos superiores se despliegan. Hasta el momento se han identificado siete sistemas emocionales nucleares. Los cuatro primeros se encuentran en todos los vertebrados: el temor, la ira, la búsqueda y el impulso sexual. Ellos constituyen los principales elementos para

los sentimientos de ansiedad, rabia, deseo y erotismo en mamíferos. En estos últimos han surgido, además, los sistemas de cuidado, pánico y juego, que son las bases esenciales para los sentimientos de cuidado y crianza, tristeza y ansiedad de separación, y alegría social, respectivamente.

Estos afectos son *memorias de evolución* que permiten anticipar y afrontar con éxito las situaciones que amenazan o favorecen la supervivencia. No se trata, entonces, de un aprendizaje efectuado durante el ciclo vital sino de un aprendizaje a lo largo de generaciones que ha sido transmitido genéticamente para la subsistencia de ciertas especies que se desenvuelven en ambientes particulares. En humanos, estos afectos son luego analizados a la luz del aprendizaje social, lo cual da lugar a una complejidad emocional que no se encuentra en los animales. Es importante recalcar que no se está sugiriendo que los cerebros tienen contenidos específicos sino disposiciones que serán desarrolladas en el transcurso del desarrollo.

Las investigaciones sugieren que las capacidades socioemocionales más tempranas son la habilidad para sentir ansiedad cuando se detecta la pérdida de apoyo social —y comunicar vocalmente ese estado para incrementar la probabilidad de reunión— y la tendencia a sentir comodidad ante el apoyo social, lo cual se aprecia a través de sonrisas y risas. En edades tempranas, existe la tendencia a generar vínculos positivos a través de, por ejemplo, juegos físicos toscos. Estas emociones esenciales en la juventud —ansiedad de separación y disposición al juego— son fundamentales para la calidad de los vínculos con los cuidadores y, posteriormente, la amistad con otros, lo cual a su vez va sentando las bases para las habilidades sociosexuales. En cambio, la sexualidad podría haber servido como una precondition evolutiva a partir de la cual los impulsos maternos eventualmente surgieron. De este modo, el entendimiento de las relaciones sociales más complejas podría basarse en la comprensión de la ansiedad de separación y la disposición al juego de los más jóvenes, por un lado, y los impulsos sexuales y maternos de los adultos por el otro.

Mientras que Panksepp se concentra en las tendencias emocionales básicas, Jellema y Perrett (2007) se ocupan de algunos de los circuitos neuronales fundamentales para la cognición social, aquellos concernientes al procesamiento visual. Se trata de sistemas que proporcionan la base mecánica para la comprensión de las acciones de los demás, mas no dependen del entendimiento de las mentes de otros. Así, son capaces de vincular las acciones de las demás personas con claves situacionales de tal modo que permiten generar una relación causal y de predecir probables eventos o resultados gracias al acceso de información perceptual inmediatamente anterior. Es a partir de estos sistemas más básicos que se construirán otros de mayor complejidad que permitirán ya no solo predecir la conducta ajena sino explicarla de un modo más sofisticado.

Existen dos corrientes corticales distintas para el procesamiento visual: una que va desde la corteza visual hasta la temporal inferior y otra que también parte de la corteza visual pero se dirige a la parietal inferior. La primera se encarga de registrar la percepción visual y la segunda se ocupa de la acción, ambas haciendo uso de una codificación espacial. Primeramente el cuerpo es dividido en sectores claves como los ojos, la boca, la cabeza, el dedo, la mano y el brazo para luego ser analizados en términos de movimiento. Finalmente, en el surco temporal superior (STS) confluyen ambos circuitos para procesar el comportamiento de otros a partir de posturas y animaciones clave. Por ejemplo, algunas células del STS se activan al observar a un agente realizando el acto de alcanzar siempre y cuando la cabeza y los ojos estén apuntando a la misma dirección, lo cual ayuda a detectar acciones intencionales. Asimismo, otras células se activan frente a actos como recoger, rasgar y manipular.

Sin embargo, esta información necesita unirse a sistemas de mayor complejidad para atribuir estados mentales a los demás. Uno de estos, que se conecta estrechamente con el STS, es el sistema de las neuronas espejo, el cual permite entender el significado de las acciones mediante una simulación interna de las mismas. Las neuronas espejo son un tipo particular de neuronas motoras que se caracterizan por codificar actos motores que contienen una meta; por ejemplo, cargar, agarrar, manipular y rasgar. Como se verá más adelante, muchos consideran que a partir de este sistema es posible descubrir las intenciones de las demás personas mediante un proceso que permite experimentar los actos del otro como si fueran propios. Naturalmente, esta relación entre el STS de las neuronas espejo es explotada por quienes, como Alvin Goldman (2006), sostienen que la interpretación de otros agentes es básicamente un proceso de simulación (como se vio en la introducción a esta sección del trabajo, se trata de las teorías de la simulación o TS).

Este tipo de neuronas fue descubierto en monos macacos primero en el área F5 (una parte de la corteza premotora ventral) y, más adelante, en la parte rostral del lóbulo parietal inferior (IPL). En los humanos, la localización de las neuronas espejo es similar a la de los macacos; se encuentran en el lóbulo parietal inferior y en la parte baja del giro precentral (corteza premotora ventral), además de la parte posterior del giro frontal inferior. Así, las regiones parietal y frontal constituyen el núcleo del sistema de las neuronas espejo (Rizzolatti & Fogassi, 2007).

Dentro de las neuronas del área F5, existe un grupo que responde a estímulos visuales (neuronas visomotoras) el cual, a su vez, se divide en dos categorías: las neuronas canónicas y las neuronas espejo. Las primeras se activan ante objetos tridimensionales mientras que las segundas lo hacen al observar a un agente realizando una acción y cuando uno mismo realiza la misma acción o una similar. De este modo, existe una congruencia entre las respuestas visuales y las respuestas motoras.

El funcionamiento del sistema de neuronas espejo es como sigue: cada individuo tiene una representación de los actos motores, codificado en las cortezas parietal y frontal. Estas representaciones pueden ser evocadas internamente cuando una persona decide realizar un acto. En este caso el agente sabe (predice) el desenlace de la acción. Cuando la misma representación motora es activada al observar a otro individuo realizando ese acto motor, el observador sabe lo que el otro está haciendo porque la observación del acto activa, en su propia corteza parietal y frontal, la representación motora de esa acción. De esta manera, la acción de otro individuo es como si fuera propia y ello permite descubrir la meta detrás del acto (2007).

Dado que las neuronas espejo se activan frente a acciones con objetivos, algunos investigadores se preguntaron si entonces estas debieran activarse cuando el significado del acto fuera comprendido, aún en ausencia de información visual. Los experimentos realizados confirmaron esta hipótesis. En uno de ellos, la parte final de la acción fue ocultada tras una pantalla y, no obstante, las neuronas espejo se activaron. En otro experimento, se les permitió a los monos ver y oír cómo se rompía un pedazo de papel y, en un segundo momento, se los expuso solo al sonido del papel rompiéndose. Como resultado, las neuronas espejo audiovisuales, que responden a sonidos de acciones específicas, se activaron. Así, este sistema permite la comprensión de las acciones independientemente de si estas son vistas, escuchadas o realizadas.

La comprensión sobre cómo los individuos pueden tener información de las intenciones de los otros se encuentra en los circuitos neuronales. Las neuronas que codifican actos motores individuales —como sujetar— se encuentran insertas en circuitos previamente formados, cada uno de los cuales codifica una acción específica (las acciones son secuencias de actos motores individuales). De esta manera, las neuronas se vinculan con otras que codifican los próximos actos motores y, posiblemente, los facilita. Es el contexto el que da las claves sobre qué acción es la que se está realizando. En un experimento se observó que las neuronas espejo del lóbulo parietal inferior se activaban de manera diferenciada cuando el acto motor de agarrar pertenecía a distintas acciones —agarrar para comer y agarrar para colocar—. Las neuronas espejo, entonces, permitieron predecir el siguiente acto motor; esto es, la intención (2007).

La posibilidad de captar la intención de las demás personas es un aspecto fundamental para la interacción social. Si bien aún queda más por descubrir con respecto a las neuronas espejo, se sabe ya, por ejemplo, que el asco se encuentra dentro de este sistema. Es posible que diversas emociones también estén codificadas en el sistema de neuronas espejo, lo cual brindaría una comprensión más cabal de la empatía y daría cuenta de ciertas bases fundamentales para las relaciones interpersonales.

1.3. La atribución psicológica en el desarrollo social primate

Como vimos, desde sus orígenes la teoría de la selección natural ha servido, entre otras cosas, como paradigma para la investigación de la historia biológica de nuestras capacidades cognitivas y sociales. Ello implica que algunos aspectos de aquellas ya se encontraban, de forma quizás más elemental, en especies anteriores. Esto motiva a la primatología a producir y contrastar hipótesis sobre posibles capacidades compartidas con los grandes simios, pues se asume que, de las especies que viven en el presente, son las más cercanas a nosotros y, por tanto a nuestros antepasados. De hecho, uno de los textos que acuñaron el término ‘teoría de la mente’ es el ya clásico artículo «Does the Chimpanzee Have a Theory of Mind?» de David Premack y Guy Woodruff (publicado por primera vez en 1978). Mediante dicho término se hace referencia a la capacidad para atribuirse a uno mismo estados mentales y de, además, atribuirles a otros individuos estados mentales que podrían ser distintos a los de uno.

Posteriormente, en «Does the chimpanzee have a theory of mind? Revisited», Premack (1988) evalúa los métodos y los resultados de las investigaciones en torno a si aquellos primates superiores pueden, o no, realizar atribuciones mentales de manera similar a los humanos. El desarrollo de estudios semejantes implica resolver o, al menos, tomar partido respecto de dos cuestiones más básicas. La primera se refiere al rol del lenguaje en el desarrollo de la mente. Hoy en día, muchos autores proponen formas de lidiar con este problema. Por ejemplo, Peter K. Smith piensa que el único modo seguro de identificar la atribución de estados mentales depende eminentemente de un sistema simbólico como el lenguaje humano. Jim Hopkins sostiene que el surgimiento de las teorías de la mente se explica a partir de la forma en la que empleamos el lenguaje para describir el entorno, al punto que nuestra forma de hablar de los estados mentales es muchas veces análoga a la que empleamos para hablar de eventos físicos. Del otro lado del espectro, para Peter Carruthers la explicación tanto evolutiva como operativamente más económica de la atribución mental debe implicar la idea de que los mecanismos metarrepresentacionales (la capacidad de representarse las representaciones ajenas) pueden operar en instancias previas al empleo del lenguaje. Por su parte, Premack, como se verá con más claridad en lo que sigue, considera que es necesario distinguir entre formas de atribución, y algunas de ellas no dependen del lenguaje.

La segunda cuestión ocupa el centro de la atención de Premack a lo largo del escrito. Se trata del valor probatorio de los experimentos en circunstancias controladas, frente al valor de las observaciones hechas en hábitats naturales (1988, p. 161). Cabe preguntarnos si los experimentos acaso no generan circunstancias tan poco naturales que las respuestas de los individuos así estudiados corren el riesgo de no explicar nada que ya haya sido puesto previamente por quienes indagan.

Un tipo de caso registrado que suele vincularse con la tesis de que los chimpancés podrían poseer alguna especie de teoría de la mente es el engaño. Lo interesante del caso radica en que estos primates son capaces de engañar en una gama muy amplia de circunstancias. El ejemplo de engaño que emplea Premack es el de la omisión del gemido que los chimpancés macho suelen emitir cuando copulan con su hembra preferida. Ello tiene como objeto no llamar la atención de otros machos para no tener que compartir la hembra. Ahora bien, el problema consiste en determinar si el individuo que emplea dicha estrategia lo hace teniendo en cuenta que con ello evita que los demás *sepan* algo y, por tanto, evita que actúen de cierta manera no deseada, o si meramente busca evitar que los demás *hagan* algo. Solo el primer escenario instancia necesariamente una teoría de la mente por parte del chimpancé, mientras que en el segundo bastaría con una explicación conductista. El autor niega que para poder discernir entre ambas posibilidades ayude recurrir a información obtenida en otras experiencias de campo, pues la nueva información también se podría interpretar de dos maneras distintas. Aquí es donde la experimentación controlada resulta ser particularmente importante. En opinión de Premack (1988, p. 171) (y, como diremos más adelante, de Whiten también) únicamente aquella nos permite realizar distinciones conceptuales finas. Frente al contraargumento del condicionamiento que podría tener lugar en un laboratorio, factor que podría volver la prueba demasiado forzada, este responde que tampoco en el campo existen garantías que protejan a los individuos contra ello.

Gracias a la experimentación en psicología del desarrollo, se sabe que los niños de tres años ya son capaces de atribuir estados mentales a otros agentes y, sin embargo, no comprenden aún la relación entre sensaciones y creencias, en particular cuando se trata de los demás (Premack, 1988, p. 164). Al parecer, esta capacidad se adquiere cabalmente a partir de los cuatro años de edad, cuando ya se posee una teoría de la mente avanzada. Ya que los chimpancés fallan pruebas sobre el papel de las sensaciones en la adquisición de conocimiento que niños de tres años y medio tampoco aprueban, Premack piensa que se puede postular como regla general que si un niño de dicha edad no puede hacerlo, entonces un chimpancé tampoco. No obstante, ciertos experimentos parecen demostrar que algunos chimpancés sí son capaces de predecir el comportamiento de otros agentes o emplear tácticas en su beneficio basándose en lo que podría ser alguna forma de conexión del comportamiento y las sensaciones. Para explicar estos contraejemplos, piensa el autor, debemos afinar la idea de atribución de estados mentales.

Según Premack, resulta necesario distinguir entre dos tipos de estados mentales, unos más complejos y otros menos. Los estados más simples son automáticos, encapsulados y poco variables. Las sensaciones se encontrarían dentro de este grupo.

La creencia, de otra parte, parece ser del tipo complejo, pues no solo es altamente variable, no automática, ni encapsulada, sino que, además, se representa mediante el lenguaje. La creencia también permite poner en duda los estados sensoriales que de otra manera estarían ligados directamente a la acción. Cuando vemos a un individuo humano vacilar, pensamos inmediatamente que, por ejemplo, se está preguntando si realmente ve lo que cree ver. Se tiene registro de vacilación en chimpancés, pero ello podría tratarse del resultado de un condicionamiento mixto, es decir, de que los primates hayan estado expuestos a experiencias diversas con resultados unas veces positivos y otras veces negativos respecto de la misma cosa. En general, las probabilidades de encontrar una explicación conductista que reemplace una postura mentalista respecto de la tenencia de las creencias simples son bastante altas.

También poseemos una gran cantidad de creencias que admitimos como posibles explicaciones sin que necesitemos recurrir, al menos directamente, a datos sensoriales. Premack (1988, p. 174) cita ejemplos como la idea de que la mente está en el cerebro, los gérmenes causan enfermedades o que las especies evolucionan. Estas creencias parecen depender de dos factores: una competencia lingüística considerable y la capacidad de realizar explicaciones causales de manera consciente. En estos casos, la brecha entre humanos y el resto de animales parece infranqueable. No obstante, el autor considera que este es un campo abierto a la investigación. Un ámbito particular donde se puede profundizar en esto es en la manera como humanos y otros primates generan (o no) creencias y expectativa respecto de la conducta de los demás. En los primeros, un encuentro con otro individuo puede producir predicciones de dos tipos: de una parte, predicciones basadas en el mero condicionamiento, de manera que frente a un rango restringido de casos se tiene la *expectación* de que el otro haga algo; de otra, los humanos también, a raíz del encuentro, adoptamos *creencias* acerca de la personalidad del otro en general, lo cual nos permite realizar predicciones en un rango bastante más amplio de contextos posibles. La experimentación ha mostrado que los chimpancés jóvenes pueden llegar a hacer predicciones basadas en la expectativa, mas parecen carecer de juicios más generales, aquellos del segundo tipo.

Hasta aquí, pareciera que los chimpancés solamente son capaces de atribuir estados mentales simples, como la expectativa o los otros estados sensoriales. De hecho, algunos experimentos muestran que estos primates no son capaces de distinguir entre sus propios estados mentales, como lo que ellos ven, y los estados de otros individuos. Esto contrasta notoriamente con la facilidad con que los niños de cuatro años aprueban muchas versiones del test de la falsa creencia, el cual es la principal herramienta para determinar la tenencia de teoría de la mente en humanos. La idea básica, desarrollada por Wimmer y Perner (1983), consiste en colocar a los infantes frente a una situación en la que deberían atribuirle a otro agente creencias basadas

en una percepción que ya no corresponde al estado de cosas actual. Por ejemplo, en la primera versión de la prueba (*The Maxi Task*) se les contaba un relato a los niños, en el cual el personaje principal, Maxi, colocaba su barra de chocolate en un lugar específico y salía a jugar. Luego, su madre cambiaba el chocolate de lugar y Maxi regresaba a buscarlo. Se les preguntaba a los niños dónde creían que Maxi buscaría su chocolate, si acaso lo haría donde originalmente lo dejó o en el nuevo lugar al que su madre lo movió. Antes de los cuatro años, los niños suelen errar la respuesta al afirmar que Maxi buscará el chocolate donde ellos saben que está (en el nuevo lugar asignado por la madre), en vez de responder poniéndose en el lugar del otro agente. No obstante, como hemos visto, la cuestión por la atribución psicológica en primates no humanos sigue abierta, y el refinamiento en los métodos de experimentación controlada resultará decisivo para el progreso de la investigación.

A pesar de lo dicho arriba respecto de las dudas en torno de la tenencia plena de teoría de la mente por parte de primates no humanos, Robin Dunbar (1988) piensa que se les puede atribuir intencionalidad de segundo orden, lo cual quiere decir que cuentan con estados mentales de la forma «yo *creo* que tú *crees* que...», en la cual el verbo mental («creer», en este ejemplo) es asignado en dos niveles distintos («yo» y «tú»). No obstante, los humanos podemos llegar considerablemente más lejos en el conteo de asignación de estados mentales. Si bien no existe consenso en torno al número máximo de niveles a los que la intencionalidad humana puede llegar, las pruebas indican que entre el cuarto y el quinto nivel de asignación de creencias la tasa de errores de los individuos experimenta un crecimiento dramático de entre 5% y 10% para el cuarto nivel a alrededor de 60% para el quinto nivel.

Según Dunbar, deberíamos estar en condiciones de elaborar un relato evolutivo que dé cuenta de la diferencia al menos cuantitativa entre los niveles de intencionalidad humanos y los de otros primates. La diferencia, además, parece ser compatible con algunos datos obtenidos de la comparación de la constitución del cerebro en monos y homínidos. Aparentemente, el tamaño de la sección del cerebro de los monos dedicada a procesar información visual, «V1», crece en paralelo con las demás áreas del neocórtex, dedicadas al procesamiento de, entre otras cosas, modos reactivos del comportamiento. Sin embargo, la proporción cambia en el caso de los homínidos, en los que las otras áreas del neocórtex se desarrollan bastante más que V1. Esto, a su vez, tiene un correlato en el comportamiento. Mientras los monos pueden leer y predecir el comportamiento de otros individuos a partir del procesamiento de *inputs* visuales, no son capaces de «leer la mente detrás del comportamiento» (1988, p. 245), cosa que los homínidos sí parecen poder (ya vimos que el ser humano en grados más elevados que el resto de las especies emparentadas).

La habilidad de lectura de mentes de los homínidos, entonces, parece depender del tamaño del neocórtex. La complejidad de un cerebro apto para producir teorías

de la mente refinadas es bastante alta, lo cual trae consigo costos biológicos elevados para mantener semejante órgano en funcionamiento. Por tanto, desde un punto de vista evolutivo, estas características cerebrales deberían haber sido seleccionadas directamente debido al gran beneficio que traen consigo para sus especies portadoras. Dunbar y muchos otros ven una correlación entre el desarrollo del cerebro en los homínidos y el número de individuos que componen sus respectivos grupos sociales. Como ya debe sospecharse, dada la función social de la lectura de mentes, es este dispositivo y otros semejantes los que se piensa que contribuyen a la organización de sociedades numerosas.

Casi todos los elementos del relato sobre la evolución de la capacidad para la teoría de la mente están listos; solo hace falta identificar la presión selectiva que impulsó el desarrollo de sociedades de homínidos cada vez más numerosas. Dunbar (1998) piensa que fue probablemente la necesidad de encontrar y mantener acceso al agua lo que ejerció la presión necesaria para la evolución de habilidades sociales más complejas. Esta presión habría ido en aumento desde que, con la aparición del *Homo erectus* (hace aproximadamente 1.5 millones de años), nuestros antepasados homínidos empezaron a desplazarse grandes distancias. Entonces habría sucedido un incremento en el tamaño del neocórtex y un correspondiente incremento en el número de integrantes de los grupos, excediendo el límite superior de las sociedades actuales de grandes simios. De la misma manera, se estima que es con la aparición de las primeras formas de *Homo sapiens* (hace aproximadamente 120 000 años) que se dio otro salto, tanto respecto de la constitución del cerebro como en la cantidad de integrantes de cada grupo.

1.4. La hipótesis de la inteligencia social

Naturalmente, poder predecir las acciones de los demás con vistas a obtener el mayor beneficio posible para uno, resulta de gran utilidad adaptativa. Esta capacidad, la *inteligencia maquiavélica*, se ha desarrollado entre los humanos gracias a la manera en que usualmente predecimos el comportamiento de los demás, es decir, gracias a que atribuimos estados mentales a otros agentes. Por ello, para poder determinar si otros primates poseen también inteligencia maquiavélica, debemos responder primero si es que son capaces de leer mentes, al menos en un nivel elemental.

1.4.1. Inteligencia maquiavélica y niveles de atribución psicológica

Aparentemente, las pruebas de la falsa creencia sirven como el mejor criterio para determinar si el individuo ya posee mentalismo considerablemente desarrollado o no. Como vimos, Premack (1988) sostiene que los experimentos en chimpancés solo alcanzan para mostrar que, dada una secuencia de acciones de un agente, los primates

pueden predecir correctamente la acción siguiente, pero porque es lo que ellos mismos harían. El mentalismo que buscamos requiere, en cambio, que el individuo sea capaz de adscribirle a otro agente estados mentales distintos de los suyos. A pesar de estos resultados, Andrew Whiten (1997) piensa que estudios realizados acerca de la atribución de ciertas clases de estados mentales en niños de edades anteriores a la requerida para aprobar las test de la falsa creencia, resultan de particular interés en la investigación del maquiavelismo de primates no humanos. Cabe recordar lo que dijimos que Premack (1988, p. 146) consideraba una regla no estricta pero útil: si el niño de tres años y medio no puede, tampoco puede el chimpancé. Pero, en el caso contrario, si el niño pequeño puede, entonces podría ocurrir que otros primates también.

Ya que se puede rastrear formas de atribución mental previas al mentalismo avanzado requerido para aprobar las pruebas de la falsa creencia, Whiten (1997, p. 150) sostiene que es preferible pensar la lectura de mentes gradualmente. Así, distingue entre cuatro posibles tipos. La forma más elemental sería aquello que llama *lectura implícita*, y se refiere a la capacidad de identificar un nexos entre la percepción de un agente y la acción, relación mediada, no explícitamente, por una mente. Un segundo tipo es el *contraengaño*, que consiste en la habilidad para reconocer las intenciones detrás de las acciones realizadas en falso por un agente. El *reconocimiento de variables intermedias* implica saber que ciertos comportamientos generan estados mentales en los demás agentes, lo cual, a su vez, permite realizar predicciones. Finalmente, la *proyección de la experiencia*, como el nombre sugiere, consiste en saber utilizar la propia experiencia en la predicción del comportamiento de otros. Whiten mismo cree que los últimos dos tipos de mecanismos son indicios más seguros de mentalismo, aunque, como dijimos, las otras opciones no pueden ser descartadas sin más.

Ya los niños de casi tres años emplean mayoritariamente explicaciones causales mentales para describir sus relaciones con el entorno. Y no solo ello, sino que dichas explicaciones son empleadas, más que en cualquier otra circunstancia, para extraer beneficios personales. De hecho, se ha registrado que son los niños que más hábilmente emplean su capacidad de lectura de mentes en contextos maquiavélicos los que posteriormente poseerán una teoría de la mente más desarrollada (Whiten, 1997, p. 153).

El maquiavelismo precoz de los niños parece contradecir los resultados de las pruebas de la falsa creencia. Además, experimentos controlados demuestran que los niños pequeños sí son engañadores capaces. El problema de lo último radica en que también se sabe, mediante experimentación, que los niños de tres años, a pesar de su capacidad para ciertas formas de engaño, no extraen todas las consecuencias mentalistas que se esperaría de alguien que comprende cabalmente en qué consiste poseer una representación falsa (como la que ellos mismos logran crear con sus

tácticas engañosas). Al parecer, la razón de ello radica en que los pequeños de tres años aún no poseen un concepto completo de creencia, y sus estrategias de engaño, en el fondo, son intentos por hacer que el agente a quien quieren aventajar participe en una realidad pretendida, para lo cual no es necesario que comprendan en qué consiste creer falsamente (Whiten, 1997, p. 156). Así, los contextos maquiavélicos tempranos sirven para desarrollar y afianzar mecanismos necesarios para el surgimiento del mentalismo maduro. Tal vez aquí es donde puede encontrarse un paralelo con el maquiavelismo de primates no humanos.

No se tiene evidencia concluyente sobre hasta qué punto los niños pequeños y los primates no humanos distinguen entre la realidad percibida y los estados mentales de los demás. En los niños, el problema se resuelve a partir de los cuatro años. No obstante, otros primates también exhiben comportamiento engañoso. Al igual que Premack, Whiten se pregunta si dicho comportamiento no se puede explicar recurriendo al mero conductismo. Mas en torno a este punto, nuevamente se puede establecer un paralelo entre lo encontrado en psicología del desarrollo con relación a los niños y algunos datos experimentales de la primatología. Nos referimos específicamente a la capacidad de leer la atención que otros agentes le prestan al entorno compartido. Si bien la información que se tiene acerca de ello en chimpancés y otros primates no es consistente, Whiten enfatiza el hecho de que en ciertas pruebas los chimpancés intentan trasladarse físicamente a un punto donde puedan visualizar qué cosa está observando otro agente. Esto podría indicar que, de hecho, los individuos así estudiados saben que la visión es siempre *acerca de algo*. La intencionalidad de la visión es considerada, además, como una posible primera puerta de los infantes hacia la teoría de la mente, pues es esencial en su aprendizaje sobre cómo nombrar objetos mediante la atención que le prestan al mismo acto en sus cuidadores (1997, pp. 164-165). Por motivos similares, el autor también considera que resulta de igual importancia para el estudio de la inteligencia maquiavélica determinar en qué medida los primates no humanos pueden distinguir entre posibles voliciones de otros agentes.

Las conclusiones a las que puede llegar Whiten luego de considerar toda la información mencionada, son prácticamente las mismas a las que llega Premack. La investigación sobre la atribución mental en primates no humanos se enriquece mucho con lo que la psicología del desarrollo descubre en los infantes, pero aún quedan muchos cabos sueltos por resolver. Esta tarea se debe realizar, sobre todo, con la ayuda de la experimentación controlada. Sin embargo, a pesar del estado aún embrionario de esta empresa científica, Whiten también considera que, en un sentido mucho más simple que en el relativo a los humanos, se aprecia en otros primates alguna forma de atribución mental. O, al menos, se aprecian mecanismos necesarios para el desarrollo posterior de teorías de la mente.

1.4.2. Complementos a la hipótesis de la inteligencia social

Si bien la hipótesis de la inteligencia social ha abierto un amplio campo de investigación evolutiva, podemos preguntarnos sobre sus límites explicativos (Byrne, 1997). Podríamos pensar que dicha idea basta para explicar las diferencias en la evolución de las capacidades cognitivas de ciertos primates, incluidos los humanos. Pero también puede pensarse que la teoría no es suficiente y, entonces, sea necesario indagar otras causas complementarias del fenómeno evolutivo.

Como vimos al discutir algunas tesis de Dunbar, parece haber una correlación entre el incremento del tamaño del *neocórtex* y la complejidad de la vida social de las especies. Naturalmente, la hipótesis de la inteligencia maquiavélica se ve reforzada por esta evidencia. Específicamente, datos de este tipo se ajustan muy bien en el caso de las diferencias entre los primates haplorrinos, como los grandes simios o los monos, y los estrepsirrininos, como los lémures. No obstante, la relación entre los datos y la teoría no siempre es armoniosa. Al interior de la rama de los haplorrinos hay diferencias importantes: existen diferencias notorias entre las capacidades cognitivas de simios y monos, pero las sociedades de los primeros no son necesariamente más complejas ni más grandes que las de los segundos. De hecho, un aspecto importante de la postura de Dunbar era la proporción del tamaño del neocórtex con relación al resto del cerebro, mas dicha proporción no varía mucho entre los haplorrinos (Byrne, 1997, pp. 294-295).

Byrne sostiene que las principales diferencias cognitivas entre los grandes simios y otras especies del mismo suborden no se manifiestan únicamente en las habilidades sociales. Sin embargo, sí resulta relevante notar que, por ejemplo, el engaño, estrategia tan importante para el estudio de la inteligencia maquiavélica, aunque presente también en monos, es ejecutado de manera cualitativamente distinta por los grandes simios. En estos parece haber cierto grado de planificación que, como también vimos en secciones anteriores, podría indicar la presencia de alguna forma débil de atribución psicológica.

De manera similar al engaño y otras estrategias sociales, si bien tanto monos como simios superiores son capaces de utilizar ciertas herramientas físicas elementales, los primeros no parecen ser capaces de representarse específicamente qué características del objeto seleccionado son relevantes para aquella actividad (usualmente extracción de material alimenticio) en la que lo emplean. Los segundos, en cambio, aparentemente sí son capaces de hacer este tipo de distinciones finas desde el momento mismo de la selección de los objetos (Byrne, 1997, p. 296). Otra habilidad que llama la atención de los estudiosos, y que se relaciona con el uso de herramientas, es la construcción de camas o nidos por parte de gorilas (de manera más simple), y orangutanes y chimpancés (con mayor complejidad). Curiosamente, esta habilidad está presente en otros primates, pero no en haplorrinos.

Byrne (1997, p. 297) también menciona la capacidad de imitar como factor diferenciador entre los grandes simios y otros primates. Incluso dándole una interpretación modesta (que no recurre a la postulación de tenencia de alguna forma de teoría de la mente), la imitación parece requerir de la capacidad de representación del comportamiento en general, lo cual solo parece presente, nuevamente, en los primeros.

Todo indica que aquello que distingue a los grandes simios es una cierta capacidad de abstracción. Sus respuestas más finas a ciertos problemas parecen implicar, al menos, que los individuos estudiados pueden simular escenarios posibles. Estas facultades, además, no se limitan a la esfera de los problemas sociales, esfera en la que los monos también son bastante hábiles. Surge, entonces, la pregunta por las presiones selectivas que distanciaron cognitivamente a los grandes simios del resto de primates.

Algunas posibles explicaciones de la evolución peculiar de los grandes simios se centran en sus limitaciones motrices en tanto braquiadores, si se tiene en cuenta que los otros primates son considerablemente más ligeros. Así, la selección natural habría favorecido en aquellos una mayor capacidad de procesamiento de datos espacio-temporales o, en otras palabras, un mapa cognitivo más complejo (Byrne, 1997, pp. 298-299). Arriba hablamos acerca de la capacidad de simulación, la cual resulta particularmente útil considerando que los grandes simios, dado su peso, corren bastante más riesgo de caer que los otros braquiadores arbóreos más ligeros. Naturalmente, la simulación permitiría al individuo ponerse en casos posibles y anticipar sus propios movimientos, reduciendo el riesgo de caída. Por ello, la simulación del movimiento propio es otra posible capacidad que determinó la evolución cognitiva de los grandes simios (p. 300). De otra parte, su tamaño también podría haber influenciado en la evolución de sus capacidades representacionales, imponiendo la necesidad, ausente en el resto de primates, de desarrollar técnicas sofisticadas de construcción de camas en los árboles (posiblemente en un antepasado común a los grandes simios contemporáneos). Esta inteligencia técnica básica luego podría haberse desarrollado en otros ámbitos de la vida de dichos simios, como se observa hoy.

Se han desarrollado teorías en torno a la sorprendente capacidad de los grandes simios para emplear herramientas que los ayudan a extraer alimentos. Dichas teorías, sin embargo, se topan con el problema de que el uso de herramientas también está presente en algunas especies de monos. Byrne (p. 303) piensa que los problemas se disipan si nos enfocamos, más allá del mero hecho del empleo de herramientas, en la manera en la que los individuos lo hacen. De hecho, si el empleo de herramientas en los grandes simios responde a esquemas sistemáticamente organizados de acción, entonces cabría esperar esquemas semejantes en actividades en las que se prescindiera de herramientas. Al parecer, sí puede observarse comportamiento organizado de este tipo en gorilas y orangutanes, específicamente en sus técnicas manuales para la obtención de alimento.

En el caso de los monos, sus técnicas manuales para la obtención de alimento no parecen responder a una organización sistemática. Por el contrario, estas parecen ser producto de una rápida capacidad de aprendizaje volcada sobre ciertos estímulos repetidos. Los grandes simios, en cambio, presentan secuencias de acciones organizadas jerárquicamente de acuerdo con objetivos primarios y subordinados. Este refinamiento en las técnicas de obtención de alimento parece denotar la capacidad de representación del comportamiento, y se debería a las exigencias alimenticias que mamíferos grandes con estómagos no particularmente resistentes deben satisfacer en hábitats muchas veces complejos.

La propuesta tentativa de Byrne sería, sintéticamente, que algún antepasado común de los grandes simios (probablemente similar a los gorilas contemporáneos) desarrolló la capacidad de organizar distintos aspectos de su comportamiento de acuerdo con planes bastante flexibles que le permitan superar limitaciones físicas no presentes en otros primates. Esto último, evidentemente, implica cierto nivel de entendimiento representacional. Ello podría ayudar a explicar, además, el incremento en el tamaño absoluto del cerebro de los grandes simios, ya que las exigencias computacionales de un nivel semejante de planificación son bastante más costosas que en el comportamiento de, por ejemplo, los monos.

1.5. Evolución de la lectura de mentes, la autoconciencia y la agencia

En años recientes, las teorías que explican la evolución de la cognición social también han sido empleadas como nuevas formas de enfrentar problemas clásicos de disciplinas como la filosofía y la psicología. En las siguientes secciones, trataremos dos casos interesantes: el de la explicación filogenética de la relación que existe entre el conocimiento de nuestras propias mentes y de las mentes de los demás, y el problema de explicar nuestra experiencia como seres capaces de modificar su propio comportamiento voluntariamente.

1.5.1. La relación de la lectura de mentes y la metacognición. Explicaciones filogenéticas

Por lo general, en circunstancias normales, los seres humanos pensamos que poseemos la facultad de acceder directamente a nuestros propios estados mentales. Aquí, *acceder directamente* quiere decir que el conocimiento que poseemos sobre lo que ocurre en la mente de uno mismo es considerablemente menos falible que aquello que conocemos perceptualmente. Por ejemplo, los estados mentales que creemos leer en el comportamiento de otros agentes, en relación con los cuales, por ello mismo, tan solo seríamos intérpretes. Pero, si pensamos que es gracias al mismo dispositivo, alimentado por fuentes similares de información, que interpretamos a los demás

y conocemos nuestras propias mentes, la asimetría entre la falibilidad de ambas actividades se disipa. Esta última es la tesis de Carruthers (2009), quien considera que tanto la metacognición como la lectura de mentes son actividades del mismo dispositivo metarrepresentacional.

A la idea del acceso directo a las propias representaciones se le suele llamar *introspección*, y es este mismo sentido del término el que Carruthers intenta abandonar. Según su modelo, todo lo que se necesita para explicar la metacognición es que el dispositivo metarrepresentacional, cuya función primaria es la lectura de mentes, tenga acceso a los datos de salida (*outputs*) de todos los otros dispositivos cognitivos que procesan la información perceptual y cuasi perceptual (como los productos de la imaginación visual o el discurso interno) de primer orden. Esto quiere decir que el dispositivo de lectura de mentes interpreta el comportamiento de los agentes sobre la base de los conceptos de primer orden que los otros dispositivos han utilizado. No obstante, jamás se tiene información perceptual sobre el proceso de generación de juicios. Vemos, por ejemplo, a un agente caminar, y el dispositivo de lectura de mentes recibe la información perceptual ya procesada con conceptos como «persona», «caminar» y todos los que sean necesarios, pero no recibe ningún dato sobre las creencias que impulsan la acción del caminante. Exactamente lo mismo ocurre en el caso de la primera persona: tampoco poseemos acceso consciente a los procesos de generación de nuestros propios juicios y decisiones. Justamente, este es el espacio que le corresponde completar al dispositivo metarrepresentacional.

El caso de la interpretación de uno es, prácticamente, el mismo. La única diferencia, cuando la hay, consiste en que la mayor parte del peso de los datos perceptuales que el dispositivo metarrepresentacional debe interpretar recae en aquello que llamamos más arriba *cuasi perceptual*. De esta manera, puede que la atención que le prestemos al entorno en un momento determinado sea mínima, pero seguimos produciendo expresiones verbales internas e imágenes visuales basadas en recuerdos. Esta información cuasi perceptual será la que sirva como principal punto de apoyo para la labor del dispositivo metarrepresentacional, con lo que surge la metacognición. Así también se explica la idea de la introspección. Es este enfocarnos en los datos cuasi perceptuales durante la metacognición lo que nos hace pensar que ello depende de un mecanismo esencialmente distinto a cualquier otro que emplee información del entorno.

De todo lo anterior se siguen consecuencias evidentes. La metacognición resulta ser tan interpretativa como la lectura de mentes, pues se trata de dos caras de un mismo mecanismo. Por lo mismo, ya no es necesario postular un dispositivo especial de introspección que garantice la metacognición; aquella actividad es, en el fondo, uno de los conceptos que conforman la teoría *folk* de nuestras mentes, pero no es una función real. Se predice, además, que somos altamente confabuladores con nosotros mismos.

Esto quiere decir que, en aras de un funcionamiento más eficiente, el dispositivo metarrepresentacional puede interpretar en falso el comportamiento de uno mismo (de hecho, acabamos de afirmar que la idea misma de un mecanismo encargado de la introspección es falsa, aunque computacionalmente provechosa).

Como vimos en la introducción a este capítulo, Carruthers piensa que la refutación final de las teorías que dependen de la existencia de un mecanismo introspectivo, como las teorías de la simulación, debe realizarse en el terreno de las explicaciones filogenéticas de los mecanismos sobre los que se discute. La razón es que, aparentemente, la información proveniente de los estudios ontogenéticos de nuestros mecanismos para la interpretación del comportamiento de los agentes puede ser interpretada desde más de un modelo, ya sea uno afín a las TT o a las TS.

El argumento filogenético a favor del modelo que postula un solo mecanismo metarrepresentacional para la metacognición y la lectura de mentes es sumamente sencillo. La idea consiste en que, desde la perspectiva de la teoría de la evolución, resulta mucho más económico postular una sola historia evolutiva para más de una función cognitiva. Biológicamente, la generación y el mantenimiento de los mecanismos implicados en la atribución psicológica son sumamente costosos. Si se postula un mecanismo metacognitivo (con capacidades introspectivas) aparte de uno de lectura de mentes, es decir, de atribución en tercera persona, se deben postular dos historias evolutivas distintas para cada mecanismo, cada uno biológicamente muy costoso. Ahora bien, otra opción, entre tantas, es sostener que solamente se requiere de un tipo de mecanismo que se encargue de todas las formas de atribución psicológica, pero que este mecanismo fue originalmente diseñado para la metacognición. Esta es la postura de Alvin Goldman (2006), defensor de las TS. Contra Goldman, Carruthers argumenta que una tesis semejante depende también de al menos dos historias evolutivas diferentes. De una parte, se tendría que dar cuenta de qué presiones evolutivas actuaron en la selección de un mecanismo metacognitivo anterior e independientemente de la metarrepresentación; de otra, se tendría que ofrecer una explicación de bajo qué presiones evolutivas aparecieron mecanismos que permitan identificar la información perceptual de otros agentes con las propias representaciones metacognitivas.

Según Carruthers, el modelo de un solo mecanismo metarrepresentacional es el que mejor compatibiliza con la hipótesis de la inteligencia maquiavélica, tal y como es postulada por autores como Premack o Dunbar. La complejidad de la vida social de nuestros antepasados homínidos parece ser la mejor alternativa de presión evolutiva para explicar la selección de nuestras capacidades metarrepresentacionales. Por ello, la historia evolutiva que imagina Carruthers parte de la selección de un mecanismo encargado de la lectura de mentes. Es decir, de un mecanismo metarrepresentacional *seleccionado para* la lectura de mentes. La metacognición sería,

entonces, un producto colateral de la evolución del dispositivo de lectura de mentes. De esta manera, obtenemos una misma historia evolutiva para explicar dos facultades cognitivas diferentes.

1.5.2. El problema de la causalidad mental

Propuestas como la de Carruthers no solo ponen en tela de juicio conceptos como el de introspección. También, ponen en duda otras intuiciones típicas de nuestra *psicología folk*, ciertamente vinculadas con el concepto de introspección, como la idea de que muchas veces nuestro comportamiento es causado por nuestra voluntad luego de un proceso deliberativo. Normalmente, creemos que tomamos decisiones y actuamos acorde a ellas. Pero, si la vía de acceso al conocimiento de la primera persona es casi tan falible como la del conocimiento de la tercera persona, si constantemente confabulamos contra nosotros mismos, entonces podemos dudar de nuestra experiencia cotidiana de la voluntad como causa de muchos actos.

Lo dicho arriba es una expresión contemporánea de un clásico problema filosófico: el de si es posible justificar alguna forma de compatibilidad entre la causalidad de los fenómenos naturales, tal como los describimos en las ciencias, con la idea común de que podemos ejercer un control consciente sobre nuestros actos. Quienes defienden que los eventos mentales, como las creencias y los deseos, sí se pueden pensar como causas legítimas de eventos físicos, deben demostrar que los nexos causales descritos en términos mentales no son reductibles a descripciones estrictamente físicas. Este tipo de defensas van de la mano de explicaciones sobre la eficacia de las descripciones mentales del comportamiento. En lo que sigue, veremos dos ejemplos recientes de dicho modo de defender la causalidad mental.

Si bien el fisicalismo (la idea de que toda propiedad y evento del universo son, al menos en una de sus posibles descripciones, propiedades y eventos físicos) es la tesis ontológica más aceptada en la filosofía contemporánea, pareciera que si se le lleva a sus últimas consecuencias nos obliga a renunciar a otras tesis que suelen considerarse importantes para explicar ámbitos específicos de la realidad. En particular, pareciera obligarnos a aceptar que todo sistema no descrito en términos de las leyes de la física es reducible, en principio, a esta. Una de las formas de señalar el problema es mediante la discusión de la idea de *superveniencia* de propiedades no físicas a propiedades físicas y su efectividad o irrelevancia causal. En filosofía, el término *superveniencia* se utiliza para hablar de la dependencia de ciertas propiedades respecto de otras más básicas. Así, por ejemplo, si admitimos, como sería consecuente con el fisicalismo, que las propiedades psicológicas de un agente *supervienen* a propiedades físicas, cualquier cambio en las primeras implica un cambio correspondiente en las segundas. Si aceptamos la superveniencia de lo psicológico sobre lo físico, entonces al considerar

que una propiedad psicológica cualquiera, digamos p , causa la aparición de otra propiedad psicológica p' ; y aceptamos que p superviene a una propiedad física f y p' a una propiedad física f' ; entonces la descripción del nexo causal entre p y p' puede reducirse a la descripción del nexo entre f y f' . En términos menos formales, al reduccionismo fiscalista subyace la idea de que, puesto que toda propiedad no física superviene a alguna propiedad física, toda descripción no física es, en principio, reducible a otra física y las relaciones causales no son una excepción a la regla.

Como notan El-Hani y Pereira (1999), el fiscalismo reduccionista entra en conflicto con una creencia cuya verdad resulta difícil de cuestionar: que los sistemas que explican distintos niveles de la realidad, como aquellos que componen las teorías de la biología o de la psicología, deben su relevancia a que responden a un dominio propio de cuestiones, por lo que debería garantizarse su independencia causal respecto de las propiedades meramente físicas. Por ello, los autores mencionados se proponen demostrar que el concepto de *emergencia* de niveles de organización más compleja de propiedades a partir de niveles menos complejos resulta útil para solventar un fiscalismo no reduccionista.

Los sistemas biológicos, por ejemplo, emergen de sistemas físicos y químicos menos complejos. En un sentido, existe continuidad entre los niveles, pues en todos casos se trata de los mismos componentes materiales. Las partículas que estudia la física son las mismas que conforman los organismos que estudia la biología. De hecho, el concepto de emergencia no puede desligarse de la idea de que los distintos niveles de organización han ido surgiendo en una misma historia evolutiva. No obstante, a pesar de poseer el mismo sustrato material, entre los distintos niveles existe una discontinuidad de propiedades, correspondiente a sus diferencias en la complejidad de organización. De ahí que el concepto de emergencia sea compatible con el de *dualismo de propiedades* (o, quizá, sea más pertinente hablar de un *pluralismo* de propiedades) (El-Hani & Pereira, 1999, p. 337).

El-Hani y Pereira (1999, p. 340) sostienen que el problema de la irrelevancia causal de las propiedades supervenientes se mantiene si es que solo se admite la idea de causalidad eficiente. Piensan que el problema puede ser afrontado si se toma en cuenta una reinterpretación contemporánea de la concepción aristotélica sobre los cuatro tipos de causalidad. De esta manera, se piensa la causalidad funcional, correspondiente a la causalidad final aristotélica, como referida a la función que un elemento juega dentro de su correspondiente sistema con relación a la mantención de su estabilidad. Asimismo, la causalidad formal aristotélica se reinterpreta en términos de los patrones de coordinación que los elementos de un sistema adoptan. La causalidad material se refiere, ahora, a las propiedades de dichos elementos, propiedades que, hemos visto, supervienen sobre las de elementos de un nivel inferior.

Lo anterior nos debería permitir evitar comprender unidireccionalmente las relaciones causales entre los distintos niveles organizativos (unidireccionalmente, pues los nexos causales de los niveles superiores son, de acuerdo con el fisicalismo reduccionista, simples epifenómenos). La posibilidad de la reducción es la otra cara de una forma de esencialismo, esto es, de la creencia de que, en el fondo, todo proceso se resuelve en el nivel de las teorías básicas de la física (p. 343). Desde una perspectiva más amplia de los tipos de causalidad, si bien aún se mantiene la influencia causal de los niveles de organización inferiores sobre los superiores, en un sentido especial estos últimos también poseen poder causal sobre los primeros. Así, respecto de un organismo cualquiera, dada la continuidad material entre la física y la biología, aquel estará necesariamente determinado por las leyes que rigen los fenómenos del nivel inmediato inferior. Sin embargo, en tanto que el todo pertenece al orden de lo biológico, se rige, como tal, de acuerdo con patrones que pueden entenderse en términos de la causalidad formal y funcional (por ejemplo, mantener la estabilidad del sistema, en este caso, mantener vivo al organismo, determina las funciones específicas de sus componentes), patrones que enmarcan la acción de las propiedades del nivel organizativo inferior.

Estas consideraciones sobre los tipos de causalidad ayudan a mantener la inteligibilidad del poder causal de las propiedades supervenientes frente a las objeciones reduccionistas. No obstante, El-Hani y Pereira van más lejos en su argumentación, y sostienen que las propiedades y principios que rigen las relaciones de los niveles organizativos inferiores solo resultan explicativas respecto de las entidades que conforman los niveles superiores si es que se las lee en virtud de estas últimas. En el caso de la biología, los autores llaman a esta noción ‘significatividad biológica’ (p. 344), y es su principal argumento contra el reduccionismo. El argumento se funda en una idea que, a pesar de que pareciera ser de sentido común, el reduccionismo olvida fácilmente. La pertinencia, en el área de, por ejemplo, la biología, del recurso a propiedades de sistemas más básicos solamente se puede determinar a la luz de las propiedades de biológicas mismas. Resulta absolutamente intuitiva, como señalan El-Hani y Pereira, la diferencia entre «una masa aleatoria de moléculas y un organismo» (p. 344). Si se va a recurrir a la reducción de algún fenómeno biológico a uno molecular, la reducción solo tiene sentido si se le asigna un puesto específico dentro del ordenamiento general de las teorías biológicas.

Los mismos autores (p. 343) son conscientes de que su argumento final se inscribe dentro de una postura metafilosófica particular. La posibilidad o imposibilidad de la reducción no sería, como admiten junto con Davidson y Rorty, una disputa ontológica, sino, más bien, sobre el estatus de posibles descripciones de la realidad (pp. 343-345 y nn. 35, 36, 40 y 41).

Según Peter Menzies (2010), la mayor parte de las opiniones tradicionales en torno del estatus epistemológico de la psicología intencional se pueden dividir en dos grandes grupos. Por una parte está la tradición que el autor denomina *naturalismo científico*, que postula que la explicación del comportamiento de los agentes depende de esquemas nomológico-deductivos. De esta manera, en tanto intérpretes aplicamos, normalmente de manera no explícita, leyes que gobiernan los nexos entre los estados mentales de los agentes y las acciones que observamos. Dichos nexos son, por lo general, pensados como nexos causales, tal y como ocurre en las ciencias de la naturaleza. De hecho, desde esta perspectiva suele asumirse que los nexos causales son parte del ámbito de lo físico, por lo cual se concluye que las relaciones que explican las leyes de la psicología intencional también son relaciones causales físicas.

De otra parte, se encuentra la tradición que el autor denomina *racionalismo kantiano*. Aquí se propone que los principios de la psicología intencional cumplen un rol normativo, ya no descriptivo como en el caso anterior. Es decir, dichos principios explican el comportamiento de los agentes según lo que solemos esperar a partir de nuestra concepción general de la racionalidad. Los autores que pertenecen a este grupo tienden a distinguir entre *racionalizar* una acción (hacerla inteligible en tanto acto de un agente) y *causarla*. En su concepción, los contenidos intencionales atribuidos en las explicaciones psicológicas de la acción solo hacen lo primero, racionalizan, mas no causan, pues, al igual que quienes pertenecen al otro gran grupo, consideran que las causas siempre pertenecen al dominio de lo meramente físico.

Como señala Menzies (2010, pp. 145-149), uno de los principales argumentos desarrollados por filósofos desde mediados del siglo XX con el fin de separar de una vez por todas la explicación causal de la psicológica (y que, por ende, sirve para justificar la oposición entre naturalistas científicos y racionalistas kantianos), es el llamado «argumento de la conexión lógica». La estrategia consiste en resaltar dos rasgos interrelacionados de las explicaciones de la psicología intencional. Las generalizaciones mediante las cuales explicamos la conducta de un ente racional no parecen ser empíricas, sino únicamente conceptuales. Pensamos los actos de alguien de acuerdo con lo que idealmente concebimos como lo más racional y le atribuimos los estados mentales correspondientes al agente de las acciones. Sin embargo, por más que el agente efectivamente haya realizado los juicios que le atribuimos de acuerdo con nuestros estándares de racionalidad, aquel podría actuar de otro modo. Los casos de debilidad de la voluntad en que los agentes no parecen decidir racionalmente son abundantes entre nosotros y ello parece mostrar que las generalizaciones de la psicología intencional no son verdades empíricas. El otro aspecto de la explicación intencional que incentiva el escepticismo con relación al poder causal de las razones queda evidenciado en las expresiones que utilizamos normalmente en semejantes contextos. Usamos, por ejemplo,

palabras como el auxiliar verbal «debería» cuando pensamos la opción según la cual el agente cumpliría mejor con los estándares de racionalidad. Como dijimos arriba, aquí se hace manifiesto el carácter normativo de la explicación intencional.

Aparte de los dos grupos tradicionalmente mayoritarios en cuanto a la discusión sobre la naturaleza de la explicación intencional, existe una tercera vía, cuyo principal representante es Donald Davidson. Menzies denomina a esta opción ‘compatibilismo’, y es la que se propone defender, aunque con ciertas diferencias respecto de la versión davidsoniana, a la que denomina ‘compatibilismo de dos niveles’ (2010, pp. 145-146 y 149-153). En dos artículos clásicos —«Actions, Reasons and Causes» (2001a) y «Mental Events» (2001b)—, Davidson intenta conciliar lo que el argumento de la conexión lógica pretende demostrar que es irreconciliable: primero, que las razones atribuidas a un agente con el fin de racionalizar sus acciones son, al mismo tiempo, las causas de sus acciones; luego, que este poder causal de las razones es compatible con una concepción nomológica de la causalidad (como la que cree el propio Davidson que pertenece a las ciencias naturales) y con la irreductibilidad de la explicación psicológica a la meramente física.

Resulta esencial para la argumentación de Davidson su concepción de la causalidad como una relación extensional entre eventos. Así, un mismo nexo causal entre dos sucesos puede ser descrito en términos mentales (empleando, por ejemplo, los llamados verbos mentales, como *creer*, *pensar*, *desear*, etcétera) y satisfacer los principios de la explicación psicológico-intencional, y, además, ser descrito en términos físicos y satisfacer las explicaciones causales científicas construidas sobre la base de leyes. Pero, como también hacen muchos otros intérpretes de la propuesta davidsoniana, Menzies considera que esta salida trae consigo un costo elevado para la psicología. El problema radica en que Davidson acepta como una de sus premisas iniciales que todo nexo causal debe instanciar, al menos en una de sus descripciones posibles, leyes que lo gobiernen. Esto parece obligar a la psicología intencional a que recurra a dos tipos de explicaciones distintas: las racionalizaciones, formuladas en términos, como vimos, propiamente psicológicos, y las explicaciones causales, que únicamente cumplen la exigencia nomológica en términos físicos.

Menzies argumenta, contra lo expuesto arriba sobre Davidson, que en la realidad la psicología únicamente requiere de un mismo nivel de explicación para cubrir tanto la función de racionalizar como el esclarecimiento del nexo causal. Tanto quienes aceptan el argumento de la conexión lógica como el propio Davidson habrían cometido el mismo error, o sea, asumir que las teorías científicas contienen explicaciones causales en virtud de explicaciones nomológico-deductivas. En lugar de ello, pueden pensarse, por ejemplo, las leyes estrictas de la física, no como representaciones de regularidades empíricas, sino como los elementos que estructuran modelos ideales

que cumplen objetivos explicativos específicos dentro de las ciencias (Menzies, 2010, pp. 154-159). En otras palabras, el poder explicativo de las teorías científicas no descansaría en si las leyes que las componen son verdaderas o falsas respecto del mundo. Su poder explicativo depende de cuán eficientemente se pueden emplear como modelos que instancian sistemas reales (p. 156).

Las teorías de la psicología intencional no se distinguirían de lo que hemos dicho sobre las condiciones de la capacidad explicativa de las teorías científicas en general. Por ende, semejantes teorías son, también, modelos. En este caso, se trata de modelos contruidos sobre la base de los principios que creemos componen la racionalidad de agentes ideales. El resultado de la aplicación de estos modelos es siempre contingente. Esto último resulta importante si recordamos que, de acuerdo con Menzies, la diferencia principal entre las explicaciones psicológicas y aquellas de las otras ciencias radica en que las generalizaciones de la psicología no parecen ser tan estrictas como lo requiere la concepción nomológico-deductiva de la causalidad. De ahí que se requieran dos niveles de explicación, uno que racionalice y otro que apele a leyes estrictas que sí satisfagan lo requerido para el esclarecimiento de los nexos causales. No solo esto debería poder ser superado por la idea de la explicación científica como la aplicación de modelos ideales, también el segundo punto del argumento de la conexión lógica debería superarse. Dependiendo de para qué se le requiera, un mismo modelo psicológico podría servirnos en la explicación causal, cumpliendo un papel descriptivo, y en consideraciones normativas sobre la conducta. La diferencia en la aplicación del modelo en uno y otro caso depende de su semejanza con el sistema real. Si el modelo difiere del sistema real en la acción resultante, diremos, normativamente, que el agente *debió* actuar de otra manera.

1.5.3. Evolución de la agencia

Tal vez sea posible trazar el camino por donde pueda empezar a explicarse el paso evolutivo de la mera detección y respuesta frente a estímulos ambientales, a la agencia y la preferencia, es decir, a representar y decidir. Kim Sterelny (1999) propone una hipótesis que intenta cumplir con ello. En primer lugar, debemos considerar que existen buenas razones para cuestionar la imagen clásica de la inteligencia. Se trata de la idea según la cual el comportamiento inteligente es producto de la capacidad del agente para representarse el mundo de la manera más completa posible. El problema radica en que, dada la complejidad de cualquier entorno, resulta difícil imaginar que una mente sea capaz de mantener una representación adecuada del mismo y, a la vez, seleccione sus factores más relevantes para coordinar el comportamiento (1999, pp. 203-204). Por ello, hay quienes se inclinan en favor de pensar la inteligencia como una facultad situada, o sea, como un sistema que surge de la interacción entre

módulos especializados y ciertos elementos del entorno. Naturalmente, se trata de una visión emergentista de la inteligencia. Esta última, en tanto todo compuesto por niveles inferiores de organización de los módulos específicos, poseería propiedades únicas. No tendría caso, entonces, tratar de comprender la inteligencia así descrita analizando por separado las funciones de sus componentes (p. 205).

Sterelny sostiene que la necesidad de determinar los datos del entorno importantes para la supervivencia del organismo, son una forma de presión evolutiva que desemboca en la selección, tanto de tipos de comportamiento que se explican únicamente como producto de una relación entre ciertos dispositivos de detección y el entorno, como de la representación. Muchas veces ocurre que los organismos deben hacerle frente a lo que el autor llama 'polución epistémica' (1999, p. 208), esto es, a la emisión de datos equívocos que exigen estrategias de desambiguación. La capacidad representacional probaría ser especialmente útil para ello. Por ejemplo, medios hostiles, como aquellos que cuentan con depredadores altamente especializados, favorecerían la selección de la representación.

La representación no basta para explicar la agencia. Para elegir, también se necesita la capacidad de desear. La complejidad del entorno puede seleccionar la representación. Pero, cuando un agente elige, lo hace encarando la complejidad de sus posibles respuestas frente a los estímulos del entorno. Otra característica fundamental de los seres que poseen agencia es que estos son capaces de representar sus propios estados internos, que actúan como motivaciones para las posibles acciones. A todo ello debe agregársele que los seres en cuestión deben ser capaces de representarse el entorno como debería ser, no tan solo como es (Sterelny, 1999, p. 211).

Para explicar el paso de la mera detección de los estados internos de los organismos a su representación, Sterelny (1999, pp. 211-212) sugiere que consideremos el caso de una rata que, frente a la deshidratación, comienza a beber agua. La rehidratación de las células demora entre 10 y 15 minutos, pero el animal dejará de beber antes de ese tiempo. Aquí el estímulo que pone en marcha la acción no puede ser el mismo que la detiene. Resulta factible, entonces, que en un sistema semejante, la representación del estado interno ayude a coordinar la acción al permitir el rastreo de dos tipos de estímulo diferentes. No obstante, un ejemplo como el de la rata es insuficiente, pues ambos estímulos podrían ser internos. El autor piensa que el caso paradigmático de situación en la que se presenta un estímulo externo para cesar la acción o controlarla es el enfrentamiento con depredadores.

Sterelny también sostiene que puede haber otro tipo de factores que complejicen los sistemas de control del comportamiento. A veces el comportamiento mismo se complejiza, como en el caso de los braquiadores pesados (los grandes simios) que vimos en secciones anteriores respecto de la hipótesis de Byrne (1997) sobre la inteligencia

técnica en primates. Inclusive, una posible acción podría interferir con el desarrollo de otra, y el animal debería poder controlar su propio desempeño mediante estrategias en las cuales las acciones se ordenen jerárquicamente. El propio Byrne se refería a esto como una cierta capacidad de planificación y, por ello, no resulta extraño que Sterelny encuentre atractiva la hipótesis de aquel.

Aparentemente, entonces, frente a opciones cada vez más complejas de comportamiento, «el control se vuelve menos discreto, menos modular» (Sterelny, 1999, p. 214). Otra conclusión importante que se extrae de los argumentos de Sterelny es que la preferencia depende de la representación, pero no se puede descartar la posibilidad de que exista representación sin preferencia, pues su tesis sobre el desarrollo de la segunda toma como un hecho inicial la primera. Sin embargo, el autor no profundiza mucho más en la relación. De la misma manera, con lo dicho hasta aquí solo nos alcanza para afirmar que, aparentemente, la complejidad de los estímulos de ciertos entornos podría seleccionar la capacidad representacional y, a su vez, en algunos casos ello conduciría al desarrollo de la agencia. Pero el vínculo entre ambas capacidades cognitivas aún no ha sido esclarecido exhaustivamente.

1.6. Síntesis

Como se vio al principio de este capítulo, Darwin no solo estableció las bases del paradigma que hasta hoy rige la biología evolutiva, sino que, más específicamente, ya en su trabajo se encuentran las primeras pistas para el estudio filogenético de la cognición social humana. Siguiendo esta misma línea de investigación evolutiva, desde las últimas décadas del siglo pasado, los avances en neurociencias y en primatología les han brindado sustento empírico a las teorías de la psicología y la filosofía acerca de los orígenes de nuestras facultades cognitivas; trabajos como los de Panksepp (2007) y Premack y Woodruff (1978), y Premack (1988) son claras muestras de ello.

La llamada «hipótesis de la inteligencia maquiavélica» es la teoría más aceptada sobre las presiones selectivas que han determinado la evolución de la cognición social. Sin embargo, a pesar de que, como se explica mediante dicha teoría, la competencia entre individuos ha sido crucial en la selección de muchas de nuestras facultades cognitivas más distintivas, la tesis requiere de algunos complementos. De esta manera, y continuando con una tradición iniciada por Darwin, autores como Dunbar (2000) destacan el rol que ciertos mecanismos de cooperación han jugado en la evolución de nuestra capacidad para atribuirles varios niveles de intencionalidad a otros miembros de nuestra especie (por ello, actualmente se prefiere sustituir el término *inteligencia maquiavélica* por el más general de *inteligencia social*). Asimismo, existen teorías complementarias que toman como punto de partida algún mecanismo de interacción social; este es el caso de Byrne (1997), cuya tesis revisamos en este mismo capítulo.

Finalmente, examinamos cómo hoy en día algunos problemas clásicos de la filosofía y la psicología son abordados desde las teorías filogenéticas de la ciencia cognitiva. Un ejemplo de ello es el debate actual respecto de nuestro acceso consciente a las mentes de los demás y a nuestras propias mentes. Para unos, como Goldman (2006), el conocimiento que alcanzamos de la mente de los demás depende de nuestra capacidad para simular en nosotros mismos los posibles estados mentales ajenos. Para otros, como Carruthers (2009), el camino es el inverso: el conocimiento que tenemos de nuestros propios estados mentales es un producto colateral de nuestra actividad como lectores de otras mentes. En buena cuenta, el debate se libra en el terreno filogenético. Otro ejemplo expuesto arriba es el viejo problema de la aparente incompatibilidad entre la imagen determinista de la naturaleza, ampliamente extendida en el pensamiento científico, y la idea de sentido común de la eficacia causal de la mente (es decir, la idea de agencia). Si una teoría naturalista de la evolución de la agencia, como la propuesta por Sterelny (1999), es exitosa, quienes aquí vimos que abogan por la irreductibilidad de conceptos psicológicos de sentido común (El-Hani & Pereira, 1999; Menzies, 2010) verían sus posiciones reforzadas.

2. EL DESARROLLO DE LA COGNICIÓN SOCIAL

Hasta aquí hemos presentado la cognición social desde el punto de vista de la evolución, es decir, en términos filogenéticos. A continuación, intentaremos relacionar esta información con la que se tiene a nivel del desarrollo del individuo, es decir, desde una perspectiva ontogenética. Un puente interesante lo proporciona la psicología evolucionista del desarrollo, que trataremos a continuación. Luego detallaremos algunos hitos en el desarrollo esenciales para la cognición social y finalizaremos profundizando en la teoría de la mentalización, que forma parte de aquellas propuestas que enfatizan que es a partir de las relaciones con otras personas que el individuo desarrolla las habilidades necesarias para la cognición social.

2.1. El desarrollo desde el punto de vista evolucionista

La psicología evolucionista del desarrollo (PED) es un campo que se interesa por el desarrollo en la infancia, niñez y adolescencia desde una postura evolucionista. Asimismo, analiza las conductas parentales que evolucionaron para brindar el cuidado requerido durante dichas etapas del desarrollo. Busca describir y comprender aquellas conductas y rasgos psicológicos del cuidado parental que son adaptaciones diseñadas para incrementar la supervivencia en la adultez. Ello implica «[...] el estudio de los mecanismos genéticos y ambientales que subyacen al desarrollo universal de las competencias cognitivas y sociales, y la epigénesis evolucionada

(interacciones genes-ambiente) que adaptan esas competencias a las condiciones locales» (Hernández & Bjorklund, 2008, p. 182). Es decir, la PED se pregunta cómo los mecanismos heredados se desarrollan y se expresan en la adultez a través de la interacción de los individuos con su ambiente. Es necesario realizar un análisis evolucionista y epigenético para entender el desarrollo humano en nuestro mundo actual. Con esta finalidad, la psicología evolucionista del desarrollo ha recogido los aportes tanto de la psicología evolucionista como de la psicología del desarrollo.

La psicología evolucionista (PE) es un modelo que intenta unificar e integrar distintas áreas y corrientes de la psicología tradicional (Dunbar & Barrett, 2007). Es una psicología que recoge información de la biología evolucionista y postula que la arquitectura de nuestra mente y, por tanto, la conducta humana, son el resultado del proceso de evolución (Barkow, Cosmides & Tooby, 1992). Ahora bien, el diseño de nuestra mente, producto del proceso de selección, no está orientado a solucionar los problemas del humano actual, está dirigido a resolver los problemas adaptativos que los individuos tuvieron que afrontar como cazadores-recolectores. Cabe señalar que un problema adaptativo es aquel cuya solución puede afectar la reproducción, aunque sea de manera lejana o indirecta. Algunos problemas adaptativos son evitar el incesto, identificar plantas y animales, cazar animales, adquirir la gramática, responder a llamadas de alarma, detectar cuando los niños requieren ayuda, querer atender las demandas de los niños, seleccionar parejas, ser atractivos para las parejas, interpretar situaciones sociales, ayudar a familiares, mantener relaciones de amistad, reconocer emociones, cooperar, entre otros (Buller, 2005). La tarea de los psicólogos evolucionistas está orientada a encontrar las conexiones necesarias entre los problemas adaptativos y la estructura de los mecanismos que evolucionaron para resolverlos (Barkow, Cosmides & Tooby, 1992). Su finalidad última es descubrir, describir y explicar la naturaleza de dichos mecanismos puesto que para ellos la evolución no selecciona conductas sino los mecanismos que generan la conducta.

Por su parte, la psicología del desarrollo o psicología evolutiva se ocupa en describir y explicar los cambios en la estructura física, el comportamiento y el funcionamiento mental que surgen a lo largo de todo el ciclo vital, es decir, desde la vida intrauterina hasta la muerte (Stassen, 2008). Compartimos la propuesta de que los seres humanos se desarrollan de una manera dinámica, continua y en constante interacción con su ambiente, lo cual incluye las relaciones interpersonales e intergrupales (Cairns, Elder & Costello, 1996).

A partir de la integración de estas dos fuentes de información, la PED ha elaborado una serie de planteamientos básicos: todos los mecanismos evolucionados se desarrollan a partir de continuas interacciones entre los genes y el ambiente, en una influencia mutua; la selección natural está presente en todos los periodos del

desarrollo, no únicamente durante la adultez; ciertas características de los infantes y los niños fueron seleccionadas para ejercer funciones adaptativas en momentos específicos del desarrollo y no como preparaciones para la adultez; las adaptaciones de los niños presentan cierto grado de plasticidad o flexibilidad, lo que implica capacidad para cambiar en respuesta a distintas condiciones ecológicas y ambientales; los individuos requieren de una niñez extendida para comprender las complejidades de las comunidades sociales humanas; y, finalmente, la PED señala que tanto los mecanismos cognitivos de dominio específico como los de dominio general fueron modelados por la selección natural.

Del mismo modo, la PED plantea la existencia de tres tipos de adaptaciones. Las adaptaciones ontogenéticas son aquellas para periodos particulares del desarrollo que desaparecen cuando ya no son necesarias. Por ejemplo, el cordón umbilical y la imitación de gestos faciales en niños menores de dos años. Las adaptaciones diferidas (*deferred*) funcionan para preparar a los niños para temas específicos de la adultez. Por ejemplo, la diferencia en los estilos de juego entre niños y niñas. Por último, las adaptaciones condicionales, que aparecen en los momentos en que un individuo responde a las condiciones de su entorno regulando su conducta de manera anticipada a las condiciones de la vida futura. Por ejemplo, el modo de crianza en la infancia predice el *timing* de la pubertad y la estrategia de elección de pareja en la adultez.

Dentro de este modelo, los autores Flinn y Ward (2005) se interesaron por encontrar un vínculo entre la niñez y la mente social. Ellos consideran que la competitividad social ha sido la presión de selección principal que ha dado forma a la combinación humana de ser físicamente altricial, pero mental y lingüísticamente precocial durante la infancia, la niñez y la adolescencia. Esta combinación es posible gracias al cuidado parental y familiar extendido.

Debido a que los humanos se volvieron tan ecológicamente dominantes, ellos mismos se tornaron en su principal fuerza hostil de la naturaleza. Así, aquellos individuos más diferenciados cognitivamente, social y conductualmente tendrían mayor oportunidad para ejercer control sobre otros individuos y tener mayor acceso a los recursos de su medio, todo lo cual estaría al servicio de la supervivencia y la reproducción. Como consecuencia, para los autores se trata de una presión selectiva esencial que moldeó los principales procesos mentales de los humanos como, por ejemplo, la teoría de la mente y el lenguaje.

Ahora bien, ¿por qué ello daría lugar a una niñez prolongada? La respuesta está en que las relaciones sociales son inmensamente complejas y cambiantes, por lo cual se requiere del desarrollo de herramientas con las mismas características; ello toma tiempo. El aprendizaje, la práctica y la experiencia son indispensables para desarrollar las herramientas sociales necesarias y así obtener éxito.

De otro lado, Flinn y Ward (2005) señalan que la variedad en las relaciones interpersonales favorece la evolución de sistemas cerebrales y cognitivos más generales, capaces de integrar el procesamiento realizado por mecanismos de dominio específico. Además, hacen hincapié en la plasticidad de los sistemas neuronales, lo cual permite que el niño se adapte a su comunidad local como, por ejemplo, a su lenguaje. Cabe mencionar que el lenguaje permite un mecanismo nuevo para el aprendizaje social y la cultura; el patrón recursivo y la representación simbólica facilitan un procesamiento de información abierto, creativo y flexible.

Finalmente, como se mencionó al inicio, los padres y otros familiares juegan un rol importantísimo en el desarrollo mental del niño pues son ellos quienes, por un lado, brindan una garantía de supervivencia mientras el niño va adquiriendo las habilidades sociales necesarias y, por otro lado, constituyen puntos de referencia confiables que influyen de manera directa en el aprendizaje social del niño.

2.2. Hitos del desarrollo de la cognición social

En esta sección, señalaremos algunos logros del desarrollo que se relacionan directamente con la cognición social (Thompson, 2006). De manera general, a lo largo del primer año de vida, los infantes aprenden a distinguir el mundo social del mundo inanimado gracias a las preferencias perceptuales que los orientan hacia eventos sociales. Asimismo, poco tiempo después del nacimiento, empiezan a diferenciar entre lo interno y lo externo. Según la teoría perceptual contemporánea, ello sucede mediante la percepción, ya que esta permite distinguir entre la acción iniciada por la propia persona y aquella generada en el entorno. A partir de este punto, poco a poco, el individuo va adquiriendo conciencia de sí mismo.

Entre los dos y tres meses, ya son capaces de distinguir características faciales más detalladas. Además, comprenden que sus acciones generan una reacción en las demás personas y que esa reacción, a su vez, puede producir excitación y ánimo positivo en ellos mismos. No solo eso, sino que esperan que las personas —familiares y desconocidas— interactúen con ellos. Es importante recalcar que las diversas respuestas de los adultos tienen efectos en las reacciones emocionales y sociales de los infantes y que ello se generaliza hacia otras personas. Es decir, se van generando pautas o modelos de relaciones que marcarán una tendencia en el infante al encontrarse con otros sujetos. Hacia los tres meses, aparecen los juegos cara a cara con cuidadores conocidos, una de las formas más tempranas de interacción social focalizada y de aprendizaje social. En este intercambio socioemocional, el bebé descubre que los encuentros sociales son dinámicos y cambiantes, no siempre claros y previsible. Por otro lado, en esta etapa, desarrollan la habilidad para representar los actos de los demás como si fueran ellos los que también podrían realizarlos. Por último, otra

capacidad que surge en este momento es aquella que permite discriminar entre las distintas expresiones emocionales de la madre, lo que hace que expresen desconcierto cuando no hay congruencia entre las expresiones faciales y vocales de la madre.

Hacia los seis meses, los bebés ya saben perfectamente que la conducta de los cuidadores para calmarlos les genera una sensación de alivio. Con ese conocimiento, se producen expectativas sobre aquellas personas, tanto así que muchos infantes se empiezan a calmar con solo escuchar los pasos de un adulto acercándose.

Alrededor de los siete meses, gracias al desarrollo motor, los bebés adquieren una mayor independencia; son capaces de acercarse —o desplazarse si gatean— y explorar el entorno sin ayuda de otros. Se producen grandes avances en el sistema perceptual y, también, en el intercambio social y con el mundo. Los bebés progresan en la comunicación referencial, la comprensión de la relación entre los medios y el fin, y la interacción social a distancia. La mayor movilidad e independencia del infante permite la expresión de sus deseos y facilita una mayor conciencia de las intenciones de las demás personas. Al mismo tiempo, los cuidadores empiezan a poner límites y prohibiciones más claros para proteger al bebé.

Entre los nueve meses y el año de vida, los infantes empiezan a entender las conductas de los demás como intencionales y dirigidas hacia una meta, ya que son interpretadas en asociación a los objetos hacia los cuales están dirigidas (Thompson, 2006). Cerca al año, el infante adquiere la capacidad para la atención conjunta, la cual consiste en una coordinación de la atención entre dos individuos y un objeto o evento, es decir, una interacción triádica. Ahora bien, no se trata solo de dos personas dirigiendo su atención hacia un mismo punto, sino que ambos han de ser conscientes de que están atendiendo a algo en común; es decir, se requiere de la conciencia de experiencia compartida. Las conductas más representativas de la atención conjunta son los gestos para señalar y las respuestas frente a los gestos para señalar de las otras personas. Así, se observan gestos protodeclarativos para dirigir la atención de otra persona —como cuando el infante alterna la mirada entre la cara del adulto y un objeto mientras lo señala— y gestos protoimperativos que transmiten un pedido —por ejemplo, cuando el infante intenta agarrar algo alternando la mirada entre el adulto y el objeto que intenta agarrar— (Carpendale & Lewis, 2010).

La habilidad para la atención conjunta es considerada fundamental por muchos pues constituye la base para una mayor interacción comunicativa, para la adquisición del lenguaje y para la participación en procesos culturales. En este sentido, la atención conjunta sería un aspecto esencial para el desarrollo de la cognición social. De hecho, en algunos estudios longitudinales, se ha encontrado que la atención conjunta correlaciona con la posterior competencia social, y el desarrollo cognitivo y del lenguaje (2010).

Para algunos autores como Trevarthen (Beebe, Sorter, Rustin & Knoblauch, 2003) las personas venimos al mundo con una intersubjetividad innata que nos permite captar intuitivamente —sin requerir del lenguaje— las motivaciones y sentimientos del otro; venimos con un potencial para comunicarnos. Trevarthen propone una secuencia para el desarrollo de este potencial. En primer lugar, nacemos con una ‘intersubjetividad primaria’ que nos permite una comunicación diádica basada en la forma, la sincronía y la intensidad. Aquí se ubica la protoconversación. Luego, en una especie de paso intermedio, se ubican los juegos que, paulatinamente, irán dando paso a la ‘intersubjetividad secundaria’. Aquí se ubica la atención conjunta descrita más arriba y se empieza a formar una teoría de la mente en tanto empieza a surgir un pensamiento simbólico.

Al año de edad, ya se aprecia con bastante claridad el uso de la referencia social, que permite comprender aquellas situaciones que resultan ambiguas a partir de una lectura de las señales emocionales de otros. Esta información captada influye directamente en las propias respuestas del infante, tanto emocionales como conductuales (Thompson, 2006).

A los dieciocho meses, las inferencias que los infantes hacen sobre las intenciones de los adultos conforman la base para el aprendizaje inicial de palabras. Asimismo, en el desarrollo del lenguaje, se observa un esfuerzo por comprender las normas del entorno social. Adicionalmente adquieren la conciencia de que los deseos de las demás personas pueden ser distintos de los propios y comprenden bastante más la relación entre el deseo y la emoción: cuando los deseos se cumplen uno está contento. Alrededor de esta edad, los infantes son capaces de reconocerse a sí mismos en el espejo y se avergüenzan si su apariencia física no se adecua a lo que se espera (por ejemplo, si tienen pintada la frente).

A los dos años, los niños son capaces de representarse eventos, como las rutinas cotidianas, y de ese modo tener expectativas sobre el futuro. La organización, la estructura y el contenido de los eventos no solo se elabora a partir de lo prelingüístico sino también por la estructura verbal que se le da en el intercambio padre-hijo. La representación de eventos en general es la base para la comprensión de eventos sociales. Por otro lado, los niños a esta edad son capaces de hacer referencias verbales sobre las emociones, sus causas e, incluso, los esfuerzos para su regulación.

A los tres años, los niños saben que las emociones se vinculan a creencias y expectativas, y son conscientes de que las creencias pueden no ser consistentes con la realidad. Es decir, comprenden que, frente a una misma experiencia, las personas pueden generar diversas interpretaciones.

A los cinco años de edad, los niños son capaces de entender que los sentimientos pueden ser reactivados a partir de la conexión entre un evento actual y recuerdos

del pasado. Por ejemplo, si vemos a alguien que se parece a un familiar fallecido podría producirnos pena. Asimismo, comienzan a percibir a los otros en términos de rasgos psicológicos y pueden predecir el comportamiento futuro a partir de las características que han inferido de una persona (Thompson, 2006).

Hasta aquí, hemos presentado algunos aspectos importantes de la primera infancia que permiten apreciar el modo en que las distintas habilidades sociales van surgiendo y que permitirán —si los logros son óptimos— una comprensión y manejo adecuados del mundo social.

2.3. La mentalización. Una perspectiva relacional

En la actualidad, no hay duda de que la cognición social es el resultado de una combinación entre lo biológico y lo ambiental. Las diferencias entre teorías radican, más bien, en el peso que se le otorga a cada aspecto. Algunas teorías, aunque no niegan la influencia del entorno social, se concentran fundamentalmente en los aspectos biológicos; estudian la cognición social desde las capacidades mentales de los individuos. La teoría de la mente se ubicaría aquí, por ejemplo. Otras, como la que presentaremos, adoptan un enfoque relacional para dar cuenta de la cognición social. Desde esta aproximación, se asume que es a partir de los procesos sociales que se desarrolla la mente.

La perspectiva relacional parte de la premisa de que el bebé, al nacer, no se diferencia de su cuidador. Esto es, no puede delimitar con claridad que él es uno y el cuidador otro. A partir de las acciones en el mundo de ambos, este sistema unificado irá perdiendo fuerza para dar paso a la diferenciación. Recién en este punto el bebé podrá darse cuenta de que el cuidador y él son dos, y no uno. Así, el desarrollo de la comprensión social y de uno mismo parte de la relación entre dos individuos. Por tanto, para entender cómo surge la mente y, por ende, la cognición social, es fundamental atender el aspecto relacional (Carpendale & Lewis, 2010).

La teoría del apego se ubica dentro de esta propuesta y considera que la diferenciación se logra a partir de ciertas relaciones de apego fundamentales. Esta teoría plantea que existe un sistema cerebral que orienta al bebé a buscar cercanía y comunicación con sus cuidadores. En principio, ello cumple un rol importante para la supervivencia, pero este sistema es más complejo aún. A nivel mental, el cerebro poco desarrollado del bebé, a partir del encuentro interpersonal, utiliza las funciones ya maduras del cerebro de sus cuidadores para organizar sus propios procesos mentales. A partir de estas relaciones emocionales se desarrolla una serie de funciones mentales importantísimas como la memoria, la narración, las emociones, las representaciones y los estados mentales (Siegel, 1999). Así, el apego podría funcionar como una base sobre la cual se desarrollaría la mente.

Para que esto se logre de manera óptima, aquello que se denomina *apego seguro*, es esencial que los cuidadores sean sensibles a las señales del bebé, pues de ese modo podrán amplificar las respuestas emocionales positivas del bebé y modular las negativas. El cuidador ayuda al bebé a disminuir la intensidad de emociones desagradables como el miedo, la tristeza o la ansiedad, con lo cual se tranquiliza, y, a la vez, le ofrece consuelo frente a ese malestar, es decir, seguridad. A partir de las repeticiones de este patrón de intercambio se van a ir generando modelos que el bebé luego interiorizará y hará propios. Cuando esto suceda, será capaz de proporcionarse la tranquilidad y seguridad por sí mismo, sin necesidad del cuidador. Varios estudios sobre este tema han encontrado que los modelos generados a partir del sistema de apego se asocian con los procesos particulares de la regulación emocional, las relaciones sociales, la disponibilidad de la memoria autobiográfica y el desarrollo de la autorreflexión y la narración (1999). Por otro lado, también se ha encontrado en investigaciones una asociación entre el apego seguro y el desarrollo social de los niños, la cual incluye el hallazgo de que los niños con apego seguro tienen más probabilidades de pasar la prueba de la falsa creencia a alrededor de los cuatro años que los niños con apego inseguro (Carpendale & Lewis, 2010).

La mentalización se encuentra muy próxima a la teoría de apego, como se verá a continuación. La mentalización o función reflexiva del *self* es la habilidad que permite a las personas percibir e interpretar su propia conducta y la de los otros en términos de estados mentales intencionales, e imaginar lo que los demás sienten y piensan (Busch, 2008). Dentro de los estados mentales se encuentran los deseos, los sentimientos, las creencias, las fantasías, las expectativas, entre otros. Esta habilidad incluye, en realidad, una serie de capacidades tanto explícitas como implícitas, es decir, que se ubican en niveles conscientes y no conscientes o automáticos. Dentro del primer grupo se ubican, por ejemplo, la interpretación verbal de un evento o la elaboración de una explicación, mientras que en el segundo se encuentran la intuición o la sintonía afectiva entre madre e hijo (Fonagy, 2008).

Hace poco, influidos por la psicología evolucionista, Fonagy y Target (1997) han redefinido la función reflexiva del *self* por la función interpretativa interpersonal del *self* (IIF, por sus siglas en inglés), en un intento por resaltar el componente intersubjetivo de esta capacidad y su impacto tanto en la evolución como en el desarrollo de la vida mental y social de los individuos (Bouchard & Lecours, 2008).

La IIF permite procesar las experiencias nuevas e interpretarlas en términos psicológicos a partir de la atribución de intencionalidad. Esta función presenta cuatro elementos que se desarrollan de manera sucesiva en el contexto del vínculo de apego: a) el reconocimiento y regulación de afectos, b) la regulación de los estados de activación internos, c) la regulación cognitiva —que implica el control de la atención

y la conceptualización de los estados mentales— y d) las capacidades específicas de la mentalización. Cabe resaltar que la habilidad para mentalizar se concretiza entre los tres y los seis años.

Fonagy postula, al igual que el biólogo evolucionista Alexander (1989, citado por Fonagy, 2008), que las habilidades mentales más sofisticadas, como la inteligencia social, fueron impulsadas por la competencia entre los seres humanos. Luego de alcanzar cierto grado de dominio sobre el ambiente natural, las personas nos convertimos en nuestra principal amenaza y, debido a ello, tuvimos que aprender a cooperar entre nosotros. Competir con otros exige entenderlos y predecir sus intenciones. De este modo, las capacidades para construir y manipular situaciones en la mente sobre los pensamientos y sentimientos se tornaron en una ventaja reproductiva.

Para Fonagy (2008), los vínculos de apego en los seres humanos generan un ambiente seguro para el adecuado desarrollo del cerebro social. Las capacidades de mentalización, al igual que las otras habilidades cognitivas necesarias para su crecimiento, evolucionan a partir de los intercambios sociales con los cuidadores principales. Como mencionamos anteriormente, el crecimiento y la complejización de las habilidades sociales evolucionaron junto con otros factores del desarrollo, como el aumento de la indefensión en los infantes y la prolongación de la infancia, y la aparición de formas de paternaje más sofisticadas.

A continuación, se presentarán los hitos ontogenéticos, agrupados en cuatro etapas secuenciales que delinearán el desarrollo de la capacidad para la mentalización (Fonagy, 2008).

La primera etapa, en la cual se aprecia una *agencia teleológica*, está conformada por la segunda mitad del año de vida. Durante este periodo el infante logra captar e identificar la relación causa-efecto. Espera desenlaces determinados para acciones conocidas y reconoce que estas son medios para lograr fines específicos. La agencia es entendida en términos de la elección de los medios más eficaces para lograr el objetivo esperado. Debido a ello, el infante tiene la expectativa de que los otros agentes llevarán a cabo las acciones que sean las más eficaces para alcanzar una meta particular. Todavía no es capaz de atribuir estados mentales como motivadores de la conducta ni de considerar otras perspectivas distintas de la suya. Al acercarse al primer año de vida, las habilidades comunicativas del infante se tornan más elaboradas; empieza a hacer uso de la referencia social, es decir, busca el rostro del adulto para obtener información sobre si debe acercarse o alejarse de determinados objetos. Asimismo, comienza a señalar los objetos alternando la mirada entre el objeto y el rostro del adulto. Finalmente, un logro muy importante de esta etapa es el de la atención compartida.

La *agencia intencional* es el modo de comprensión que corresponde a la segunda etapa, que se extiende desde el primer año hasta, aproximadamente, los tres años de edad. El infante empieza a atribuir estados mentales como deseos y sentimientos; es capaz de razonar de un modo no egocéntrico sobre los sentimientos y los deseos de las demás personas, y logra diferenciar entre las metas intencionales de una acción y sus consecuencias accidentales. Aún no puede representar estados mentales de forma independiente a la realidad física. La relación realidad y apariencia, mundo interno y mundo externo, todavía no se encuentra delimitada. Fonagy (2008) y Fonagy y Target (1997) sostienen que, en este periodo, el infante vive de manera simultánea dos formas de experimentar la relación entre la realidad externa y la realidad interna: el modo aparente (*pretend mode*) y el modo de equivalencia psíquica. En este último, la mente y el mundo es lo mismo, la relación entre ambos es especular, los estados mentales son reflejo del mundo y viceversa. No hay perspectivas. El modo aparente surge en el juego; sorprendentemente, en este caso el niño opera representacionalmente, puede creer que él es un superhéroe que va a salvar a una persona con su espada, la cual es, en realidad, un bastón. No obstante, se mantiene desconectado de la realidad; requiere separar la realidad interna de la realidad externa. La comprensión mentalística se logra cuando ambos modos se integran dando como resultado que los estados mentales sean representaciones que se relacionan con el mundo pero no de un modo equivalente. Así, en la siguiente etapa, el entendimiento se sofisticada e incluye la atribución de creencias y la consideración de perspectivas diferentes acerca del mundo.

En la etapa de la *agencia representacional*, entre los tres y los cinco años, el niño incluye a las creencias como estados mentales que motivan la acción. Asimismo, es capaz de diferenciar entre lo que percibe en un individuo y lo que podría estar sucediendo en la mente de ese mismo individuo. Logra atribuir falsas creencias a sí mismo y a los otros, y adquiere las capacidades necesarias para intercambios sociales nuevos como bromear, engañar y hacer trampa. Además, logra captar las emociones de los otros a través de la empatía. De otro lado, el niño empieza a preferir el juego con sus pares en lugar del juego con adultos. Es así como se va cerrando el periodo en el que la adquisición de las capacidades de mentalización se daba a partir de la agencia de una mente adulta y surge un nuevo periodo más largo en el que se promueve la comprensión de las mentes de otros con quienes se comparte similitudes en el sentido del humor y en los intereses.

Por último, la etapa de la *agencia autobiográfica*, que se presenta alrededor de los seis años de edad. En este momento hay grandes avances en el desarrollo de las habilidades para recordar y narrar experiencias y actividades intencionales de un modo organizado y coherente. El niño logra hacer un uso adecuado de la causalidad

y la temporalidad. Del mismo modo, la teoría de la mente se vuelve más sofisticada con la aparición de nuevas capacidades como las representaciones de segundo orden —comprender creencias erróneas sobre otras creencias—, la comunicación de sentimientos mezclados o encontrados —estar en conflicto—, modos de engaño más sutiles como las mentiras blancas, y la conciencia de que las expectativas de uno mismo pueden influir o sesgar la interpretación de un evento ambiguo. A partir de la elaboración de narrativas autobiográficas, los niños logran la comprensión de sí mismos y de los demás.

Existen dos funciones parentales esenciales que dan inicio a la secuencia de etapas antes mencionada: el reflejo (*mirroring*) y el *marking* (no hay traducción). El *marking* es la tendencia de los adultos a enfatizar con la voz y los gestos la emoción que el bebé transmite. Esta conducta le proporciona al infante la primera representación de su propia emoción, la cual usará como base para la identificación de sus emociones. Por su parte, el reflejo o *mirroring* se refiere a la transmisión del estado emocional del bebé a partir de la exageración de los gestos y las tonalidades de voz (Gergely, 2007).

Fonagy sostiene que el apego es fundamental para la adquisición de la teoría de la mente, y critica las posturas nativistas que plantean que el entorno social del infante dispara pero no determina su desarrollo. Si bien existe evidencia de una programación universal para el desarrollo de la teoría de la mente en un momento particular del desarrollo, la mayor parte de las investigaciones han encontrado que existe una variación transcultural con respecto al tiempo de su adquisición. Asimismo, factores como el modo de interacción entre los padres, la manera como estos hablan de las emociones a sus hijos, el espacio que le otorgan a las emociones en sus explicaciones sobre las acciones, y las creencias de los padres sobre la parentalidad, influyen en la teoría de la mente del niño (Fonagy, 2008).

Según Fonagy y Target (1997), las influencias sociales más importantes en el desarrollo de la teoría de la mente son el juego, el tipo de comunicación entre padres e hijos y el intercambio con los pares. El juego de simulación (*pretend*) favorece la comprensión de los estados mentales. Las representaciones que se originan durante dicho juego son, a la vez, diferentes de la realidad externa y compartida con otros individuos que son parte de la realidad: los compañeros de juego. Surge una conexión entre realidad y fantasía que empieza a perfilarse. El niño juega con la realidad y confía en que el otro con quien juega, que es real, puede compartir esa vivencia con él. De otro lado, el lenguaje tiene un rol central en la adquisición de la habilidad para mentalizar. Hablar de las intenciones y los sentimientos detrás de las acciones son prácticas cotidianas que se relacionan positivamente con la capacidad reflexiva del niño. Por último, el intercambio con hermanos y otros pares facilita la familiarización con perspectivas distintas a la propia, y permite la interacción con otras mentes.

No obstante, el señalamiento más importante para resaltar el papel del ambiente en la adquisición de la teoría de la mente se encuentra en los estudios que vinculan apego y mentalización. El punto principal es que la seguridad del apego entre el niño y sus cuidadores produce comprensión psicológica en la vida futura del niño.

Fonagy y Target (1997) plantean que los cuidadores ‘seguros’ muestran una tendencia a ‘contener’ los afectos de sus hijos al transformarlos de modo tal que sean más manejables y comprensibles para ellos. En contraste, el vínculo con un cuidador inseguro hace que el niño se identifique con un comportamiento defensivo, proveniente del cuidador, ya que si no lo hace perdería la cercanía con este. Cuando los padres son despreocupados tienden a desatender el afecto del niño, dado que les resulta demasiado movilizante y los hace recordar vivencias dolorosas. En el caso de los padres preocupados, la tendencia es a exagerar o tergiversar la experiencia emocional del niño. En cualquiera de los casos, no hay una sincronía entre la manera de decepcionar el afecto del padre y el modo de sentir del niño, lo cual da lugar a una sensación de falta de coherencia y de sentido del afecto al interior del sí mismo. Ello podría ocasionar en el niño una dificultad para representarse la experiencia afectiva (en el caso de cuidadores despreocupados) o la interiorización de construcciones extrañas sobre sí mismo (en el caso de cuidadores preocupados).

La adquisición de la teoría de la mente es parte de un proceso intersubjetivo entre el niño y sus cuidadores. La habilidad del niño para desarrollar una teoría de la mente depende de la observación de sus propios estados mentales en el mundo mental del adulto cuidador. Para que esto ocurra, el cuidador debe ofrecerse como alguien confiable que promueva la exploración mental del niño, y ser lo suficientemente sensible para captar y comunicar el sentido de los estados mentales en el niño durante las interacciones de rutina. Esta sensibilidad en los padres o cuidadores es un elemento esencial para el surgimiento de un vínculo de apego seguro en los hijos.

Cuando un cuidador es confiable, el niño siente la confianza suficiente para poder interpretar su conducta, ya que no encuentra nada de qué asustarse. En cambio, el niño con apego inseguro tiende a empobrecer sus capacidades reflexivas sobre los estados mentales de las otras personas, se aleja del intercambio intersubjetivo y, con ello, de la posibilidad de encontrar en el otro el medio para identificar sus propios estados mentales. El movimiento es defensivo, es decir, que el niño lo que está intentado hacer es protegerse. Así, prefiere refugiarse en la fantasía y la desconexión (características similares a las del cuidador) para poder defenderse y, a la vez, conservar el vínculo con sus cuidadores. El niño desorganizado o ambivalente desarrolla de manera prematura la capacidad para interpretar los estados mentales detrás de la conducta de sus padres en ciertas situaciones (particularmente las ansiógenas o violentas), pero no orienta la función reflexiva hacia sí mismo. Hay un desequilibrio

en el uso de las facultades mentales reflexivas ya que estas se concentran en los cuidadores, quienes, al ser percibidos como impredecibles, ambivalentes y hasta violentos, generan en el niño una hipervigilancia en relación a sus estados mentales. Adquiere algo como una teoría-teoría para comprender la conducta de los otros, pero no una capacidad para mentalizar propiamente dicha dado que el papel que esta tiene en la organización del *self* no se logra. El niño, en cambio, se protege ante la posibilidad de representarse a sí mismo en base a la imagen distorsionada (y a menudo maligna) que los padres o cuidadores le transmiten en las interacciones cotidianas.

Los niños que han sido expuestos a experiencias traumáticas suelen tener dificultades para delimitar adecuadamente entre realidad y fantasía. La sensación de agencia sobre la propia conducta se ve limitada, lo que afecta de manera sustancial la capacidad para diferenciar la propia experiencia de la ajena. Como resultado, algunas capacidades como la empatía se ven afectadas ya que el individuo no comprende la dimensión mental que subyace a la conducta de los otros, no produce una teoría del dolor ni la resonancia afectiva que esta tiene en su propia vivencia. Las personas con estructuras limítrofes de personalidad son un ejemplo para dar cuenta de una adquisición disfuncional de esta capacidad. La falta de sintonía de los padres ocasiona que sus estados mentales sean malinterpretados y reflejados por los mismos, y la persona los experimente como alienados. Intentará liberarse de ellos externalizándolos en sus vínculos con los demás. Las funciones prerreflexivas seguirán influyendo, de manera que la vivencia de la realidad quedará sujeta a las distorsiones de la fantasía interna y, de igual modo, la realidad interna será experimentada con la intensidad de la realidad externa.

La mentalización, entonces, abarca una serie de habilidades que permiten la autorreflexión y la comprensión de los demás en base a la atribución de estados mentales como los deseos, las creencias, los sentimientos, entre otros. Sin esta posibilidad de reflexionar sobre nosotros mismos y sobre los demás, como agentes intencionales, sería casi imposible entender el mundo social a nuestro alrededor.

2.4. Síntesis

La psicología evolucionista del desarrollo (PED) constituye un puente interesante entre los aspectos filogenéticos y los ontogenéticos de la cognición social. Busca analizar el desarrollo en la infancia, la niñez y la adolescencia desde una posición evolucionista por un lado y, por el otro, estudiar las conductas parentales que evolucionaron para proporcionar los cuidados necesarios durante esos periodos de desarrollo. Le interesa comprender cómo los mecanismos heredados —moldeados por la selección natural— se desarrollan y expresan a lo largo del ciclo vital de los individuos en interacción con su ambiente (Hernández & Bjorklund, 2008).

Dentro de este modelo, Flinn y Ward (2005) consideran que la competitividad social ha sido la presión de selección principal —desde que los seres humanos dominaran su medio ambiente—, que moldeó los procesos mentales centrales de los humanos, como la teoría de la mente y el lenguaje, ya que los más diferenciados cognitivamente, social y conductualmente tendrían mayor posibilidad de ejercer control sobre los demás individuos y, así, tener mayor acceso a los recursos.

Casi desde el inicio de la vida se observa una serie de logros íntimamente relacionados con la cognición social. Ya entre los dos y tres meses, los bebés entienden que sus actos producen una reacción en los demás y esperan que los adultos interactúen con ellos; van generando modelos de relación y son capaces de diferenciar entre las distintas expresiones emocionales de la madre. Asimismo, surgen los juegos cara a cara. Entre los nueve y doce meses, los infantes empiezan a entender la conducta de los otros como intencional y dirigida hacia una meta; va concretándose la capacidad para la atención conjunta, la cual es esencial para el desarrollo de la cognición social, y se manifiesta fundamentalmente con gestos protodeclarativos y gestos protoimperativos. Al año de edad, el uso de la referencia social es bastante evidente (Thompson, 2006; Carpendale & Lewis, 2010). A partir del año y medio comprenden que los deseos de los demás pueden ser diferentes de los de uno y que hay una relación entre el deseo y la emoción. A los dos años, los niños son capaces de representarse eventos y, así, tener expectativas sobre el futuro. De igual modo hacen referencias verbales sobre las emociones, sus causas y, también, los esfuerzos para su regulación. A partir de los tres años saben que las emociones se asocian a creencias y expectativas, y que las creencias pueden no ser consistentes con la realidad. Ya a los cinco años, los niños empiezan a percibir a los demás en términos de rasgos psicológicos y logran predecir el comportamiento futuro a partir de las características atribuidas.

Existen algunas teorías que se enfocan, sobre todo, en los procesos interpersonales que posibilitan el desarrollo de la cognición social. La teoría de la mentalización se ubica en esta categoría. La mentalización requiere del desarrollo de una serie de habilidades que permite al niño interpretar su conducta y la de los otros en términos de estados mentales intencionales —como los sentimientos, los deseos, las creencias, etcétera. (Fonagy, 2008 y Busch, 2008)—. Para que ello se logre es necesario que el niño tenga un apego seguro, que le permitirá la exploración mental con un sentimiento de protección. Fundamentalmente, el cuidador deberá captar y comunicar el sentido de los estados mentales del niño durante los encuentros cotidianos puesto que el infante necesita hacer uso de las funciones ya maduras del cerebro adulto para organizar sus propios estados mentales. Es a partir de estos intercambios con otros que el individuo logra desarrollar las funciones mentales esenciales para un adecuado desarrollo y desenvolvimiento social (Siegel, 1999).

EVOLUCIÓN Y DESARROLLO DEL LENGUAJE

Mente y Lenguaje

Grupo Interdisciplinario de Investigación

Andrés Abugattás, Ricardo Braun, Paola Cépeda, César Escajadillo,
María de los Ángeles Fernández Flecha, Marcos Herrera Burstein,
Carla Mantilla, Luis Manuel Olguín, Jorge Pérez Silva,
Pablo Quintanilla y Carolina Romero

En el primer capítulo de esta revisión teórica nos ocupamos de la cognición social, es decir, de las capacidades cognitivas que nos permiten relacionarnos socialmente con otros individuos. Elaboramos en él un breve recuento de las principales teorías que explican su evolución en la especie, así como su adquisición y desarrollo en el niño. En este segundo capítulo pretendemos ocuparnos específicamente de nuestra capacidad lingüística. De manera similar, partiremos de una definición inicial de lenguaje que nos permita presentar las propuestas y debates más interesantes sobre su evolución y luego sobre el proceso de su desarrollo en el desarrollo.

El capítulo a continuación está dividido en tres subtítulos. El primero ofrece una definición de lenguaje a partir de describir una lengua, tanto como fenómeno sociohistórico, cuanto como conocimiento adquirido. El segundo presenta un recuento de las propuestas y debates más resaltantes sobre la aparición y evolución de la facultad del lenguaje en la especie. Finalmente, en el tercero, presentamos los hitos más importantes en el desarrollo del lenguaje en el niño y las principales teorías que explican la adquisición de una lengua materna. Las referencias bibliográficas de ambos capítulos aparecen en conjunto como último apartado.

1. ¿QUÉ ES EL LENGUAJE?

Este apartado está dividido en dos partes. La primera ofrece una definición de la noción de lengua como fenómeno social, cultural e histórico. Destacamos, primero, la asombrosa diversidad lingüística en el mundo para presentar, después, las características del cambio natural por el que pasan las lenguas a lo largo del tiempo.

Nos interesa particularmente distinguir con claridad la noción de *cambio lingüístico* de la de *evolución del lenguaje*, la que será elaborada más adelante. En la segunda parte definimos a una lengua como el conocimiento que le permite a un individuo producir y comprender enunciados en ella. Hacemos notar que este conocimiento está organizado en componentes específicos encargados del procesamiento de diferentes tipos de información. El subtítulo concluye con una pequeña síntesis de lo revisado.

1.1. Las lenguas y el cambio lingüístico

Un rasgo exclusivo y fundamental de los seres humanos es el de poder hablar. Muchas veces se denomina a esta capacidad como *la facultad del lenguaje*. Pero hablar es siempre hablar una lengua. Por ello es útil iniciar nuestra introducción a esta parte presentando la noción de lengua y explicando, además, una propiedad fundamental de las lenguas del mundo: el hecho de que estas varían geográfica, social e históricamente.

1.1.1. ¿Qué son las lenguas?

Los estudiosos calculan que en la actualidad se hablan unas 6000 lenguas o idiomas en el mundo (Comrie, 2009; Brown & Ogilvie, 2009). Aunque las nociones de *lengua* o *idioma* son bastante intuitivas, delimitar una lengua o definir un idioma no son asuntos sencillos, como podría parecer a primera vista. Si bien resulta claro que algunas personas hablan, por ejemplo, chino; otras, castellano; y otras, quechua; de esto no se sigue que existan tres entidades nítidamente diferenciables: el chino, el castellano y el quechua. Lo que existe en concreto son individuos, y suele decirse que hablan una misma lengua o idioma cuando son capaces de interactuar pronunciando e interpretando enunciados lingüísticos de manera fluida, generalmente con fines comunicativos. Sobre la base de este comportamiento de las personas —miembros, por lo general, de una misma comunidad—, suele pensarse que cada una de ellas, como producto de su socialización, ha desarrollado un *saber hacer* o un *saber representacional* compartido con los otros miembros de su comunidad. Es este saber abstracto el que se identifica con la lengua y, por ello, suele decirse que las personas «aprenden una lengua».

Así pues, podría caracterizarse una lengua como el saber compartido por los miembros de una comunidad que les permite pronunciar e interpretar enunciados. Si bien esta definición es fundamentalmente adecuada, hay que precisarla en dos aspectos. Por una parte, no todos los hablantes de una lengua usan su idioma exactamente de la misma manera y, por tanto, su saber lingüístico no es idéntico sino tan solo parcialmente compartido. Los 400 000 000 de castellanohablantes no hablan todos igual, y lo mismo ocurre con los 1 500 000 000 de hablantes de chino y con

los 8 000 000 que hablan quechua. Todas las lenguas están formadas por variedades o dialectos distintos según los diferentes lugares en que se usan —no SonIguales el castellano de México y el de Buenos Aires, ni el chino de Pekín y el de Cantón, ni el quechua de Huaraz y el de La Paz—. Incluso en un mismo lugar, no usan igual su idioma miembros de grupos sociales diferentes, como los hombres y las mujeres, los ancianos y los jóvenes, etcétera. De esta manera, habría que precisar que la lengua es un saber compartido en gran medida por sus usuarios y no un saber idéntico en cada uno de ellos.

Por otra parte, debemos precisar la definición de lengua o idioma como saber compartido tomando en cuenta un segundo aspecto. No es solo el tipo de saber el que permite definir una lengua y distinguirla de otra, sino también un criterio de base social e histórica. Por ejemplo, lo que distingue el chino del castellano no es solamente que los saberes de sus hablantes son muy diferentes y, por tanto, no pueden interactuar lingüísticamente de manera exitosa, dado que esto último también ocurre con hablantes de una misma lengua. En efecto, los hablantes de una variedad de chino no necesariamente entienden a los hablantes de otra variedad, y lo mismo ocurre con los hablantes de quechua y hasta con los de castellano. Y, de otro lado, hablantes de idiomas diferentes pueden entenderse entre sí, como ocurre, por ejemplo, con quienes hablan castellano, gallego y portugués, o sueco, danés y noruego. La delimitación de una lengua y su diferenciación con respecto a otra depende, principalmente, de la tradición sociohistórica de sus hablantes: por lo general, las personas nacemos en una comunidad que identifica su manera de hablar como propia distinguiéndola de la manera de hablar de otros grupos y esto es todavía más claro en el caso de las naciones-estado contemporáneas, que suelen declarar una o más lenguas como «idioma oficial».

Así, a nuestra definición de lengua o idioma como saber en gran medida compartido por los miembros de una comunidad —una definición basada en un criterio de índole cognitiva—, habría que añadirle la precisión de que su delimitación no depende solo de las características de este saber sino también, como veremos a continuación, de una tradición sociohistórica.

1.1.2. El cambio histórico y la diversificación de las lenguas

Hemos señalado que se calcula que actualmente se hablan unas 6000 lenguas en el mundo y que cada una de estas, a su vez, está conformada por cierto número de variedades. La variación en el lenguaje humano es la norma antes que la excepción. La evidencia conocida sobre la manera en que se hablaban algunas lenguas en el pasado muestra que estas están en constante cambio y lo mismo muestra la evidencia sociolingüística del presente: los miembros de grupos sociales tienden a adoptar

rasgos lingüísticos para identificarse entre sí y para distinguirse de otros; el habla de los jóvenes en comparación con la de sus padres y abuelos es, probablemente, el mejor ejemplo de cómo una lengua puede cambiar poco a poco de generación en generación.

El cambio lingüístico del castellano a través de la historia se encuentra bastante bien documentado. La diversidad lingüística de tipo geográfico y social que muestra en la actualidad el castellano americano es el resultado de 500 años de transformaciones que ha sufrido esta lengua desde que fue traída de la península Ibérica por los conquistadores españoles. Los diferentes rasgos de tipo léxico, fonológico y gramatical que distinguen a argentinos, mexicanos, bolivianos, cubanos, nicaragüenses, colombianos, etcétera, así como los rasgos que distinguen entre sí a diferentes grupos sociales de peruanos, a diferentes grupos de dominicanos, de chilenos, etcétera, fueron adquiriendo la forma que tienen en cada una de sus comunidades a lo largo de estos años. Por cierto que los primeros españoles llegados a América no hablaban todos igual —en la península, en aquella época (como hasta ahora), también había variedades lingüísticas—, pero la fisonomía actual de las variedades lingüísticas del castellano americano actual se debe, principalmente, al cambio paulatino del que ha sido objeto cada una, en su propia dirección y a su propio ritmo, por causa de sus hablantes.

Podemos darnos una idea indirecta de cómo ha cambiado el castellano desde principios del siglo XIII hasta hoy examinando un texto de aquella época:

De los sos ojos tan fuertementre
llorando tornava la cabeça i
estávalos catando

De sus ojos fuertemente llorando
volteaba la cabeza y los estaba
mirando

Si bien se trata de un poema escrito y no de cómo hablaba la gente normalmente, los estudiosos de la historia del castellano proponen hipótesis sobre la pronunciación que debieron representar estas letras. Así, por ejemplo, la <j> correspondía aproximadamente a como los uruguayos pronuncian la <y>; la <ll> correspondía al sonido lateral que todavía se usa en la zona andina para las palabras que llevan esta grafía; por último la <ç> era más o menos equivalente a <ts>, mientras que en la actualidad este sonido ha sido reemplazado por la zeta española en partes de la península y por el mismo sonido que representa la letra <s> en América. También podemos observar que sintácticamente el castellano contemporáneo no utiliza la construcción *los sus ojos* sino *los ojos* o *sus ojos*. Finalmente, podemos observar que los signos léxicos *tornar* y *catar*, si bien existen en castellano contemporáneo, tienen en la actualidad otros significados, y que, más bien, el significado que tenían en el siglo XIII se expresa en castellano contemporáneo con los verbos *voltear* y *mirar*, respectivamente.

El castellano del siglo XIII era el resultado de la manera en que había terminado hablándose en la región de Castilla la lengua latina que habían llevado los romanos, al menos quince siglos atrás, en su conquista de la península Ibérica. En tanto tiempo el latín se transformó, en boca de sus usuarios, fonológica, léxica y gramaticalmente. El siguiente texto del siglo I d. C., que recoge parte de la fábula de *La zorra y las uvas* de Fedro, nos da una idea del latín de ese momento y nos permite hacernos una idea sobre la paulatina transformación de esta lengua al castellano contemporáneo a lo largo de veinte siglos:

*Fame coacta vulpes alta in vinea
uvam appetebat, summis saliens
viribus.*

Acuciada por el hambre, una zorra
intentaba alcanzar unas uvas que
estaban en viña alta saltando con
todas sus fuerzas.

Si bien es casi imposible entender a cabalidad el texto, sí se pueden reconocer en él algunas características que sobreviven en nuestra lengua. La frase *alta in vinea*, por ejemplo, presenta un par de diferencias sintácticas con respecto a su versión actual, *en una viña alta*: el orden de las palabras y la presencia del artículo. Asimismo, muestra pequeñas diferencias fonológicas: *in*, en contraste con *en* y *vinea* en contraste con *viña*. No cabe duda de que estas palabras son las mismas que se usaban hace 2000 años, solo que su pronunciación ha variado ligeramente (a diferencia de la palabra *alta*, ¡que se sigue pronunciando igual!). Por su parte, las palabras *vulpes* y *viribus* han sido reemplazadas por otros elementos léxicos contemporáneos: *zorra* y *fuerzas*, respectivamente.

Este panorama nos ayuda a entender cómo se transforma y diversifica una lengua con el paso del tiempo y por acción de sus hablantes. El latín hablado en diferentes partes fue transformándose de distintas maneras, dando lugar a las diferentes lenguas romances contemporáneas: italiano, portugués, francés, catalán, etcétera, cada una con sus propias variedades lingüísticas. Así, pues, en un sentido muy fundamental, todas las variedades de estos idiomas que se hablan en la actualidad son una misma lengua solo que transformada históricamente por sus hablantes.

Ahora bien, el latín del siglo I d. C. también tenía variedades geográficas y sociales, y también era el resultado de un proceso histórico de transformación practicado por generaciones de hablantes. En efecto, según los estudiosos, esta lengua, así como el griego, el persa, el sanscrito y muchas otras, tienen un origen común en una lengua ancestral conocida como protoindoeuropeo. Esta habría sido la lengua hablada hace unos 5000 o 7000 años por un grupo de pobladores de Europa oriental que empezaron a migrar en distintas direcciones en busca de nuevas tierras.

Los pueblos indoeuropeos habrían llevado con ellos, además de la agricultura y otras técnicas, su lengua, la cual, con el paso del tiempo y el cambio natural producto del uso de sus hablantes, se habría transformado en la diversidad de lenguas que hoy se conoce como la *familia indoeuropea*. El español, por ejemplo, pertenece, con el portugués, el francés, el rumano y otras lenguas de la Europa occidental, al subgrupo denominado *romance*. Con otros subgrupos, ahora extintos, el romance formaba la subfamilia de lenguas itálicas, propias de la península apenina.

Algunos estudiosos del cambio y la diversificación de las lenguas (Greenberg, 1987 y Ruhlen, 1994) han propuesto que las 6000 lenguas del mundo pueden agruparse en menos de dos decenas de familias, lo que sugeriría la existencia anterior de un número similar de protolenguas que habrían dado lugar a todos los idiomas que conocemos hoy. Aunque no todos los expertos están de acuerdo con esta propuesta específica, todos creen que las lenguas contemporáneas derivan de un proceso continuo de transformaciones y diversificaciones de estados anteriores. El caso de la familia indoeuropea es probablemente el mejor estudiado por la calidad de la evidencia con la que se cuenta. En otros casos, las hipótesis son más difíciles de verificar, lo que las deja en el terreno de la especulación.

En este mismo terreno se encuentran los intentos de respuesta a la pregunta que surge lógicamente de toda esta discusión: ¿existió una primera lengua que dio origen a todas las demás por cambio y diversificación? Si, como sugieren los estudios genéticos y paleoantropológicos, asumimos que todos los *homo sapiens* que poblamos la Tierra en la actualidad tenemos un mismo origen africano y que nuestros ancestros empezaron a poblar el mundo hace aproximadamente 100 000 años, entonces, resulta muy posible que el lenguaje humano haya existido ya desde entonces. Así, todas las lenguas del mundo no serían sino variedades de esa lengua original que habría ido cambiando a lo largo de miles de generaciones, conforme nuestros antepasados iban separándose en grupos buscando tierras nuevas. Quizá una evidencia notable a favor de esta propuesta es que el diseño fundamental del lenguaje humano se encuentra en todas las lenguas: todas son sistemas cognitivos que incluyen un componente léxico, uno fonológico y uno gramatical¹.

¹ No obstante, esta afirmación debe matizarse debido a una importante excepción: las lenguas de signos manuales y gestuales de los sordos no contienen propiamente un componente fonológico debido a la naturaleza visual de sus signos. Estos, de todos modos, contienen elementos de expresión organizados por lo que en las lenguas orales sería el componente fonológico. Por otro lado, como veremos en la siguiente sección, al describir el sistema cognitivo en el que consiste una lengua también se considera, al lado de estos componentes, a un componente semántico y a un componente pragmático.

1.2. La estructura del saber lingüístico

Más arriba hemos caracterizado una lengua como el saber compartido por los miembros de una comunidad que les permite pronunciar e interpretar enunciados. Dicho saber presenta para todas las lenguas aproximadamente la misma estructura general, que consta de distintos componentes interrelacionados: el componente fonológico, el léxico, el gramatical, el semántico y el pragmático². El *componente fonológico* de una lengua está constituido por las unidades que conforman la expresión de los signos del léxico, así como por las reglas que gobiernan su pronunciación y audición. Las unidades del *componente léxico* o vocabulario son expresiones sonoras (en la mayoría de lenguas) o gestuales (en las lenguas de los sordos) que sus usuarios asocian, por convención, con algún tipo de contenido mental, es decir, son expresiones interpretables o con significado —o, simplemente, ‘signos o símbolos lingüísticos’—. Los signos del léxico pueden combinarse para conformar signos de mayor complejidad de acuerdo con reglas morfológicas (que forman palabras) y sintácticas (que forman frases) que, en conjunto, conforman el *componente gramatical*. Por último, nuestro saber o capacidad para interpretar enunciados lingüísticos se divide en los componentes semántico y pragmático. El *componente semántico*, a su vez, se subdivide en nuestro conocimiento del significado de los signos simples del componente léxico y en nuestra capacidad para construir composicionalmente el significado de los signos complejos producidos gramaticalmente a partir de aquellos. El *componente pragmático* es la capacidad del hablante para inferir la intención última del interlocutor evaluando la información semántica de sus enunciados en relación con las características particulares del contexto de la comunicación. A continuación se examina cada uno de los componentes con un poco más de detalle.

1.2.1. El componente fonológico

El papel que cumple el componente fonológico de nuestro saber lingüístico es doble: proporciona las unidades que sirven para construir la expresión de las unidades del léxico (*d, u, r, a, n, t, e; c, o, r, t, é, s;* etcétera) y gobierna la pronunciación, así como la percepción auditiva, de los enunciados. En lo que respecta a la primera función, para poder memorizar, reproducir y reconocer los miles de morfemas que componen un léxico, la expresión de cada uno debe ser claramente distinguible de la de los otros. Por ejemplo, la diferencia entre los morfemas *sal* y *sol* es claramente distinguible gracias a las unidades *a* y *o*; de manera análoga, *sal* y *cal* se distinguen fácilmente gracias a la diferencia entre *s* y *c*, al igual que *sol* y *son*, gracias a *l* y *n*. Tales unidades

² No todas las descripciones de los sistemas lingüísticos coinciden en el número de componentes. Aquí presentamos una de las propuestas más aceptada.

discretas o discontinuas se conocen como *fonemas*. Su combinación en distinto orden y número permite formar infinitas expresiones nítidamente distinguibles unas de otras a partir de un repertorio finito de elementos combinatorios. Conforme aprenden los signos lingüísticos utilizados en su comunidad, los hablantes interiorizan el inventario de fonemas propios de su variedad lingüística. Así, los hablantes de quechua de Cusco deben aprender un inventario de fonemas consonánticos diferente que los hablantes de quechua de Junín³.

Inventario de fonemas consonánticos del quechua de Cusco

	bilabiales	alveolares	palatales	velares	postvelares	glotal
oclusivos simples	p	t	tʃ	k	q	
oclusivos aspiradas	p ^h	t ^h	tʃ ^h	k ^h	q ^h	
oclusivos globalizadas	p'	t'	tʃ'	k'	q'	
fricativos		s				h
nasales	m	n	ɲ			
laterales		l	ʎ			
vibrante		r				

Inventario de fonemas consonánticos del quechua de Junín

	bilabiales	alveolares	palatales	retrofleja	velar	glotales
oclusivos	p	t	tʃ	ʂ	k	ʔ
fricativos		s				h
nasales	m	n	ɲ			
laterales		l	ʎ			
vibrante		r				

La segunda función del sistema fonológico es gobernar la manera en que pronunciamos y percibimos auditivamente los enunciados. La observación cuidadosa de la forma en que se pronuncian los fonemas demuestra que esta está condicionada por el contexto fonológico en que se encuentran. Veamos el siguiente ejemplo en que se examina la pronunciación del morfema castellano *dos* propia de los hablantes de Lima.

³ Los signos gráficos que representan los fonemas son los del Alfabeto Fonético Internacional y su caracterización en términos de punto de articulación (columnas) y modo de articulación (filas) se basa en la forma en que se pronuncian más normalmente los fonemas.

Castellano de Lima

[dosarkos]	<i>dos arcos</i>	[dohpalos]	<i>dos palos</i>
[dosenanos]	<i>dos enanos</i>	[dohtamaños]	<i>dos tamaños</i>
[dosilos]	<i>dos hilos</i>	[dohtfinos]	<i>dos chinos</i>
[dososos]	<i>dos osos</i>	[dohkaros]	<i>dos carros</i>
[dosurones]	<i>dos hurones</i>	[dohfaros]	<i>dos faros</i>

Cuando el morfema *dos* aparece antes de un morfema que empiece con un fonema vocálico, su último fonema se pronuncia [s]; en cambio, cuando aparece antes de una palabra que empiece con un fonema consonántico, su último fonema se pronuncia [h]. Esto sugiere que los fonemas, como unidades que sirven para codificar los morfemas, son unidades abstractas y que pueden pronunciarse, de manera concreta, como sonidos distintos en contextos distintos de acuerdo con reglas.

Los fonemas, a su vez, están conformados por unidades constituyentes conocidas como *rasgos*, los que son responsables de las dos funciones del componente fonológico de distinguir y pronunciar los morfemas. Con respecto a la primera, vimos que los morfemas *sal*, *sol*, *cal* y *son* se distinguen por los fonemas *a*, *o*, *s*, *c*, *l* y *n*. Si bien esto es cierto, los fonemas no se diferencian unos de otros de manera integral, sino que comparten algunas propiedades y se distinguen por otras. Así, los fonemas *a* y *o* tienen en común el ser [vocálicos] y [no altos], pero se diferencian por ser el primero [central] y [no redondeado], mientras el segundo es [posterior] y [redondeado]. Por su parte, los fonemas *s* y *c* tienen en común el ser [consonánticos] y [no sonorantes], pero se diferencian por sus rasgos de modo y punto de articulación: el primero es [fricativo] y [alveolar], y el segundo es [oclusivo] y [velar]. Por último, los fonemas *l* y *n* son [consonánticos], [sonorantes] y [alveolares], pero el primero es [lateral] y el segundo, [nasal]. En lo que se refiere a la segunda función, son también los rasgos las unidades fundamentales con que operan las reglas que determinan la pronunciación diferente de los fonemas según el contexto en que se realizan. En el ejemplo anterior podemos apreciar que las dos realizaciones del fonema final del morfema *dos* —[s] y [h]— comparten los rasgos [consonántico], [fricativo] y [no sonoro], y se distinguen porque la primera es [alveolar] y la segunda es [glotal]. Cabe agregar finalmente que los componentes fonológicos de las variedades lingüísticas e idiomas seleccionan y organizan a su manera los rasgos que utilizan para construir y pronunciar los signos de su léxico.

1.2.2. El componente léxico

Las unidades del componente léxico se conocen como *morfemas* y como indicamos anteriormente se pueden caracterizar como expresiones interpretables, es decir, capaces de evocar contenidos mentales. Por esta razón, suele decirse que los morfemas son asociaciones de una expresión y un contenido o de un significante y un significado. Aparte de esta doble dimensión, los morfemas presentan asociada información de tipo gramatical, que corresponde a sus propiedades combinatorias de tipo morfológico y sintáctico. Los morfemas son entonces unidades de nuestro saber lingüístico formadas por características o rasgos de distintos componentes lingüísticos. Así, por ejemplo, el morfema *cortés* presenta una expresión formada por las seis unidades fonológicas *c, o, r, t, é, s*. Asimismo, esta expresión está asociada a un contenido semántico o significado léxico más o menos compartido por quienes usan este morfema. Por último, los hablantes saben también que el morfema *cortés* puede asociarse morfológicamente con un prefijo como *des-*, para formar la palabra *descortés*, o con un sufijo como *-es*, para formar la palabra *cortesés*, pero no con un sufijo de tiempo o de persona, ya que no podemos decir, por ejemplo, **cortes-ába-mos* para significar ‘éramos cortesés’. De manera similar, los hablantes saben que el morfema *cortés* se puede combinar sintácticamente con un adverbio como *excesivamente* en una frase como *excesivamente cortés* o que puede formar parte de una frase junto con un artículo y un sustantivo, como en *una persona cortés*.

Señalamos también que los morfemas son ‘signos o símbolos lingüísticos’, es decir expresiones que los hablantes asocian convencionalmente con algún tipo de contenido mental. De Saussure definió en ese sentido al signo lingüístico como «la combinación del concepto y de la imagen acústica» (1974[1916], p. 129), llamando luego al concepto *significado* y a la imagen acústica *significante*. Si por *imagen acústica* o *significante* entendemos la particular secuencia de fonemas que identifica a un signo lingüístico, ¿qué podemos entender por *concepto* o *significado léxico*? Si llamamos *referente* a aquel ítem particular al que aludimos con una unidad léxica en el contexto de un acto comunicativo, podemos definir el significado léxico como un conjunto de criterios, compartidos por una comunidad de habla, para seleccionar ítems en el entorno extralingüístico como referentes para dicha unidad léxica (Herrera, 2002, p. 359). Así, probablemente al escuchar el enunciado «Pásame la taza», un hablante del castellano preferirá seleccionar al ítem (A) antes que al ítem (B) como referente para la unidad léxica *taza*, porque los criterios que comparte con su comunidad de habla le indican que tales ítems deben tener un asa:



(A)



(B)

Es importante tener en cuenta que tales criterios reflejan el saber del mundo de los hablantes y no tienen que corresponder necesariamente a lo que los expertos nos pueden decir acerca de cómo es el mundo: expresan más bien las creencias que los miembros de la comunidad de habla comparten acerca de cómo es el mundo, las que hacen posible la comunicación. No se trata por ello de si las tazas tienen «en realidad» un asa o no, sino de que si los hablantes que usan la palabra *taza* no comparten esta creencia, pues no van a seleccionar en el entorno extralingüístico los mismos ítems como referentes para esa unidad léxica, con lo que pueden surgir problemas en la comunicación (De Saussure, 1974[1916]; Aitchison, 1994 y Herrera, 2002).

1.2.3. El componente gramatical

En la introducción indicamos que las unidades del componente léxico o morfemas se combinan formando signos de mayor complejidad mediante las reglas del componente gramatical. Puede decirse por tanto que la función principal de este componente es la de proporcionar los medios para construir signos complejos a partir de los signos simples del léxico. El componente gramatical se divide en un componente morfológico y un componente sintáctico. Las reglas del componente morfológico permiten combinar los signos del léxico o morfemas para formar palabras, y las reglas del componente sintáctico permiten combinar tales palabras para formar frases.

Podemos comprender el rol central que juega el componente gramatical si tomamos en cuenta que una proposición como la expresada por el signo *la descortés gata maulló durante tres horas* no se puede expresar en castellano con *un solo* signo léxico —pues no existe convencionalmente tal signo—, sino que debemos combinar distintos signos simples hasta conseguir una estructura mayor que signifique dicha proposición. Los signos simples de que disponemos —*la, des, cortés, gat, a, maull, ó, durante, tres, hora, s*—, sin embargo, no pueden combinarse de cualquier manera para lograr una estructura significativa. Ninguna de las tres combinaciones siguientes resulta interpretable por los hablantes de castellano:

- a) la cortés des gat a ó maull tres durante s hora
- b) la des gat a ó maull tres durante hora s cortés
- c) tres gat a s maull ó durante la cortés des hora

En cambio, cualquiera de las tres siguientes sí lo es:

- a) la des cortés gat a maull ó durante tres horas
- b) durante tres horas maull ó la des cortés gat a
- c) maull ó durante tres horas la des cortés gat a

La razón de esta diferencia en la posibilidad de interpretar las combinaciones de signos de (X) y (Y) radica en el conocimiento (parcialmente) compartido del componente gramatical: en (Y) reconocemos estructuras interpretables; en (X) no. Las estructuras más obvias de (Y) son [*la [des cortés gat a]*], [*maull ó*] y [*durante [tres hora s]*], que aparecen en órdenes sintácticos alternativos. Otro orden sintáctico que permitiría la interpretación de la combinación de signos sería [*la [gat a des cortés]*], pero no ocurriría lo mismo con [*[gat a des cortés] la*] ni con [*gat ala des cortés*].

1.2.3.1. El componente morfológico

Señalamos que las reglas del componente morfológico permiten combinar los signos del léxico o morfemas para formar unidades mayores que conocemos como palabras. En ese sentido, otras estructuras interpretables que reconocemos en (Y) pero no en (X) son las de índole morfológica: [*des cortés*], [*gat a*], [*maull ó*] y [*hora s*]. Para expresar lo contrario de *cortés*, el castellano cuenta con un signo léxico que debe colocarse justo antes de esta raíz: el prefijo *des-* (nótese que no basta la adyacencia puesto que no decimos [*cortés des*]). Esta misma información, por supuesto, puede expresarse mediante la combinación sintáctica de otros signos —[*lo contrario de cortés*]—, pero la construcción o palabra morfológicamente compuesta [*des cortés*] lo hace de manera más sintética.

Algo semejante ocurre con los otros tres ejemplos. La palabra [*gat a*] combina morfológicamente la raíz *gat-*, que denota una especie animal, con el sufijo *-a*, que restringe la denotación a las hembras de dicha especie. Por su parte, la palabra [*hora s*] combina morfológicamente la raíz *hor-* con el sufijo *-s*, que indica que el número de unidades de tiempo denotadas por la raíz es mayor de uno. En ambos casos, la información expresada por las construcciones morfológicas podría canalizarse mediante construcciones sintácticas —[*el ejemplar de felino doméstico hembra*] o [*un número de horas mayor a uno*]— pero, como se ve, la posibilidad de combinar signos simples para formar palabras compuestas resulta un recurso muy eficiente.

El último ejemplo —[*maull ó*]— es particularmente ilustrativo porque muestra la combinación morfológica de dos signos simples que no tienen una relación semántica directa. En efecto, el signo *-ó* expresa el tiempo pasado en que se desarrolla todo el evento descrito por la oración *la descortés gata maulló durante tres horas* —algo así como *hubo un tiempo anterior al momento de enunciación de esta oración* en el que la descortés gata maulló y ese tiempo duró tres horas— y, sin embargo, aparece formalmente sufijado a la raíz verbal.

Así pues, como parte del componente gramatical, muchas lenguas (no todas) incluyen, además del componente sintáctico, uno morfológico para la construcción

de palabras que incluyan más de un signo del léxico. Para ello, los signos se categorizan en tipos combinatorios como raíces, prefijos y sufijos. En algunas lenguas, inclusive, existen infijos que son una categoría combinatoria que se acomoda «dentro» de la raíz.

1.2.3.2. *El componente sintáctico*

Vimos que mientras las reglas del componente morfológico permiten combinar los signos del léxico o morfemas para formar palabras, las reglas del componente sintáctico permiten combinar tales palabras para formar frases. Así, al hablar (o escribir) escogemos y ordenamos palabras en secuencias; pero debemos hacerlo de cierta manera, pues no toda combinación de palabras comunica lo que queremos decir: para ello debemos formar frases, que son combinaciones de palabras bien ordenadas de acuerdo con reglas o patrones de construcción (Coral & Pérez, 2004, p. 29). Tales reglas o patrones de construcción constituyen el componente sintáctico del saber lingüístico. Aquellas combinaciones de palabras que no siguen tales reglas o patrones de construcción, como (1), nos resultan incomprensibles:

(1) Por entraba cuando agua abierto ellos compró desde.

No podemos atribuir un significado a (1) porque dicha oración no está construida siguiendo las reglas o patrones de construcción propios del castellano. Los hablantes suelen considerar inaceptables a tales combinaciones de palabras; los lingüistas dirán que son agramaticales. El componente sintáctico nos permite por tanto construir diferentes tipos de frases, combinando palabras pertenecientes a ciertas categorías gramaticales en base a reglas o patrones. Así, combinando un sustantivo como *casa*, un determinante como *la* y un adjetivo como *grande* podemos formar la frase nominal (o frase sustantiva) *la casa grande*; combinando una preposición como *de* con dicha frase nominal podemos formar la frase preposicional *de la casa grande*. Esta a su vez puede combinarse nuevamente con el determinante *la* y el sustantivo *puerta* para formar la frase nominal *la puerta de la casa grande*. Por otro lado, combinando un verbo como *abrir* con tal frase nominal podemos construir la frase verbal *abrió la puerta de la casa grande*. Si por otro lado combinamos el determinante *la* y el sustantivo *niña* para formar la frase nominal *la niña*, podemos finalmente combinar dicha frase nominal con la anterior frase verbal y formar la oración (2)⁴:

⁴ Para los conceptos de frase nominal o sustantiva, frase verbal, frase preposicional y oración ver Coral y Pérez (2004).

(2) La niña abrió la puerta de la casa grande.

Quizás el aspecto más distintivo del lenguaje humano es la capacidad de los hablantes de comprender y producir creativamente y de modo ilimitado expresiones lingüísticas en su lengua materna. A esta capacidad se le da el nombre de creatividad lingüística. El gran truco del lenguaje que hace posible la creatividad lingüística es que el hablante posee medios finitos (elementos léxicos y la posibilidad de combinarlos de acuerdo con reglas o patrones de construcción), que le permiten producir infinitas expresiones en su lengua. Se llama también *productividad sintáctica* a esta capacidad del componente sintáctico de combinar creativamente las palabras o elementos del léxico en frases y oraciones. Pero si bien las diferentes lenguas se diferencian por la forma en que se pueden construir frases y oraciones, la investigación lingüística ha encontrado que hay ciertas propiedades fundamentales que parecen estar presentes en la sintaxis de todas las lenguas. A continuación presentamos algunas de ellas.

– *Autonomía de la sintaxis*

Analicemos los siguientes ejemplos:

(3) a. La muchacha de vestido rojo ha comprado pantalones de tres diseñadores.

b. *¿De qué ha comprado la muchacha ___ pantalones de tres diseñadores?

Mientras el enunciado (3a) está bien construido, los hablantes de castellano no formulan preguntas como (3b) a partir de enunciados como (3a), y de hecho eso no ocurre casi en ninguna lengua. Pero no existe ningún problema lógico en formular una pregunta por una característica del agente que participa en el evento descrito (el vestido rojo de la muchacha); lo que falla en (3b) no es el contenido informativo que expresa, sino la forma en que está construido (específicamente, la extracción de una frase como *de qué* desde dentro de un sujeto). Por tanto, el componente sintáctico prohíbe ciertas estructuras aun cuando su contenido informativo sería de fácil interpretación. Esto muestra que dicho componente es autónomo. Esto se refleja también en el hecho de que podemos reconocer los elementos oracionales en el ejemplo (4), aun cuando desconozcamos algunos de los significados de los elementos léxicos (porque son términos técnicos o inventados).

(4) Los galebos soritosos mucularán las nepimas ogamente.

– *Jerarquía de constituyentes*

Como hemos visto anteriormente, los morfemas se combinan en constituyentes cada vez mayores (que corresponden a las frases nominales, preposicionales, verbales, etcétera, que hemos descrito más arriba). Un aspecto fundamental de la estructura

sintáctica es que tales constituyentes forman una estructura jerárquica, es decir que tienen entre sí relaciones de subordinación. Por ejemplo, en la oración *El profesor conversa con los cuatro estudiantes*, podemos encontrar las siguientes agrupaciones jerárquicas:

- (5) a. El profesor conversa con los cuatro estudiantes.
- b. [El profesor][conversa con los cuatro estudiantes]
- c. [El profesor][conversa [con los cuatro estudiantes]]
- d. [El profesor][conversa [con [los cuatro estudiantes]]]

– *Infinitud discreta y recursividad*

Hemos explicado que los signos simples del léxico, los morfemas, constituyen un inventario finito pero se pueden combinar gramaticalmente para formar infinitos signos complejos. Lo que hace que el producto de esas combinaciones sea infinito es el hecho de que existen reglas de combinación que pueden tomar una nueva expresión lingüística recién construida y combinarla con otra, y así sucesivamente. Incluso un tipo de frase puede estar contenido en otra de su mismo tipo. A esto se le llama *recursividad* y es lo que hace al lenguaje altamente productivo: con pocos elementos, se pueden producir muchos enunciados. Por ejemplo, una oración cualquiera siempre puede quedar subordinada a otra oración:

- (6) a. Jorge come arroz.
- b. Marcos cree que Jorge come arroz.
- c. Pablo piensa que Marcos cree que Jorge come arroz.
- d. César dice que Pablo piensa que Marcos cree que Jorge come arroz.
- e. Paola cree que César dice que...

– *Movimiento*

Por lo general, en las expresiones lingüísticas complejas, los constituyentes, como el sujeto o el objeto de la oración, se interpretan en relación con otros constituyentes con los que se encuentran relacionados estructuralmente (típicamente, con el verbo). Por ejemplo en *Jorge come arroz*, *Jorge* es el que come y *arroz* es lo comido: el sujeto de un verbo de acción se interpreta como el agente de la acción y su objeto:

- | | |
|---------------|------------|
| (7) 1. Sujeto | 2. Objeto |
| El que come | Lo comido |
| Jorge | come arroz |

Nótese que lo mismo se interpreta cuando el objeto es una palabra interrogativa:

(8) 1. Sujeto	2. Objeto
El que come	Lo comido
¿Jorge	come qué?

Y lo mismo se interpreta cuando el sujeto y el objeto se encuentran en posiciones distintas de las que tienen usualmente en las oraciones declarativas simples:

(9) 2. Objeto	1. Sujeto
Lo comido	El que come
¿Qué	come Jorge?

La interpretación de (9) es consistente con (7) y (8) aun cuando ha habido «movimiento» de los constituyentes de la oración. Cuando construimos preguntas en castellano, el constituyente *qué* se desplaza desde su posición de objeto del verbo hasta el comienzo de la oración y el verbo aparece delante del sujeto. Sin embargo, el movimiento de constituyentes no es irrestricto, sino que existen condiciones que lo limitan, como vimos con el ejemplo (3b).

– *Categorías vacías*

Tradicionalmente, se dice que hay elementos sobreentendidos, que no escuchamos, pero que podemos interpretar de alguna manera. Un caso típico es el sujeto tácito del castellano, que se ejemplifica en (10):

- (10) a. *cv* Quiero a mi hija.
 b. Paola quiere *cv* comprar un libro.

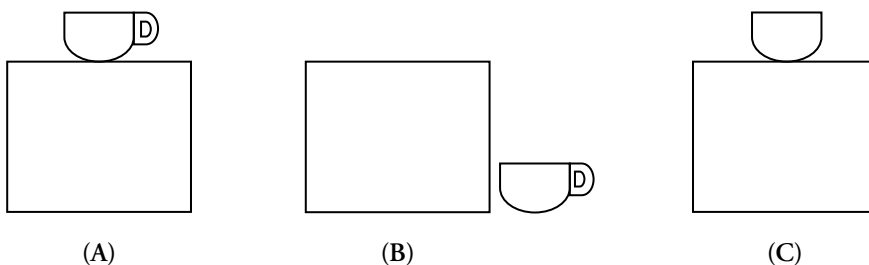
En (10a) hay un sujeto, pero realizado en una categoría vacía o *cv*, la que permite establecer la flexión verbal a través de la concordancia. En (10b) el predicado *querer* satura su requisito de dos argumentos con *Paola* y *comprar un libro*. Análogamente, el predicado *comprar* satura su requisito con el argumento *un libro* y con la categoría vacía *cv*, cuya interpretación, en este ejemplo específico, depende del sujeto de la cláusula principal, *Paola*.

1.2.4. El componente semántico

El *componente semántico* o *saber semántico* puede definirse como ‘el saber que hace posible al hablante comprender el significado de las expresiones lingüísticas’. Como vimos en la parte inicial de esta sección, este componente se subdivide en nuestro conocimiento del significado de los signos simples del componente léxico y en nuestra capacidad para construir composicionalmente el significado de los signos complejos producidos gramaticalmente a partir de aquellos.

El saber semántico tiene dos aspectos. El primero es el permitir al hablante relacionar oraciones con situaciones, o *referencia*. Así, si un hablante sabe el significado de la oración (11), seguramente dirá que se refiere a la situación (A) y no a (B) ni a (C):

(11) La taza está sobre la caja.



Los hablantes solemos caracterizar la particular relación que existe entre un enunciado como (11) y una situación como (A) con la familiar noción de *verdad*: decimos que (11) es verdadera en (A) y que no es verdadera en (B) ni en (C). Cuando un hablante conoce el significado de una oración, sabe cómo tiene que ser el mundo para que esa oración sea verdadera.

El segundo aspecto del saber semántico es el permitir al hablante relacionar oraciones entre sí, o establecer *relaciones semánticas* entre las oraciones, como implicación, equivalencia o contradicción. Por ejemplo, si un hablante sabe que (11) es verdadera, entonces sabe que (12) y (13) también son verdaderas, o que (11) implica a (12) y a (13):

(12) Algo está sobre la caja.

(13) Un recipiente para tomar líquidos está sobre la caja.

Hemos visto que el saber sintáctico es productivo: a partir un inventario finito de elementos léxicos y de un conjunto finito de reglas sintácticas es posible producir creativamente infinitas oraciones. Pero además los hablantes somos capaces de comprender cualquier expresión que nuestro saber sintáctico puede producir, por lo que nuestro saber semántico también debe ser productivo. Esto tiene que ver con una propiedad muy importante del significado de las expresiones lingüísticas, que es la *composicionalidad*: el significado de una expresión lingüística compuesta es función del significado individual de sus partes y de cómo están combinadas. La composicionalidad del significado lingüístico es la que hace posible la productividad del saber semántico: la manera en que otorgamos un significado a cualquier expresión construida por la sintaxis (como frases u oraciones) es asignar un significado a cada uno

de los elementos léxicos que la componen, y luego construir composicionalmente a partir de ellos el significado de la expresión en su conjunto. Para explicar cómo ocurre este proceso retomemos la siguiente oración, vista en relación al componente sintáctico, y que repetimos aquí como (14):

(14) Los galebos soritosos mucularán las nepimas ogamente.

A pesar de que desconocemos el significado de varias de las palabras contenidas en (14), somos capaces de reconocer, por ejemplo, que la oración describe un evento denotado por el verbo *mucular*, que dicho evento tendrá lugar en el futuro, que en él hay dos participantes (dos entidades), denotados por las frases nominales *los galebos soritosos* y *las nepimas*, que el primero tiene el rol de agente y el segundo el de paciente, que *soritosos* describe una propiedad de la entidad *galebos* y que *ogamente* describe una propiedad del evento *mucular*. Toda esta información, de carácter claramente semántico, forma parte de lo que se conoce como *significado estructural* y es derivada a partir de elementos estructurales o formales de la oración, proporcionados por la sintaxis. Este constituye el marco de relaciones conceptuales en el que se insertan los significados de las palabras individuales o *significados léxicos*, que en este ejemplo desconocemos, como los sustantivos *gálivos nepimas*, el adjetivo *soritoso*, el verbo *mucular* o el adverbio *ogamente*, que pertenecen a las llamadas *palabras de contenido*. Si reemplazamos tales palabras por otras palabras de contenido conocidas, como los sustantivos *marineros* y *olas*, el adjetivo *intrépidos*, el verbo *enfrentar* y el adverbio *valientemente*, obtendremos la oración (15):

(15) Los intrépidos marineros enfrentarán las olas valientemente.

Aquí podemos ver cómo estos significados léxicos individuales proporcionados por el componente léxico se combinan en el marco brindado por el significado estructural, constituido por las relaciones de significado transmitidas por los elementos gramaticales de la oración, para producir composicionalmente el significado completo de la oración o significado oracional (Chierchia & McConnell-Ginet, 2000; Cruse, 2000; Saeed, 1997).

1.2.5. El componente pragmático

El componente pragmático o saber pragmático puede definirse como ‘el saber que hace posible al hablante interpretar el sentido de los enunciados lingüísticos en contextos concretos de comunicación’. Hemos visto que el componente semántico permite al hablante derivar composicionalmente el significado de la oración o *significado oracional*. Sin embargo, hay situaciones en las que el oyente debe ir más allá del significado oracional si quiere descubrir la intención comunicativa del hablante.

Así, la oración (1) expresa el significado oracional de que hoy es el día de la semana que está entre el sábado y el lunes. Pero, en el contexto del diálogo (i), (1) adquiere un sentido adicional:

(16) Hoy es domingo.

(i)

(01) A: ¿Me acompañas a cobrar un cheque?

(02) B: Hoy es domingo.

Dicho significado adicional puede parafrasearse más o menos así: ‘no vale la pena ir, pues no vas a poder cobrar tu cheque’. Este significado adicional se conoce también como *significado del hablante*. Si definimos el significado oracional como ‘lo que la oración dice’, podemos definir al significado del hablante como ‘lo que el hablante quiere decir al emitir esa oración en ese contexto’. El significado del hablante expresa por tanto la intención comunicativa del emisor. La derivación del significado oracional, aplicando nuestro saber fonológico, sintáctico y semántico, puede caracterizarse como un proceso de *decodificación*. Frente a ello, la derivación del significado del hablante constituye más bien un proceso de *interpretación*, que requiere que hablante y oyente compartan un saber del mundo, en el ejemplo las creencias de que (a) los cheques se cobran en los bancos y (b) los domingos los bancos están cerrados.

De este modo, el componente pragmático nos permite interpretar los enunciados en el contexto comunicativo más allá de su significado oracional, el que por su lado es independiente del contexto y que es producto del componente semántico. En ejemplos como este el significado oracional y el del hablante no coinciden y la interpretación es necesaria: hablamos aquí de comunicación implícita. Pero hay ocasiones en que sí coinciden y en las decimos lo que queremos decir, de modo que no tenemos que interpretar nada (por ejemplo, si B hubiera respondido directamente «No te acompañe»). Hablamos aquí de comunicación explícita.

En los casos de comunicación implícita juega un rol central la expectativa de los participantes en el diálogo de que sus interlocutores reaccionarán ante sus intervenciones con contribuciones que sean relevantes (es decir, que vengan al caso). En el ejemplo, la falta de relevancia del significado oracional de (02) señala al oyente que debe someter a dicho enunciado a un proceso de interpretación para encontrar un significado del hablante que sí sea relevante; en dicho proceso, la búsqueda de relevancia funciona como una brújula que orienta al oyente y le permite reconocer cuáles son las creencias (que forman parte de su saber del mundo) que debe activar: son aquellas que permitan derivar el significado del hablante más relevante posible.

Otro aspecto muy importante de nuestro saber pragmático tiene que ver con ser capaces de reconocer qué acto de habla está tratando de realizar el hablante al emitir un enunciado en un contexto en particular. Sabemos así que la función principal del lenguaje no es representar la realidad, sino realizar acciones de diverso tipo. Enunciados como *Pepe está en la casa* o *Abro la ventana* describen situaciones en relación a las cuales pueden ser verdaderos o falsos, si bien en el segundo caso dicha situación es una acción no verbal del propio hablante. Pero cuando el Presidente del Comité Olímpico Internacional emite el enunciado *Declaro inaugurados los juegos olímpicos* en un estadio rodeado por varios miles de personas, dicho enunciado no describe o representa la acción de inaugurar los juegos olímpicos: su emisión misma, si se cumplen ciertas condiciones, constituye dicha acción. Este tipo de enunciados, para los que no vale la dicotomía verdadero/falso, se conocen como *performativos*: si no se cumplen las condiciones exigidas, no diremos que el enunciado es falso, sino que no tiene éxito. Pero este tipo de enunciados no son un fenómeno marginal del lenguaje, sino que cada vez que decimos algo, estamos realizando un tipo de acción llamado *acto de habla*, como prometer, ordenar, preguntar, aconsejar, amenazar, perdonar, informar, advertir, agradecer, felicitar, etcétera. Si digo «Prometo que tendré mi artículo listo la próxima semana», no estoy describiendo mi acto de prometer, sino que estoy realizando dicho acto mediante la emisión misma de dicho enunciado. Ni siquiera es necesario que utilice el verbo *prometer* para realizar dicho acto de habla, basta con decir «Tendré mi artículo listo la próxima semana». La relación entre el enunciado lingüístico y el tipo de acto de habla que se realiza con él puede describirse en los siguientes términos: la emisión de un determinado enunciado en un contexto dado (también llamada acto locucionario) *cuenta como* la realización de un tipo de acto de habla (también llamado acto ilocucionario), si se cumplen ciertas condiciones. Una parte muy importante de nuestro saber pragmático consiste, entonces, en conocer qué tipo de actos de habla son posibles y cuáles son las condiciones que deben cumplirse para su realización. Esto es necesario tanto para producir los actos de habla como para reconocer el tipo de acto de habla (acto ilocucionario) que nuestro interlocutor está tratando de realizar mediante la emisión de un determinado enunciado.

Finalmente, hay que recordar que en la comunicación verbal no nos limitamos a realizar actos de habla individuales y aislados, sino que interactuamos verbalmente con otros hablantes. La unidad básica de la comunicación verbal es por ello el diálogo. Por ello, el saber pragmático también incluye la capacidad de iniciar, mantener y finalizar diálogos o conversaciones, de participar en secuencias conversacionales y de introducir, negociar y cambiar temas conversacionales (Blakemore, 1992; Escandell, 1993; Heritage, 2001; Levinson, 1983 y Renkema, 1993).

1.3. Síntesis

Un rasgo propio de los seres humanos es que podemos hablar y para hacerlo, hacemos uso de una lengua natural. Iniciamos este capítulo con una distinción necesaria entre *lengua* y *lenguaje* para explicar esta característica exclusiva de nuestra especie. Comenzamos definiendo una lengua como un fenómeno cultural transmitido y mantenido por una comunidad a lo largo de su historia, pero también como un fenómeno cognitivo, como conocimiento particular que un hablante posee del conjunto de reglas y unidades lingüísticas que utiliza para comunicarse. A partir de esta distinción, presentamos el lenguaje como la capacidad propiamente humana que nos permite adquirir ese conocimiento.

En tanto fenómeno cultural, notamos que las lenguas modernas —como el castellano, el quechua y el inglés— constituyen el estado actual de un proceso de *cambio lingüístico*, solamente reconstruible, según algunos estudiosos, hasta un estadio no mayor a 5000 o 7000 años atrás, pero que se presume comenzó con la primera salida de nuestros antepasados homínidos del África septentrional. Hicimos hincapié en que el cambio es una característica propia de las lenguas humanas y en que no debe confundirse con la evolución del lenguaje, de la que trataremos, exclusivamente, en el capítulo a continuación.

En tanto fenómeno cognitivo, observamos que una lengua puede ser entendida como un sistema estructurado a partir de cinco componentes: el *componente fonológico*, que recoge los sonidos y las reglas de pronunciación de una lengua; el *componente léxico*, que almacena el conjunto de expresiones lingüísticas interpretables; el *componente gramatical*, encargado de la combinatoria de las piezas léxicas y, a su vez, dividido en dos sistemas especializados: el componente *morfológico*, para la formación de palabras, y el *sintáctico*, para la generación de oraciones. Por último, el *componente semántico*, que hace posible al hablante interpretar el significado de las expresiones lingüísticas, y el *pragmático*, que posibilita la interpretación del sentido de tales expresiones en contextos concretos de comunicación. Estos componentes lingüísticos constituyen el conocimiento que un hablante tiene de su lengua, conforman así la estructura del lenguaje. En el capítulo a continuación, exploramos las teorías más resaltantes en torno a la aparición de la capacidad lingüística en la evolución de la especie.

2. LA EVOLUCIÓN DEL LENGUAJE

Los textos introductorios al estudio de la evolución humana suelen dedicar un capítulo entero o una sección al origen del lenguaje en nuestra especie (Lewin, 1999; Delson y otros, 2000). Luego de la mención obligada al hecho de que no es posible encontrar restos fósiles de esta característica propia del ser humano, se examinan

los dos tipos de evidencia material que podrían darnos luces acerca de cuándo aparece la capacidad lingüística en nuestro linaje: primero, huesos de nuestros ancestros que indicarían de manera indirecta su capacidad para producir y procesar el habla y, segundo, logros culturales que indicarían una capacidad simbólica y de organización grupal solo obtenible mediante el lenguaje.

Los estudios de la base del cráneo, por ejemplo, permiten reconstruir tentativamente la forma y posición que debió de tener en nuestros ancestros el aparato fonador, principalmente constituido de tejido blando, y calcular aproximadamente desde cuándo habrían podido aquellos articular la diversidad de sonidos vocálicos y consonánticos característicos del lenguaje. Los resultados sugieren que ya el *Homo erectus* presentaba un cambio en dirección a la estructura actual, pero es recién el *Homo sapiens* el que muestra (hace entre 300 000 y 400 000 años) una configuración similar a la nuestra (Laitman, 2000)⁵. De otro lado, los estudios de la parte interior del cráneo ayudan a construir un modelo, aunque imperfecto, de cómo era el cerebro de nuestros ancestros y permiten conjeturar sobre el origen de las estructuras neuronales que en el cerebro actual se asocian con la producción del habla. Los resultados no son muy precisos: Deacon (1992) reporta que se han encontrado huellas correspondientes al área de Broca, en el *Homo habilis* y en el *Homo erectus*, pero advierte que esto no demuestra que esta estructura ya cumpliera una función lingüística; Lewin (1999) reporta que la reorganización cerebral que se encuentra en los australopitecinos hace pensar a algunos investigadores que la capacidad lingüística ya estaba presente en estos, pero afirma también que esta idea no cuenta con consenso en la comunidad científica.

Por su parte, los estudios sobre el desarrollo cultural señalan que la especie humana debía dominar un medio de comunicación bastante expresivo para poder construir herramientas orientadas hacia propósitos específicos, para desarrollar sistemas complejos de cooperación y competencia entre grupos, y para producir arte, mitos y ritos. La evidencia tampoco es concluyente en este ámbito de nuestro pasado: resulta claro que hace entre 50 000 y 30 000 años ocurrió un importante salto en nuestro desarrollo cultural, que, en general, no deja duda de que nuestros ancestros

⁵ Los estudios de la relación entre la forma de la base craneal y la posición del aparato fonador, en particular de la laringe, fueron impulsados por Philip Lieberman a comienzos de la década de 1970. Lieberman, sin embargo, sostiene que la tecnología de aquella época no permitía hacer mediciones precisas, lo que llevó a conjeturas hoy comprobadamente falsas. Según él, «la curvatura de la base del cráneo en sí misma no se puede usar para predecir si un fósil tenía un tracto vocal supralaríngeo propio de un adulto humano» (2007, p. 46). Y añade: «a pesar del foco en la *laringe* de muchos estudios sobre la evolución del habla, los factores principales en el desarrollo y evolución del tracto vocal humano son el descenso y cambio de forma de la *lengua*[...] y la posición y forma de la lengua no pueden inferirse del ángulo de la base craneal» (pp. 46-47; las cursivas son nuestras).

ya contaban con un dominio lingüístico sofisticado. Sin embargo, es difícil rastrear desde cuándo dominaban un sistema así o determinar si el origen y desarrollo de este sistema complejo acompañó gradualmente el progreso social y tecnológico de nuestra especie.

Ahora bien, en este capítulo presentamos otro tipo de aproximación al estudio de la evolución del lenguaje. Nos interesa caracterizar la lógica de su aparición en nuestros ancestros, así como su ventaja adaptativa, antes que fechar el momento de su aparición o determinar con precisión la especie homínido con la que pudo haber emergido. En §2.1., presentamos algunas de las hipótesis sobre el origen y evolución del lenguaje que se han propuesto principalmente desde la lingüística. Estas propuestas se concentran en el aspecto formal del lenguaje —en la organización de los componentes lingüísticos que revisamos en el capítulo anterior— antes que en su función comunicativa. Seguidamente, en §2.2., presentamos algunas propuestas sobre la aparición filogenética del lenguaje que priorizan sus ventajas adaptativas para la organización social y cohesión grupal en nuestros ancestros.

2.1. Las principales propuestas sobre el origen y evolución del lenguaje

Hemos convenido en dividir este apartado en dos partes. La primera reseña dos propuestas importantes sobre la evolución del lenguaje en etapas que coinciden con los diferentes aspectos que definen nuestra capacidad lingüística. Así, los autores que revisaremos se preocupan, por ejemplo, por el surgimiento de la representación simbólica en la especie; por la aparición de la combinatoria recursiva, característica de las lenguas naturales; o por los mecanismos que llevaron a la adaptación del tracto vocal homínido para la producción del habla. La segunda parte se concentra en un debate, surgido hacia finales de la década de 1980, sobre qué componentes de la facultad del lenguaje fueron, en su origen, producto del moldeamiento uniforme de la selección natural. Se trata así de un intercambio de argumentos sobre qué mecanismo de la evolución habría conducido a la aparición de elementos específicos del lenguaje.

2.1.1. La evolución del lenguaje en etapas

Presentamos aquí las propuestas de Derek Bickerton (1990, 2009 y 2010) y Ray Jackendoff (1999 y 2002) en torno a la aparición de componentes característicos del lenguaje. Bickerton fue uno de los primeros investigadores en preguntarse por la evolución del lenguaje desde la lingüística. Su propuesta sobre la evolución del lenguaje se enfoca en dos de sus aspectos característicos: la simbolización y la combinatoria recursiva. La propuesta de Jackendoff, por su parte, distingue nueve etapas en la evolución de lenguaje que van desde la aparición de vocalización simbólicas hasta el refinamiento de rasgos prototípicos de la sintaxis.

2.1.1.1. *La propuesta de Derek Bickerton*

Bickerton propone abordar la evolución del lenguaje de manera distinta de como lo han hecho los evolucionistas funcionalistas. El autor tomará observará al lenguaje como un mecanismo formal para la representación del pensamiento. Tres aspectos son centrales en su propuesta: la diferencia entre el lenguaje y los sistemas de comunicación animal, la representación simbólica y la combinatoria de unidades lingüísticas o sintaxis.

- *Diferencias entre la comunicación animal y el lenguaje*

Bickerton sostiene que el lenguaje cumple funciones comunicativas, pero lo hace de manera diferente a como lo hacen los sistemas de comunicación animal. La forma que tiene el lenguaje actualmente no puede ser el resultado de la evolución gradual de esos sistemas, ya que sus características son cualitativamente diferentes.

En primer lugar, la comunicación animal se vincula exclusivamente con la supervivencia de los individuos: enfrentar el peligro, buscar emparejamiento y reproducirse, y sus relaciones sociales. El lenguaje, en cambio, se vincula con muchos más y diversos aspectos de la vida en sociedad. En segundo lugar, la comunicación animal expresa información estrictamente subjetiva (por ejemplo, las emociones actuales del animal, sus intenciones, lo que desea que otros animales hagan, etcétera); el lenguaje, por su parte, expresa todo lo anterior y también información objetiva acerca del mundo. En tercer lugar, la comunicación animal es indexical o icónica, y finita; en contraste, el lenguaje no es finito y posee una dimensión simbólica única en el mundo animal. En cuarto lugar, la diferencia más resaltante radica en las unidades que se emplean. De un lado, en la comunicación animal, estas unidades aparecen en forma de sonidos o secuencias de sonidos, gestos y movimientos corporales, o cambios de color y contorno, todos los cuales pueden clasificarse semánticamente. Sin embargo, no son unidades descomponibles ni pueden conectarse unas con otras, no generan nuevos significados si son enunciadas en secuencia y refieren únicamente a situaciones específicas del mundo real en el aquí y el ahora. De otro lado, las unidades del lenguaje son semánticamente discretas y aparecen en forma de secuencias de signos sonoros o de signos manuales (palabras). Asimismo, son descomponibles y, normalmente, aparecen conectadas con otras unidades en secuencias que pueden cambiar de significado dependiendo de la forma en que estén combinadas las unidades o de la información sobresaliente del contexto. Es más, las unidades y las secuencias pueden relacionarse con cualquier momento o lugar y se refieren solo indirectamente a entidades, acciones o situaciones del mundo real a través de representaciones conceptuales.

Esta comparación lleva a Bickerton a afirmar que, si se busca un antecedente del lenguaje, lo más intuitivo será buscarlo en los sistemas de comunicación animal; sin embargo, la naturaleza del lenguaje no guarda relación evolutiva con los sistemas de comunicación animal, pues las diferencias son más cualitativas que cuantitativas. A este problema lo denomina *paradoja de la continuidad*. Para superar la paradoja de la continuidad, es necesario asumir que el lenguaje es un sistema de representación, antes que un sistema eminentemente comunicativo. Así, el antepasado del lenguaje no está en los anteriores sistemas de comunicación animal, sino en sus sistemas de representación. En efecto, el simbolismo del lenguaje es un rasgo crucial de él y constituye, según Bickerton, un paso importante en su evolución.

– *El lenguaje y su dimensión simbólica*

Para Bickerton, dos son los rasgos del lenguaje que deben ser explicados para poder dar cuenta de su evolución: el simbolismo y la sintaxis. En relación con la capacidad simbólica del lenguaje, Bickerton emprende la tarea de rastrear el origen del lenguaje en los sistemas de representación⁶ de los seres vivos. El autor propone que los sistemas de representación en la línea evolutiva de la especie humana se habrían hecho cada vez más complejos. Así, pueden distinguirse sistemas de representación primarios y secundarios: los primeros corresponden a representaciones surgidas a partir de la relación directa con el entorno físico y social; los segundos, a representaciones de las representaciones. El lenguaje (junto con la conciencia y la capacidad metarrepresentacional) sería, pues, un sistema de representación secundario.

Para explicar el surgimiento de la capacidad simbólica que da pie a la evolución del lenguaje, Bickerton se concentra en el rasgo de desplazamiento, es decir, en la habilidad para separar los signos lingüísticos del entorno inmediato en que son emitidos. Como señalamos en el apartado anterior, las llamadas de los animales dependen de la situación en la que son expresadas para comunicar información específica dentro de ella o generar reacciones particulares en sus receptores, siempre de interés crucial para la vida de los individuos del grupo en el que se utilizan.

⁶ La noción de representación que maneja Bickerton es bastante amplia: «significa, simplemente, responder o tener una propensión permanente a responder a x , una entidad o un suceso en el mundo exterior, en términos de y , una pauta determinada de actividad neurológica» (1990, p. 106). No debe asombrarnos, entonces, que hablemos del sistema de representación de las larvas, por ejemplo, que han de reaccionar de manera diferente con respecto a la cantidad de luz que ingrese a través de los pliegues de las cortezas de los árboles en donde se esconden. Un ligero cambio en la intensidad de luz recibida le indicará al organismo que se encuentra o no ante la amenaza de un depredador que hurga por comida entre los pliegues de la corteza. Estas diferencias en el input del ambiente determinarán que la larva reaccione con el propósito de salvar su vida o no reaccione en absoluto.

Esto significa que las llamadas de los sistemas de comunicación animal no pueden divorciarse de su aptación (*fitness*) al entorno. El lenguaje contrasta con esta situación.

Bickerton propone que el desplazamiento fue el primer paso hacia la aparición del lenguaje. Por ello, emprende la búsqueda de un escenario específico que pudo haber ejercido una presión selectiva para el surgimiento del desplazamiento exclusivamente en nuestra especie. De acuerdo con el autor, ninguna otra especie primate ha creado o habitado un nicho ecológico en el que se haya podido requerir el desplazamiento⁷.

La comunidad de nuestros antepasados homínidos hace dos millones de años (*Homo habilis* y *Homo erectus*) habría necesitado cohesionarse como grupo para poder explotar recursos ecológicos, principalmente, aprovechar los cadáveres de la megafauna como fuente de alimento. Dado que estas fuentes de alimento distaban espacial y temporalmente del lugar de asentamiento del grupo, el desplazamiento se volvió un elemento necesario. Para poder vencer a sus competidores carroñeros, el grupo que había descubierto un cadáver enorme debía convencer a los otros grupos que lo encontrado valía la pena más que otros hallazgos y, dado que la mayoría de los primates son altamente competitivos y poco cooperativos, solo podía lograrse este objetivo con señales que transmitieran el tipo de comida, el lugar, la distancia y la dirección para encontrarla, así como el cuidado que habría que tomar y la celeridad para realizar la búsqueda. El uso de señales que fueran al menos potencialmente simbólicas habría sido esencial para poder reclutar individuos que pudieran emplear herramientas como proyectiles o como instrumentos para cortar. En otras palabras, reclutar individuos en ausencia de evidencia sensorial exige que el sistema de comunicación que se emplea cuente con el rasgo de desplazamiento.

Bickerton señala que este tipo de 'lenguaje' arcaico habría requerido únicamente un pequeño conjunto de unidades simbólicas, sin reglas de combinación, lo que, sin duda, estaba al alcance de las capacidades cognitivas de nuestros ancestros. La presión selectiva para el simbolismo habría sido, pues, un nicho ecológico en el que había mucha competencia por conseguir carroña y que, por esta razón, exigía que nuestros ancestros reclutaran individuos a gran escala para conseguir el alimento para el grupo. La hipótesis presentada por Bickerton permite explicar no solo el surgimiento del lenguaje, sino también la aparición de otro rasgo cognitivo sobresaliente que distingue a los humanos de las otras especies primates: la cooperación.

⁷ Hay otras especies no primates que habitan un nicho ecológico en el que han requerido desplazamiento en su comunicación, como las hormigas y las abejas. Sin embargo, no podemos decir que su forma de comunicación es un precursor del lenguaje, sino, simplemente, que se trata de un tipo particular de respuesta a un nicho con ciertas características (Bickerton, 2010).

- *Lenguaje y protolenguaje*

Además de unidades simbólicas, que expresan información en ausencia de evidencia sensorial, el lenguaje necesita sintaxis, es decir, un procedimiento que permita la combinatoria, potencialmente infinita, de esas unidades. Sobre el origen de la sintaxis, Bickerton propone que, primero, debió existir una combinatoria mínima, muy simple, que incluso se mantiene hasta la actualidad: el protolenguaje.

El protolenguaje es un modo de concatenación de unidades simbólicas que puede estar al alcance de otras especies sin estas son entrenadas. Se trata de un mecanismo de yuxtaposición de elementos que no cuenta con indicios de una morfología compleja, como la que podemos observar en las lenguas actuales, con finos matices para indicar tiempo, aspecto, modo, concordancia de tiempo o modalidad. En claro contraste, en el lenguaje, el orden de las expresiones obedecen a una estructura, existen elementos fonéticamente nulos pero interpretables, existen predicados (que tienen una estructura argumental), hay mecanismos de expansión y recursividad, y abundan los elementos gramaticales. El protolenguaje carece de estas propiedades.

El autor considera que, actualmente, podemos encontrar indicadores de cómo podrían haber sido los estados iniciales de la sintaxis en el lenguaje en la especie. Entre estos *fósiles* del lenguaje, se encuentran los niveles de combinatoria que alcanzan los simios entrenados, los niños menores de dos años, los humanos que no han estado expuestos a una experiencia lingüística suficiente (como la niña Genie, que fue privada de contacto lingüístico durante el comienzo de su vida) y aquellos que generan sistemas simples que puedan ser empleados como lenguas francas (*pidgins*).

La distinción entre protolenguaje y lenguaje pleno (que ostenta una sintaxis) está vinculada con estadios diferentes de la evolución del lenguaje: debe haber existido una fase intermedia entre la ausencia de lenguaje y el lenguaje pleno. En este sentido, Bickerton asume que la aparición del protolenguaje y la de la sintaxis son hechos distintos, bastante separados en el tiempo.

El lenguaje apareció como un conjunto de símbolos, ya sean palabras o gestos. Su concatenación no mostraba ningún tipo de estructura entre ellos, sino únicamente una sucesión en el tiempo. Aunque el *Homo habilis* y el *Homo erectus* compartían más rasgos de conducta con los simios actuales que con los humanos actuales, sus cerebros eran solamente un poco más pequeños que los nuestros (hasta 1200 centímetros cúbicos). Por esta razón, se podría especular que ellos tendrían un protolenguaje, pero esto no es nada seguro. Lo que sí se puede determinar es que, solamente una vez que el protolenguaje alcanzó un cierto nivel de sofisticación (aunque no muy elevado), pudo ser utilizado exitosamente no solo para comunicar, sino también para engañar y manipular, para cooperar y cohesionar grupos, para chismear, y para todas las otras funciones que cumple el lenguaje de acuerdo con los estudiosos de la inteligencia social.

En otras palabras, no es posible que estas funciones dieran origen al protolenguaje, sino al revés: el protolenguaje debe ser previo a estas funciones.

Para Bickerton, la sintaxis aparecería de manera abrupta en la historia evolutiva de la especie. Las estructuras cognitivas del cerebro permitían a nuestros antepasados hacer un cálculo social, es decir, contar con una habilidad para distinguir no solo a los individuos del entorno y los tipos de acción y eventos, sino también crear una representación de los roles de cada individuo en un evento determinado. En algunas oportunidades, un individuo era el agente de una acción; en otras, era el paciente; en otras, era el beneficiario. Estos datos del entorno generan estructuras relacionales semánticas, que el individuo almacena para poder manejarse de manera exitosa en su grupo social. Dado que no hay nada completamente nuevo en la evolución y que la semántica existía antes que la sintaxis, Bickerton propone observar es esas relaciones semánticas la base para las estructuras oracionales.

En efecto, la sintaxis aparece hace 150 mil años cuando se comienzan a emplear las relaciones semánticas como elementos obligatorios en los enunciados. Esto es lo que Bickerton denomina *cartografía del cálculo social en el protolenguaje*. Los elementos primordiales de la sintaxis son, entonces, exaptaciones de la representación de las relaciones sociales.

La aparición de dos de los rasgos más representativos de la sintaxis de las lenguas puede explicarse en esta perspectiva. En primer lugar, el orden jerárquico es una consecuencia de las relaciones semánticas entre individuos y eventos; el orden lineal en que se producen los enunciados es una imposición del medio físico en el que se expresa el lenguaje. En segundo lugar, la recursividad es consecuencia también de la estructura de eventos. Dadas las condiciones, cualquier evento puede ser participante de otro evento. Estas relaciones entre eventos permiten que su expresión lingüística sea altamente recursiva: toda oración siempre puede ser incluida en otra más grande que ella.

Una vez aparecida la sintaxis hace 150 mil años, poco a poco va complejizándose y tomando la forma actual. Bickerton supone que diversos sistemas sintácticos entraron en competencia entre ellos con el fin de eliminar la ambigüedad de las oraciones y hacerlas más aptas. Su propuesta es que la sintaxis habría operado ella misma como un elemento de presión selectiva favoreciendo cualquier cambio en el sistema nervioso que facilitara la construcción e interpretación de enunciados. Un conjunto de sucesivas adaptaciones habrían mejorado la aptitud de los individuos y su pool genético. Bickerton cree que la selección orgánica o efecto Baldwin⁸ jugó un papel

⁸ La selección orgánica consiste en un proceso evolutivo por el cual se van convirtiendo en un instinto aquellos rasgos que, en un principio, constituían conductas aprendidas.

importante en la evolución de la sintaxis. Generación tras generación, los niños realizaban de forma espontánea lo que inicialmente era un proceso consciente en sus mayores. La sintaxis queda instalada en la especie un par de miles de años después de su aparición.

2.1.1.2. *La propuesta de Ray Jackendoff*

Sobre la base de la hipótesis del protolenguaje planteada por Derek Bickerton, el lingüista Ray Jackendoff ha propuesto una serie de posibles etapas en la evolución del lenguaje, moldeadas por la selección natural. El protolenguaje de Bickerton parece una etapa plausible en la evolución del lenguaje, pero no la única: desde el símbolo hasta el lenguaje pleno, Jackendoff postula nueve etapas sobre la base de la evidencia que encuentra en la estructura de las lenguas actuales, a manera de los *fósiles* de Bickerton. Su evidencia se concentra en el desarrollo lingüístico infantil, en los resultados de la enseñanza del lenguaje a los primates, en el aprendizaje de una segunda lengua en contextos no formales, en las formas lingüísticas de los niños sordos hijos de padres no sordos, en los pidgins y en las formas lingüísticas de los pacientes afásicos.

La argumentación del autor parte de tres asunciones explícitas. En primer lugar, cualquier incremento en el poder de expresividad de un sistema comunicativo tendría que ser necesariamente adaptativo (en el sentido de que serviría para la cooperación en la caza, el agrupamiento y la defensa, así como para el establecimiento de roles y actividades en sociedades primates complejas). En segundo lugar, las adaptaciones lingüísticas *primero* habrían mejorado la comunicación y *después* habrían refinado el pensamiento. Finalmente, la evolución del lenguaje habría procedido mediante un canal vocal-auditorio, lo que no imposibilita un estadio previo en el que pudiera haberse hecho uso de un sistema de signos gestual. Las dos primeras asunciones constituyen un apoyo a la teoría adaptacionista presentada años antes por Steven Pinker y Paul Bloom (1990).

La primera etapa muestra que las vocalizaciones simbólicas voluntarias se habrían concentrado en un solo símbolo. Sin embargo, a diferencia de las llamadas de otras especies (que emplean elementos indexicales muy ligados al contexto específico en que son proferidos), las vocalizaciones simbólicas no estaban condicionadas por una situación específica. Sus rasgos más relevantes eran el desplazamiento (la evocación sin evidencia sensorial) y la manipulación voluntaria. Aun cuando se tratara de un solo símbolo por enunciado, estas expresiones habrían sido ventajosas para la especie. Según Jackendoff, los fósiles de esta etapa son las exclamaciones «ay», «caramba», etcétera), que no se pueden integrar en construcciones sintácticas más grandes y que son empleadas no comunicativamente; algunas llamadas parecidas

a las de los primates («sh», «hey», «pst»), las cuales son usadas comunicativamente para llamar al interlocutor voluntariamente; o enunciados específicos para un contexto («hola», «chao», «sí», «no»).

La segunda etapa supone la manipulación de grandes repertorios de símbolos (palabras) aprendidos a gracias a la habilidad para imitar. Estos símbolos deben ser almacenados en la memoria de largo plazo y deben ser rápidamente recuperados. Del mismo modo, incluye la habilidad para crear nuevos símbolos a partir de un sistema generativo que permitiera la combinatoria de unidades. La generación de una clase abierta de símbolos solo podría haber sido posible, de acuerdo con el autor, como resultado del refinamiento y categorización de los sonidos utilizados para su creación. De esta manera, se pasa, sucesivamente, de la sílaba como unidad de almacenamiento a la fonología, lo que constituye un avance cognitivo importante. Un vocabulario abierto podría generarse con solo pocas distinciones entre sonidos del habla (por ejemplo, con diez fonemas, se pueden construir miles de palabras). El tracto vocal pudo adaptarse para generar un vocabulario más amplio y una articulación más rápida.

La tercera etapa corresponde a la concatenación de símbolos para formar enunciados más grandes. Esta etapa es compatible con la idea de Bickerton de que se puede ir más allá de una sola palabra sin tener que emplear la sintaxis moderna. En esta etapa, se trata de unir símbolos significativos para formar unidades más grandes también significativas pero sin organización interna. La interpretación sería aportada por la pragmática y sus limitaciones. Por ejemplo, *Pablo manzana* podría significar 'Pablo comió una manzana', 'Una manzana golpeó a Pablo' o 'la manzana es de Pablo', pero nunca 'Pabló besó a la chica que comió una manzana'. Los chimpancés entrenados alcanzan este nivel y solo este.

La cuarta etapa se encuentra cerca del protolenguaje de Bickerton. La innovación de esta etapa, el empleo del orden lineal para expresar relaciones semánticas, resulta adaptativo para la comunicación (siguiendo la propuesta de Pinker & Bloom, 1990) y se mantiene en las formas lingüísticas de los bilingües incipientes en una segunda lengua. En efecto, la denominada *variedad básica* presenta rasgos de ordenamiento, regulares y característicos, como el posicionamiento del agente del evento descrito a inicio de cláusula y del foco referencial al final. Un fósil contemporáneo de esta etapa corresponde a los agrupamientos que emplean el orden lineal para identificar qué elemento modifica a qué otro (como en los compuestos de dos nombres). De acuerdo con Jackendoff, el protolenguaje bickertoniano es muy parecido a la variedad básica, pues se caracterizaría por una competencia léxica, la ausencia de morfología flexiva y la ausencia de subordinación oracional; sin embargo, Bickerton no da evidencia de que haya orden lineal en el protolenguaje (tal vez porque incluyó a los pidgins, cuyo ordenamiento quedaría influido por la lengua meta).

La quinta etapa está marcada por la aparición de la estructura de frase. El agrupamiento de palabras en unidades más grandes y con núcleo permite pasar del orden lineal al orden jerárquico. Este logro permite mayor complejidad en mensajes: ahora se atribuye, por ejemplo, agente no a una palabra sino a una frase, un conjunto jerarquizado de palabras. De hecho, la estructura de frase permite que una frase pueda contener otras diversas.

Hasta este punto, el lenguaje ya habría desarrollado gran parte de la complejidad combinatoria y recursiva que hoy observamos en él. Las siguientes tres etapas que Jackendoff menciona (y que preceden inmediatamente al estado actual del lenguaje) se relacionan con precisiones de tipo léxico y sintáctico para expresar relaciones semánticas complejas.

La sexta etapa presenta un conjunto de términos específicos para conceptos de relación. Se trata de la aparición de nuevo vocabulario que se utiliza para expresar las relaciones semánticas de una manera explícita. Aparecen, pues, los elementos gramaticales como las preposiciones, las fórmulas de pregunta, los cuantificadores, los elementos ilativos de causas y efectos, etcétera. En esta etapa, Jackendoff incluye también la mejora de la actividad de pensamiento gracias al lenguaje, pues, a través de los enunciados lingüísticos, se examinan las actitudes proposicionales (como «suponer que p », «preguntarse si p », etcétera).

La séptima etapa va más allá de la estructura de frase y está caracterizada por la aparición de la flexión. Esta estrategia también ayuda a establecer, de manera explícita, las relaciones semánticas entre los componentes de un enunciado. Así, se establecen relaciones de concordancia entre el verbo con su sujeto o con su objeto para indicar relaciones temáticas; o también se puede marcar las frases argumentales con Caso con el mismo objetivo. Las relaciones formales quedan, entonces, visibles a través de la flexión, que bien podría ser irregular.

La octava etapa muestra ya una sintaxis avanzada. Jackendoff afirma que Bickerton trataría la etapa anterior y esta como un solo momento. La estructura de la frase puede marcar relaciones semánticas a través del orden de las frases en relación con el verbo, de manera relacionada con la flexión, pero independiente de ella. Así, cuando el sistema de flexión es muy rico, las lenguas permiten un orden más libre, sobre todo para marcar foco y tópico; sin embargo, también abundan las lenguas con estrategias redundantes de marcado de relaciones a distancia.

La novena y última etapa corresponde al estado actual del lenguaje. En efecto, el lenguaje moderno contiene las innovaciones sintácticas de las etapas previas, que se hacen explícitas en las lenguas del mundo de diferente manera y proporción. Jackendoff afirma que no existe ninguna prioridad temporal para las características de las etapas sexta, séptima y octava, por lo que podrían haber surgido al mismo tiempo.

Una vez aparecidas, fueron evolucionando de manera incremental, tal como los órganos complejos del cuerpo humano (como el ojo).

Como hemos visto, Jackendoff toma las ideas de Bickerton sobre el protolenguaje y las incluye en un panorama más amplio y mucho más elaborado. Sin embargo, a diferencia de este último, no intenta situar su propuesta dentro de periodos de tiempo determinado ni ofrece pistas sobre en qué momento de la evolución humana habrían tenido lugar las etapas que formula.

2.2.1.2. Un debate en torno a los componentes específicos del lenguaje y su evolución

Una de las contribuciones más importantes de Noam Chomsky al desarrollo de la ciencia cognitiva, entendida como el estudio de la mente, fue su cuestionamiento del supuesto central del conductismo⁹ de no invocar conceptos mentalistas (es decir, que presupongan la existencia de una mente o de estructuras, representaciones o procesos mentales) en la explicación de las capacidades humanas. En contra de dicho supuesto, Chomsky mostró¹⁰ que no era posible dar cuenta del aspecto más característico de la capacidad humana del lenguaje: la creatividad lingüística (es decir, la posibilidad del hablante para producir y comprender infinitos enunciados en su lengua materna), sin postular la existencia de una estructura mental sumamente compleja capaz de generar las estructuras sintácticas de todas las oraciones en su lengua. Pero una de las ideas más polémicas y difundidas de Chomsky (1957, 1965, 1981 y 1986) es que dicha estructura mental, o *gramática* de una lengua particular, no se adquiere a partir de mecanismos generales del aprendizaje, sino que es el producto de una facultad innata del lenguaje que, al entrar en contacto con los estímulos lingüísticos provenientes del entorno, construye en la mente una gramática correspondiente a la lengua particular a la que el niño se encuentra expuesto. Si la propuesta de Chomsky es correcta, entonces el lenguaje no sería, como tradicionalmente se ha asumido, una invención cultural como la agricultura, el derecho o la escritura, sino un rasgo biológico de nuestra especie, una idea que ha ido poco a poco ganando aceptación entre muchos especialistas¹¹.

⁹ Planteado, por ejemplo, por Skinner (1957).

¹⁰ En particular en su reseña de Skinner (1957) en Chomsky (1967 [1959]).

¹¹ Esta tesis de Chomsky ha sido difundida entre otros por Pinker (1994) en su libro *El instinto del lenguaje*. En nuestro idioma es defendida por ejemplo por De Vega y Cuetos (1999): «estos hechos [...] sugieren que el lenguaje es un rasgo biológico específico de nuestra especie. Es más un ‘instinto’ (como la eco-localización del delfín, la habilidad tejedora de la araña o la visión estereoscópica de los primates) que una invención cultural (como la rueda o la escritura)». (1999, p. 14). Es importante señalar, sin embargo, que no hay un consenso completo al respecto. Así, para investigadores de la cognición como Tomasello (1999 y 2001), el lenguaje es parte del conjunto de habilidades cognitivas que constituyen la cultura humana y que son conservadas, modificadas y transmitidas de generación en generación.

Sin embargo, un nuevo debate ha surgido en el campo de los que consideran, siguiendo a Chomsky, que el lenguaje es un rasgo biológico de la especie humana; pues si es un rasgo biológico, tiene que ser un producto de la evolución de nuestra especie. La discrepancia concierne a la manera en la cual dicho rasgo habría aparecido en el marco de la evolución. Pues como veremos, el propio Chomsky, así como otros investigadores como Piattelli-Palmarini, Hauser y Fitch, consideran que el núcleo de la facultad del lenguaje no habría evolucionado mediante el mecanismo convencional de la selección natural, que selecciona aquellos rasgos que contribuyen a la supervivencia de los organismos en un determinado ambiente, sino que habría evolucionado por otro tipo de mecanismos, como la exaptación (en el que un rasgo evolucionado para otro propósito asume una nueva función). De ser así, el lenguaje (en su aspecto más nuclear) no cumpliría ninguna función adaptativa por la cual habría sido seleccionado, y en particular no estaría diseñado para la comunicación. Frente a esta posición un grupo de investigadores, inicialmente seguidores de Chomsky, como Pinker, Bloom y Jackendoff, han tomado distancia en relación a este punto de vista de su maestro, y afirman que siendo la facultad del lenguaje un rasgo biológico de la especie, sí tiene que haber evolucionado mediante el mecanismo de la selección natural y por tanto sí debe cumplir una función adaptativa, que sería la comunicación de proposiciones por un canal serial. A continuación presentaremos en mayor detalle este debate, que es fundamental para el tema de nuestra investigación.

Una propuesta importante sobre el origen y la evolución del lenguaje desde las ciencias cognitivas modernas es la que proponen el psicólogo evolucionista Steven Pinker y su estudiante doctoral Paul Bloom en un artículo publicado en *Behavioral and Brain Sciences* en 1990. El artículo puede considerarse una respuesta frontal a un texto publicado por Massimo Piattelli-Palmarini un año antes, en el que el último esboza una serie de razones por las que la facultad del lenguaje no debería ser considerada un producto del diseño gradual de la selección natural, sino una capacidad resultante del ejercicio de otros mecanismos de la evolución en nuestra especie. En contraposición, Pinker y Bloom (1990) defienden en su texto un escenario en el que el origen y desarrollo del lenguaje se explican por presiones selectivas específicas en nuestra especie que habrían posibilitado su adaptación progresiva al entorno. Se podría considerar que ambos textos dan origen a un debate todavía vigente sobre qué aspectos del lenguaje habrían evolucionado específicamente para su funcionamiento y qué requisitos tendrían que haber estado presentes para hacerlo.

Piattelli-Palmarini (1989) suscribe en gran parte las ideas de Chomsky (1957, 1965, 1981 y 1986) con respecto a la naturaleza de la facultad del lenguaje; como, por ejemplo, la existencia de un dispositivo innato de complejas características estructurales y específicas para la adquisición de una lengua en el desarrollo del niño.

Es sobre la base de esta concepción del lenguaje y de las recientes investigaciones en microbiología que Piattelli-Palmarini pretende mostrar que la evolución del lenguaje parece no responder a un moldeamiento paulatino de la facultad del lenguaje a cargo de la selección natural. La retórica del autor parte de la siguiente pregunta: ¿cómo podemos explicar que características tan específicas —como los principios de proyección del sujeto, subyacencia, ligamiento, entre otros— hayan sido moldeadas por la selección natural o hayan cumplido algún tipo de función adaptativa?

A partir de una analogía con el desarrollo inmunológico en nuestra especie, Piattelli-Palmarini inicia el texto presentando un debate sobre las diferencias entre procesos de aprendizaje mediante la instrucción y «aprendizaje» (1989; comillas en original) conducido por un dispositivo interno, que el autor denomina un proceso de *adquisición*. El cuerpo humano viene preparado para lidiar con innumerables microorganismos en el ambiente; sin embargo, el desarrollo inmunológico de un tipo de defensas biológicas y no otras dependerá del tipo de microorganismo con el que el individuo se encuentre a lo largo de su desarrollo.

De manera similar, el autor pretende defender un nuevo modelo de cognición, de manera general, y de la adquisición del lenguaje, de manera particular, que no sea *únicamente* dependiente del ambiente. Una posición equilibrada comparada con las propuestas empiristas radicales para el desarrollo del lenguaje de mediados del siglo XX (por ejemplo, Skinner, 1957). Para el caso del lenguaje —y es en este aspecto que las ideas chomskianas se hacen evidentes— no es descabellado proponer que el niño, como en el caso de los anticuerpos que forman parte de su dotación genética, nazca con un mecanismo especializado para la adquisición de una lengua; sobre todo, si consideramos que el proceso concluye de manera excelente pese a la escasa y fragmentada experiencia que tiene de ella durante sus primeros años de vida (véase §3.1.). El mecanismo de adquisición es entendido entonces como una serie de principios innatos, comunes a todas las lenguas naturales, y un conjunto de parámetros que deben ser activados o desechados de acuerdo con los rasgos específicos de la lengua que adquiere el hablante. Es en este sentido que no hay nada que se *aprenda* —en el sentido instructivo de la palabra— o que, de alguna manera, sea transferido del exterior al organismo. Es el organismo el que responde de manera determinada ante ciertos estímulos del exterior.

La complejidad y especificidad lingüística del mecanismo innato que se propone lleva a Piattelli-Palmarini a concluir la imposibilidad de siquiera conjeturar una presión selectiva tal que haya permitido la aparición de los componentes formales computacionales específicos del lenguaje. Si el lenguaje hubiera sido diseñado para la comunicación, contaríamos con una facultad de características infinitamente más sencillas (1989). A partir de esto, en las conclusiones del texto, el autor se aventura

a proponer, más bien, un origen exaptativo del lenguaje, es decir, considera su origen como una suerte de *spandrel* biológico en la especie (Gould & Lewontin, 1984[1979]).

El concepto de *exaptación* describe un proceso por el cual un rasgo determinado en la evolución de un organismo pasa a cumplir una función adaptativa que antes no tenía. Un ejemplo de este fenómeno, que proponen Gould y Lewontin en el artículo de 1979 y que Piattelli-Palmarini (1989) recuenta, consiste en la aparición de las alas de los insectos, las que habrían cumplido la función de servir como enfriadores térmicos inicialmente para luego convertirse en estructuras que permiten el vuelo a través de un proceso de exaptación. En un escenario imaginario, la utilización para el vuelo de los enfriadores térmicos originales habría sido cada vez más adaptativo para la especie, lo que habría resultado en el establecimiento de una nueva función del rasgo y en su especialización consecuente con el devenir del tiempo.

Debemos hacer aquí una diferencia entre los tipos de exaptación que pueden darse (Gould & Vrba, 1982). En primer lugar, un tipo de proceso exaptativo es aquel en el que un rasgo pasa de cumplir una función *x* a cumplir una función *y* en la especie (como en el caso de las alas de los insectos en el ejemplo del párrafo anterior). Sin embargo, también es posible que el rasgo no haya cumplido una función previamente, sino sea una suerte de *producto colateral* o corolario de la evolución de algún otro rasgo al que se encuentre relacionado. Pensemos, por ejemplo, en el color rojo de la sangre, en la formación en V de los dedos de la mano, o que su número sea impar en cada extremidad de nuestro cuerpo. No podemos afirmar con certeza que existieran presiones selectivas para la aparición de estos rasgos a partir de la manera en la que se encuentran configurados actualmente. En un afán por precisión terminológica, resaltaremos que este último tipo de proceso exaptativo —aquel que constituye el paso de un rasgo carente de una función original rastreable a cumplir una función adaptativa en la especie— es comúnmente denominado un *spandrel biológico*, siguiendo la metáfora arquitectónica de las albanegas de la Basílica veneciana de San Marco propuesta en el texto de Gould y Lewontin (1984 [1979]) mencionado líneas arriba.

Sostener que el lenguaje se originó por exaptación libera a la teoría de problemas tales como explicar una posible presión selectiva inicial o esbozar estados intermedios de desarrollo previos a su aparición en la especie. Es por esta razón que el texto de Piattelli-Palmarini no presenta un panorama del posible origen del lenguaje en la especie, se plantea escenarios hipotéticos, contrasta condiciones necesarias tanto biológicas como presentes en el entorno que pudieran haber dado lugar a presiones selectivas específicas para su aparición, como lo reseñamos para Bickerton (1990, 2007 y 2010) líneas arriba.

En 1990, Steven Pinker y Paul Bloom presentaron un artículo en el que confrontaban la hipótesis en torno al origen exaptativo del lenguaje, esbozada por Piattelli-Palmarini un año antes, y en el que más bien defendían que su aparición sería el resultado del moldeamiento progresivo de la selección natural a lo largo del tiempo.

El artículo constituye, en primer lugar, en una crítica a la conclusión de Piattelli-Palmarini (1989) de que el lenguaje no puede ser resultado del ejercicio acompasado de la selección natural por constituir una facultad de características complejas y de máxima especificidad operativa. Los autores señalan que existen evidencias de otros rasgos en los organismos cuya especificidad estructural y funcional actual ha sido moldeada por la selección natural (por ejemplo, el ojo de los vertebrados).

Recordemos que la propuesta de Piattelli-Palmarini (1989) se sostiene en la imposibilidad de conjeturar algún tipo de presión selectiva que dé cuenta de la complejidad de los principios específicos de la facultad del lenguaje. Por otro lado, Pinker y Bloom arguyen que justamente esa complejidad en el *diseño* del lenguaje es evidencia de la sofisticación característica de un rasgo modelado por la selección natural con el propósito de cumplir una función adaptativa en la especie; en este caso, la comunicación de proposiciones complejas mediante un canal serial (1990, p. 714).

Ahora bien, la ‘complejidad’ a la que Piattelli-Palmarini, por un lado, y Pinker y Bloom, por el otro, se refieren parece estar describiendo diferentes aspectos o concepciones del lenguaje, como se verá de manera más clara años después con la publicación de un artículo de Hauser, Chomsky y Fitch —que elabora las intuiciones sobre el origen exaptativo del lenguaje— y una respuesta a este por el propio Pinker y Ray Jackendoff tres años después.

En la propuesta de Piattelli-Palmarini, la complejidad que presenta el lenguaje se refiere, básicamente, a principios sintácticos y formales imposibles de adquirir durante el desarrollo ontogenético mediante repetición, instrucción o analogía (ver §3.2.1.) por lo que se proponen características particulares del mecanismo innato de la facultad del lenguaje (ver §1.2.). Por otra parte, Pinker y Bloom parecen entender la complejidad del lenguaje en un sentido más amplio de lo que lo hace Piattelli-Palmarini. La complejidad del mecanismo para estos autores es comparada con la del ojo de los vertebrados (1990, p. 710), cuya especificidad de diseño —en otras palabras, la precisión estructural de la constitución de sus partes— habría sido moldeada de manera conjunta para cumplir una función determinada. De manera similar se argumenta para el lenguaje y la pregunta se dirige así hacia las condiciones cognitivas y estructurales que tendrían que haber estado presentes para su evolución.

Pinker y Bloom hacen explícita la noción de *complejidad adaptativa* que mencionan en el texto. Señalan así que esta «describe cualquier sistema compuesto de muchas partes que interactúan entre sí y donde los detalles de la estructura y composición

de las mismas sugieren un diseño específico para cumplir una determinada función» (1990, p. 709). De aquí que desestimar la función en el afán por explicar la forma de la estructura sea una empresa destinada al fracaso. Según los autores, considerar que la aparición del lenguaje se haya dado por razones diferentes a la selección natural —es decir, que un mecanismo así de complejo haya aparecido como producto colateral de la selección de algún otro rasgo o como la cooptación de uno que cumpliría una función diferente— no ofrece ninguna posibilidad de explicación sobre la función que el lenguaje habría cumplido en el desarrollo de la especie. La agudeza del argumento de Pinker y Bloom se observa cuando se evalúa, desde esta óptica, la analogía de Gould y Lewontin en relación con las albanegas de la basílica veneciana presentada por Piattelli-Palmarini: si bien la acomodación de un domo encima de cuatro arcos nos ofrece un *spandrel* como resultado, esto no explica el detalle y la perfección estilística del fresco pintado en él (1990, p. 710). El minucioso artista, se concluye, habría sido la selección natural.

De acuerdo con los autores, la comunicación de contenido proposicional se hace relevante al considerar las ventajas que esta función del lenguaje podría haberle otorgado a nuestros antepasados. Pensemos tan solo en la posibilidad de obtener información de segunda mano sobre peligros potenciales o hacia qué dirección caminar para encontrar el repositorio de agua más cercano. Pinker y Bloom defienden que poder comunicar este tipo de contenido habría sido una presión selectiva para la evolución del lenguaje en el mecanismo que conocemos ahora. A diferencia de los sistemas de comunicación animal —cuyo contenido significativo no puede ser desligado de la situación en la que se utilizan—, los autores precisan que la presión selectiva para la aparición del lenguaje habría sido la comunicación de proposiciones complejas mediante un canal de articulación sería capaz de servir para el intercambio de información muy fina.

La necesidad comunicativa de nuestros ancestros homínidos no se reducía únicamente a la señalización de comida y peligros en el entorno —es más, se sugiere que concebir un escenario de este tipo constituye una caricaturización de las condiciones agrestes de la sabana africana—; por otro lado, la expresión de proposiciones complejas sobre estados internos del individuo necesarios para la cooperación en el grupo habría sido un posible presión para el ejercicio de la selección natural. Nuestros ancestros fueron criaturas de características sociales suficientemente complejas, para quienes un mecanismo diseñado para expresar información precisa acerca del tiempo, el espacio y las relaciones entre sujetos en sociedad (‘quién le hace qué a quién’) habría sido sumamente adaptativo (1990, p. 724). Limitar nuestra visión de la comunicación de los primeros homínidos a la señalización de peligro o el hallazgo de comida mediante gritos y ademanes desordenados es una conjetura simplista

de las necesidades expresivas que sin duda tenían. Como se ejemplifica en el texto, contar con un mecanismo con precisión suficiente para distinguir si la región a la que nos dirigimos tiene animales que *pueden comerse* o que *pueden comernos* o si la fruta que ahí encontraremos *estuvo* madura, *está* madura o *estará* madura, habría traído, sin duda, diversas ventajas adaptativas.

Hauser, Chomsky y Fitch (2002) extienden las aseveraciones de la propuesta de Piattelli-Palmarini (1989) sobre el origen exaptativo del lenguaje y las fundamentan en una nueva concepción de los rasgos característicos del lenguaje, elaborada en los salones del MIT a lo largo de la década previa a su publicación dentro del denominado Programa Minimista. De acuerdo con Hauser, Chomsky y Fitch, la facultad del lenguaje puede ser dividida en dos componentes: la *facultad del lenguaje en sentido amplio* (FLB, por sus siglas en inglés) y la *facultad del lenguaje en sentido estrecho* (FLN). Grosso modo, la primera incluiría la segunda y *por lo menos* dos sistemas cognitivos adicionales: el sistema Articulatorio-Perceptual (A-P) y el sistema Conceptual-Intencional (C-I) (2002, p. 1570).

La FLB estaría compuesta entonces por todas las capacidades, exclusivamente humanas o compartidas con otros vertebrados, que nos permiten adquirir y ser expertos en una lengua natural sin instrucción explícita. Se deja aquí de lado otro tipo de capacidades, exclusivamente humanas o compartidas, que si bien necesarias para el lenguaje, no son suficientes para su desarrollo (por ejemplo, memoria, circulación, respiración, etcétera) (2002, p. 1571). Esta es una primera delimitación entre las capacidades humanas que sirven directamente para la implementación de la *competencia lingüística*, y otras capacidades y componentes cognitivos que parame-tran y limitan la *actuación lingüística*.

La FLN forma parte de la FLB y se concibe como un sistema lingüístico computacional aislado e independiente de otros sistemas con los que interactúa y en los que se proyecta, llamados *interfaces*, y que podemos reconocer relacionados a los dos sistemas cognitivos de los que se hace mención explícita en la definición de la FLB: el sistema A-P y el C-I. Se sostiene, además, que un componente esencial de la FLN es un algoritmo computacional, o sintaxis estrecha (*narrow syntax*), que generaría representaciones internas de características recursivas y sería capaz de proyectarlas en las interfaces mencionadas a través de los sistemas fonológico y semántico, respectivamente. Cada expresión es, en este sentido, un emparejamiento de sonido y significado (2002, p. 1571).

A este respecto, se sostiene que esta capacidad recursiva de la sintaxis estrecha sería una particularidad también hallada en los números naturales. Como se defenderá más adelante en el texto de los autores, el aprendizaje de los números por los niños humanos difiere de su aprendizaje por chimpancés entrenados en tanto

los niños parecen hacerlo mediante una capacidad computacional recursiva. En este sentido, por lo menos, el lenguaje sería análogo a los números naturales. Con esta afirmación es que comienza a revelarse una diferencia sustancial entre las capacidades humanas y las de otros animales y se marca una distinción tajante entre las capacidades cognitivas de nuestra especie y las de otros animales, incluso primates entrenados con los que compartimos poco más del 95% del ordenamiento de la secuencia de nuestro ADN.

A propósito de la inclusión de este dato, debemos señalar que el texto de Hauser, Chomsky y Fitch es un rico material de datos y evidencia empírica desde diferentes perspectivas de análisis. Especialmente, grande es el interés en demostrar, sobre la base de la biología comparativa, la particularidad de algunas capacidades en nuestra especie al punto que esto constituye las líneas de partida y la metodología de la investigación que se proponen: «[...] una tarea ineludible, básica y lógica de la investigación comparativa sobre la evolución del lenguaje es esta, simple y de inferencia negativa: un rasgo presente en un animal no humano no evolucionó específicamente para el lenguaje humano, pese a que pueda ser parte de la facultad del lenguaje y jugar un papel íntimo en el procesamiento lingüístico» (2002, p. 1572). De esta manera, cualquier rasgo que se encuentre en nuestra especie y que sea compartido con otras será descartado de antemano y no será tomado en cuenta para la explicación de la capacidad humana para el procesamiento lingüístico.

La razón para la separación de la facultad del lenguaje en componentes internos se hace explícita en la página 1571 del texto, cuando los autores señalan que se puede investigar ricamente la evolución de la FLN sin tener que preguntarnos por las historias evolutivas particulares de los demás aspectos que conforman FLB o que se encuentran fuera de ella y que limitan la implementación de la FLN. Si consideramos que el aspecto recursivo esencial de la FLN parece ser una cualidad que no presenta analogía con otra especie del mundo, debemos preguntarnos por las particularidades de su evolución en nuestra especie.

El examen de los componentes de la FLB: el sistema A-P, el C-I y la FLN, conducen al desenlace final de que solo el tercero —el mecanismo computacional para la combinación recursiva de piezas léxicas— constituiría una característica particular del lenguaje humano. Tanto las características del sistema A-P, como la percepción categorial y la laringe descendida, por ejemplo, así como las características del sistema C-I parecerían ser compartidas con otros animales. El mecanismo computacional sería lo único exclusivamente humano y exclusivamente lingüístico en la facultad del lenguaje, por lo que es por este componente que debemos preguntarnos en nuestra pesquisa por la aparición del lenguaje en la especie.

Hauser, Chomsky y Fitch sostienen que parecen existir pocas razones para señalar que FLN sea un rasgo adaptado por la selección natural. Esto, principalmente por dos razones. Primero, pues como ya se señaló, parece no haber evidencia contundente que demuestre que otras especies poseen algo similar a recursividad observada en la sintaxis estrecha; y, segundo, ya que no parecer ser una opción plausible dividir la FLN en rasgos más específicos y trazar, para cada uno, una historia evolutiva independiente (2002, p. 1574).

Es por esta razón que, de acuerdo con las líneas generales de la argumentación de Piattelli-Palmarini en 1989, se propone que la FLN es un *spandrel* o un producto colateral de rasgos preexistentes en la especie. Sin embargo, se señala que esta afirmación es perfectamente compatible con la aceptación de la teoría adaptacionista (2002, p. 1574). La pregunta no es si la FLN es, en general, un rasgo adaptativo, que sin duda lo es, pues sirve para la comunicación; sino, si componentes particulares del funcionamiento de la FLN son adaptaciones lingüísticas. Dada la complejidad y especificidad de los componentes del mecanismo se concluye que este no podría haber sido modelado por la selección natural sino, más bien, sería el resultado de la exaptación de algún otro rasgo en la especie.

Steven Pinker y Ray Jackendoff responden a la propuesta de Hauser, Chomsky y Fitch, presentada arriba, en un artículo publicado en *Cognition* en 2005. En él, señalan que la hipótesis de los autores entraña dos faltas graves: primero, una reducción del lenguaje a un mecanismo recursivo y una simplificación del componente sintáctico del lenguaje a un mecanismo recursivo.

La respuesta constituye una defensa de las hipótesis de Pinker y Bloom (1990) sobre el diseño del lenguaje producto del modelamiento gradual de la selección natural, esta vez, con evidencia tanto lingüística como genética y de imagería cerebral.

La primera parte del texto está dedicada a analizar, uno por uno, aspectos gramaticales que, de acuerdo con los autores, también deberían formar parte central de la facultad del lenguaje en tanto se reconocen como específicamente lingüísticos y exclusivamente humanos (Pinker & Jackendoff, 2005, p. 217). De acuerdo con los autores, la razón fundamental para la afirmación de Hauser, Chomsky y Fitch de que «el único componente exclusivamente lingüístico y exclusivamente humano de la facultad del lenguaje es la FLN» (2002, p. 1573) es resultado de la nueva aproximación de Chomsky a la facultad del lenguaje: el Programa Minimista, desarrollado en la década de 1990, y que propone, en su versión más estricta, que el único mecanismo necesario para el emparejamiento de sonido y significado sería una operación recursiva denominada *Merge* (Pinker & Jackendoff, 2005, pp. 218-219).

La segunda parte del texto presenta las implicancias de la afirmación de esta afirmación de Hauser, Chomsky y Fitch, en tanto de ella se desprenden las tres hipótesis

siguientes: (1) que el lenguaje está mal diseñado para la comunicación, (2) que el lenguaje es *perfecto* en tanto está rigurosamente especializado, y que (3) la única manera en la que el lenguaje podría haber surgido en la especie es por exaptación de alguna otra habilidad recursiva como la numeración, la ubicación espacial o la atribución psicológica producto de la complejización social. En esta parte del texto, Pinker y Jackendoff se encargan de falsear cada una de estas hipótesis a partir de la evidencia presentada en la primera parte.

A diferencia de lo que ocurre con Hauser, Chomsky y Fitch (2002), la propuesta de Pinker y Jackendoff partirá de una premisa de base bastante diferente: aquella que sostiene que la especificidad y complejidad del diseño de un rasgo es indicio del modelamiento progresivo de la selección natural en su evolución. En palabras de los autores: «[s]i la facultad del lenguaje tiene muchos rasgos que son particulares al lenguaje mismo, esto sugeriría que la facultad fue blanco de la selección natural. Pero, si representa una pequeña extensión de capacidades que existieron en el linaje primate ancestral, podría ser el resultado de una mutación azarosa que se fijó en la especie a través de la deriva genética o mediante otro mecanismo evolutivo no adaptativo» (2005, p. 202).

Lo que se propone con la cita es que si encontramos en el lenguaje rasgos que no podemos afirmar que sean compartidos con otras especies o que provengan de otras capacidades cognitivas humanas, entonces debemos inferir que estos rasgos de la facultad del lenguaje fueron modelados por la selección natural para cumplir una función determinada en él, como ya se advertía de manera similar en el texto de Pinker y Bloom de 1990. Lo que los autores pretenderán seguidamente será, entonces, mostrar evidencia empírica de que existen suficientes rasgos que son exclusivamente humanos y excepcionalmente lingüísticos en la facultad del lenguaje, y que no solo la recursividad sintáctica, como proponen Hauser, Chomsky y Fitch (2002), sería central en el lenguaje.

Componentes lingüísticos como la fonología —que si bien jerárquica, no es recursiva—; aspectos sintácticos como el marcado de caso, la dependencia a distancia, la concordancia entre elementos dentro de una misma frase, entre otras particularidades no recursivas del componente sintáctico son tomados como evidencia empírica de la investigación lingüística para señalar su importancia y centralidad en el diseño del lenguaje pese a no ser componentes recursivos. Con esto, se pretende sostener el complejo diseño del lenguaje para cumplir una función adaptativa. Como ya se aventuraba en el texto de 1990, esta sería la comunicación de proposiciones complejas a través de construcciones seriales.

A partir de la premisa de trabajo de la teoría evolutiva y la biología comparativa que sostiene que «[l]a pregunta clave para caracterizar una función biológica

no es para qué un rasgo es típicamente usado sino para qué fue diseñado» (Pinker & Jackendoff, 2005, p. 224), los autores ofrecen un ejemplo a este respecto. Pese a que las manos sean usadas, por ejemplo, con mayor frecuencia para tamborilear una carpeta que para asir un objeto, nadie pensaría, sugieren, que su diseño responda al tamborileo y afirmarían, en consecuencia, que esta es la razón por la que habrían sido seleccionadas en la especie. Si consideramos que el lenguaje es, como Chomsky señala, un sistema para la representación del pensamiento y no un sistema de comunicación, ¿cuál sería la razón, tan solo para poner un ejemplo, para la precisión de las reglas fonéticas y fonológicas encontradas en las lenguas sino generar una señal claramente perceptible que beneficie al oyente en el intercambio de información?

Todo conduce a la afirmación por parte de los autores, así como hicieran Pinker y Bloom un década antes, de que el lenguaje tendría que haber evolucionado a través de la selección natural y habría constituido una adaptación ventajosa para la especie en tanto permite la comunicación de proposiciones complejas. Esta aseveración cuestiona la idea central de Hauser, Chomsky y Fitch en el texto de 2002 de acuerdo con la cual la característica central del lenguaje, la FLN, sería una exaptación de algún otro rasgo recursivo previo en la constitución de la mente humana. Recordemos la cita en la que los autores señalan esto: «[d]urante la evolución, el sistema recursivo modular y específico en su dominio podría haberse vuelto penetrable y de dominio general. Esto habría abierto la posibilidad para los humanos, quizás excepcionalmente, para aplicar el poder de la recursividad a otros problemas» (p. 1578).

En relación con este punto, Pinker y Jackendoff sostienen que el surgimiento exaptativo de un rasgo de la facultad del lenguaje no constituye, necesariamente, una razón concluyente para la exclusión de un escenario adaptacionista posterior. De acuerdo con los autores, la hipótesis sobre la exaptación del lenguaje en Hauser, Chomsky y Fitch (2002) dejaría de lado la posibilidad de que un rasgo, en principio seleccionado para una función, pueda haber sido nuevamente *seleccionado* para cumplir otra función y ser modelado, posterior y consecuentemente, por la selección natural (Pinker & Jackendoff, 2005, p. 229) —un proceso común en la evolución de las especies. Un ejemplo de este caso de exaptación dirigida sería la de los antebrazos de los animales vertebrados, por ejemplo, en principio seleccionados para la estabilidad en el agua y, posteriormente, para el vuelo en las aves.

En una réplica al artículo de Pinker y Jackendoff (2005), publicada el mismo año en *Cognition*, Fitch, Hauser y Chomsky (2005) les reprochan no haber comprendido adecuadamente su propuesta y haber hecho borrosa la distinción central entre FLB y FLN (2005, p. 180). Estos autores conceden así que FLB puede ser una adaptación que evolucionó como producto de la selección natural para diferentes fines, entre ellos la comunicación, y que en su artículo del año 2002 han sostenido claramente

que sí consideran que el lenguaje es el resultado de la evolución y que muestra signos de diseño adaptativo (p. 185). Pero insisten en que a diferencia de FLB, FLN no es una adaptación (en todo caso para la comunicación) (p. 189). Observan igualmente que el hecho de que FLN esté compuesta por un único mecanismo simple (la recursión) nulifica el argumento del diseño en el que se basan los adaptacionistas (p. 190). Presentan además evidencia empírica contraria a la tesis central de Pinker y Jackendoff (2005) de que hay otros mecanismos en FLN distintos de la recursión, y consideran que dichos aspectos del lenguaje forman parte de FLB, al no ser exclusivamente lingüísticos y exclusivamente humanos.

En su contra-réplica, también publicada ese mismo año en *Cognition*, Jackendoff y Pinker (2005) sostienen que la caracterización de FLN de Fitch, Hauser y Chomsky (2005) es problemática, pues presupone una dicotomización de las capacidades cognitivas en las que son absolutamente únicas al lenguaje y a los seres humanos y aquellas que son idénticas a capacidades cognitivas no lingüísticas o no humanas, dejando de lado aquellas capacidades cognitivas que han sido modificadas durante la evolución humana. Señalan que Fitch, Hauser y Chomsky (2005) suelen aplicar la distinción entre FLN/FLB de manera absoluta, y de ese modo utilizan cualquier similitud entre una función lingüística y cualquier otra capacidad cognitiva, desterrar a dicha función de FLN y situarla en FLB (Jackendoff & Pinker, 2005, p. 214).

En un volumen reciente sobre la evolución del lenguaje (Larson, Déprez & Yamakido, 2010) Chomsky presenta una formulación más precisa de su tesis sobre la evolución del lenguaje. Una diferencia importante con los artículos publicados con Hauser y Fitch en 2002 y en 2005, en que habla de modo general de 'la recursión' como el componente central y único de FLN, es que aquí precisa que la recursión es producto del mecanismo específico denominado *Merge*, que toma estructuras preexistentes X e Y, y las combina en una nueva estructura Z, la que aplicada de manera ilimitada a un lexicón de átomos léxico-conceptuales produciría una infinitud discreta de expresiones estructuradas. *Merge* sería posiblemente la única operación recursiva dentro de la Gramática Universal, parte del componente genético de la facultad del lenguaje, y por tanto un producto de la evolución de este órgano cognitivo (Chomsky, 2010, p. 52). En este mismo artículo Chomsky plantea una asimetría en la relación entre el lenguaje-I y los dos sistemas de interface (conceptual-intencional y sensorio-motor), sosteniendo que su relación primaria es con la interface semántica, es decir, con los sistemas del pensamiento, antes que con la interface sensorio-motriz que hace posible la externalización y la comunicación (2010, p. 55). Chomsky imagina así que un cambio aleatorio en las redes neuronales del cerebro habría dado origen, en uno de nuestros ancestros, a la operación de *Merge* ilimitado, la que se habría aplicado a conceptos con propiedades complejas

y poco entendidas. El lenguaje así surgido habría proporcionado a este ancestro un rango infinito de expresiones estructuradas, dándole muchas ventajas como la capacidad para el pensamiento complejo, el planeamiento, la interpretación, etcétera las que habrían sido transmitidas a su descendencia, la que predominó. Recién en un momento ulterior habría sido ventajosa, según Chomsky, la externalización de esas expresiones, de modo que la capacidad del lenguaje se habría conectado solo secundariamente con el sistema sensorio-motor para la externalización y por tanto para la comunicación (p. 59).

En un trabajo publicado en el mismo volumen, Jackendoff (2010) contrasta por su lado dos arquitecturas posibles acerca del lenguaje y de la relación entre sintaxis y semántica, que en su opinión llevan a dos maneras muy distintas de concebir la evolución del lenguaje, que son por un lado una arquitectura centrada en la sintaxis, asumida por Chomsky y la corriente principal de la gramática generativa, y por otro lado una arquitectura paralela, como la que él mismo ha propuesto (Jackendoff, 2002). En una arquitectura centrada en la sintaxis, la capacidad generativa del lenguaje es localizada en el componente sintáctico, el que media necesariamente las correspondencias entre fonología y semántica. En una arquitectura paralela existen en cambio principios combinatorios independientes en la fonología, en la sintaxis y en la semántica, y las estructuras de los tres componentes se vinculan mediante reglas de interface (2010, p. 67).

De acuerdo con Jackendoff, el enfoque centrado en la sintaxis implica que la estructura combinatoria del pensamiento humano es una consecuencia de la capacidad combinatoria proporcionada por la sintaxis. Desde el punto de vista evolutivo, es difícil imaginar algún tipo de adaptación previa sobre cuya base hubiese surgido una capacidad del lenguaje de este tipo. El sistema generativo sintáctico en su totalidad, así como las proyecciones a la fonética y a la semántica, tendrían que haber aparecido prácticamente de la nada. Es esta situación teórica la que motivaría a aquellos investigadores que procuran reducir la evolución de la capacidad del lenguaje a un paso único, como Hauser, Chomsky y Fitch (2002). En cambio una estructura paralela admite que el pensamiento pudiese haber estado estructurado combinatoriamente antes de la sintaxis. Un protolenguaje en el sentido de Bickerton (1990) permitiría así expresar en forma cruda relaciones semánticas, por ejemplo mediante el orden de las palabras. Dentro de esta concepción (también presentada en Jackendoff, 1999), la estructura sintáctica representaría la innovación más elevada dentro de este proceso evolutivo, al posibilitar una forma más reglamentada de expresar convencionalmente relaciones semánticas entre las palabras, si bien tanto las relaciones semánticas como las palabras ya estarían presentes en el sistema anterior (2010, p. 70).

2.2. Las principales propuestas sobre las ventajas adaptativas del lenguaje para la comunicación y la cooperación

El lenguaje presenta una serie de ventajas adaptativas relacionadas a la comunicación en la especie. Dentro de las ventajas adaptativas más relevantes se encuentran la organización grupal de la especie y el mantenimiento de la cohesión social en ella. Autores provenientes de diferentes disciplinas de investigación han escrito sobre estas ventajas del lenguaje; de esta manera, su enfoque ha estado en la función de esta facultad única en nuestra especie y no, directamente, en el origen y la evolución de los diferentes mecanismos que la constituyen.

2.2.1. La propuesta de Michael Tomasello

El núcleo de la propuesta de Tomasello consiste en que la comunicación cooperativa humana resultó adaptativa porque surgió en el contexto de actividades colaborativas mutuas, en el marco de las cuales los individuos que ayudaban a otro se estaban ayudando, a la vez, a sí mismos. En un inicio, las habilidades de comunicación cooperativa surgieron y se usaron en actividades totalmente colaborativas —estructuradas a partir de la existencia de metas y atención conjuntas, compartidas entre los participantes, las cuales los proveerían de un campo conceptual común (*common conceptual ground*).

El autor presenta dos razones para justificar la necesidad de estudiar mecanismos no lingüísticos previos al lenguaje, en comparación con los hallazgos en la primatología, como camino hacia comprender con profundidad la naturaleza de la comunicación humana. La primera razón consiste en que todo uso comunicativo del lenguaje depende, necesariamente, de mecanismos extralingüísticos; la segunda, en que no se puede explicar el origen del lenguaje partiendo de elementos que lo presuponen, pues esto implica, evidentemente, un vicio lógico (2008, pp. 57-58).

Entre los dispositivos que parecemos compartir con nuestro parientes primates, se destacan el señalamiento y la pantomima (p. 60). Con respecto al primero, tanto en nosotros como en los grandes simios, la función básica del gesto sirve para llamar la atención de otros respecto de algo que se encuentra en nuestro rango perceptual. La diferencia consiste en que nuestra capacidad para realizar exitosamente el acto comunicativo depende de nuestra habilidad para seguir la mirada del otro, y en los simios ello depende de su capacidad de reconocer ruidos y determinado contacto. En el caso de la pantomima, los primates y nosotros buscamos llamar la atención sobre algo que está fuera de nuestro rango perceptual, para lo que utilizamos capacidades diferentes. En el caso de los simios, esto depende de su capacidad para anticipar los movimientos de los demás; en el nuestro, lo hacemos sobre la base de la comprensión de actos intencionales (p. 62).

Tanto las actividades colaborativas, en un sentido general, como la comunicación cooperativa, en un sentido particular, descansan, según Tomasello, en la misma infraestructura subyacente de metas y atención conjunta recursivamente estructuradas, así como en otras manifestaciones de intencionalidad compartida. La habilidad cognitiva básica de la intencionalidad compartida (esto es, lectura recursiva de mentes) habría surgido como una adaptación para la actividad colaborativa de manera específica (dada una adaptación inicial para la tolerancia que habría hecho a los individuos más generosos con la comida, con lo que habrían llegado a ser capaces, por ejemplo, de compartir con otros el botín de sus cacerías). Esto último, habría permitido el establecimiento de marcos de atención conjunta y campos conceptuales compartidos. La combinación de la voluntad de ayudar y de la lectura recursiva de mentes habría dado lugar al surgimiento de expectativas mutuas de colaboración y, eventualmente, a la intención comunicativa griceana como guía de las inferencias de relevancia.

Si bien sobre el origen de la comunicación cooperativa el autor confiesa no tener una historia evolutiva detallada que narrar, sí propone ciertos pasos lógicos que habrían seguido los componentes de la comunicación humana en su curso evolutivo. Así, propone tres elementos básicos por los cuales habría evolucionado la cooperación: (1) la colaboración mutua, (2) la reciprocidad indirecta y (3) la selección cultural grupal.

Tomasello propone que, en cierto momento de la evolución humana, los individuos se habrían vuelto más tolerantes el uno con el otro y, en ese punto, habrían empezado a orientarse hacia una verdadera cooperación a través, probablemente, del uso imperativo del gesto de señalar —paso que no supondría ninguna habilidad cognitiva adicional a aquellas con las que cuentan los grandes simios hoy en día. El paso siguiente implicaría que estos individuos, ya capaces de coordinar acciones entre sí de forma más regular y tolerante, se hallen en una situación tal que la selección natural pudiera favorecer la maquinaria cognitiva y motivacional necesaria para llevar a cabo interacciones colaborativas más complejas.

La colaboración mutua sería el lugar de nacimiento del campo conceptual común necesario para que surja una comunicación cooperativa inferencialmente rica como la humana. En términos específicos de la comunicación, dado que estamos esforzándonos hacia una meta conjunta en colaboración mutua, será ventajoso para nosotros que ayudemos al otro. En este contexto, la tendencia del comunicador de solicitar ayuda y la del receptor de simplemente brindarla podrían surgir, naturalmente, como una manera de facilitar el progreso hacia una meta común.

Como ya se vio líneas arriba, los actos de acceder a los pedidos de los demás y, además, de ofrecerles ayuda habrían surgido en el marco de actividades de colaboración mutua: en este contexto, el que un individuo acceda al pedido de alguien

resulta adaptativo siempre en tanto lo beneficia a él mismo (dado que comparte la misma meta con su interlocutor). Sin embargo, más allá de la situación mutua en curso, el brindar ayuda a los demás resulta adaptativo también en tanto pues repercute positivamente en la reputación del individuo colaborador, quien será percibido como alguien dispuesto a ayudar a los demás —y, en ese sentido, alguien cuyo favor conviene ganar, por ejemplo, ayudándolo también.

De este modo, en un grupo de individuos capaces de llevar a cabo una lectura de mentes recursiva (y de conocer, pues, las intenciones de los demás, a la vez que son conscientes de que estos conocen también sus propias intenciones) y preocupados, además, por su reputación, las expectativas de colaboración mutua habrían surgido sin problema: la idea de que el otro estará dispuesto, naturalmente, a ayudarnos.

Finalmente, el deseo de cultivar afiliaciones con los demás miembros del grupo constituye la base de una de las motivaciones básicas de la comunicación cooperativa humana: el deseo de compartir emociones y/o actitudes. Este afán por compartir emociones y opiniones tendría una función de identidad grupal en el caso de los seres humanos y se trataría de una función ausente en el resto del reino animal.

A partir de esta motivación, basada en el deseo de ser reconocidos como miembros auténticos del grupo social, surgiría la presión necesaria para la aparición de las normas sociales. En este sentido, sobre la base de la comprensión mutua de que todos quieren ayudar (mutualismo) y de que, además, todos se encuentran preocupados por su reputación en el grupo (reciprocidad indirecta), habría evolucionado la expectativa de cooperación en las situaciones comunicativas. Si a esto se añade la presión de adaptarse a las expectativas del grupo (selección cultural grupal), emergen normas, por ejemplo de cooperación comunicativa, y, además, las sanciones respectivas en caso de incumplimiento.

Se concluye entonces que la comunicación humana depende necesariamente de formas de cooperación entre el grupo. Esto se debe a que tanto quien se expresa, como quien interpreta, deben hacer el esfuerzo de asumir la perspectiva del otro. Mientras más recursos tengan para hacerlo, tanto mejor para la comunicación.

Una manera de pensar los procesos cognitivos involucrados en la comunicación cooperativa específicamente humana es como «aquello que agrega un nivel adicional de intencionalidad a la trasmisión de información» (p. 82) —nivel que puede caracterizarse como un *para ti*, pues la efectividad de la comunicación depende de cuán bien los actores involucrados en ella son capaces de acomodar sus elementos en función de lo que consideran pertinente para el contacto con el otro. El objeto real de la comunicación es, entonces, un *nosotros* que emerge al mismo tiempo que una serie de normas sociales que regulan las formas de intercambio de información en el grupo.

En suma, de acuerdo con Tomasello, la estructura cooperativa de la comunicación constituye una manifestación más de la extrema capacidad humana para la cooperación. Las habilidades y motivaciones humanas de intencionalidad compartida habrían aparecido en el marco de actividades colaborativas mutuas, que solo habrían sido posibles cuando los humanos se volvieron más tolerantes y generosos. De este modo, habría evolucionado una nueva habilidad cognitiva: la lectura recursiva de mentes, que habría permitido la generación de metas conjuntas, las cuales, a su vez, habrían hecho posibles marcos atencionales compartidos relevantes para la meta compartida. Finalmente, a partir de la construcción de campos conceptuales comunes a los interlocutores, gestos como el de señalar y otros actos comunicativos cooperativos habrían adquirido un significado.

2.2.2. La propuesta de Esther Goody

Goody (1997) sostiene que el lenguaje es la poderosa herramienta que, desde sus primeros rudimentos como un *protolenguaje* (diferente al propuesto por Bickerton; véase §2.1.1.1.) enriqueció la vida social de los antiguos homínidos, los cuales, como hemos visto en otras secciones, ya contaban con estrategias maquiavélicas heredadas de la complejidad social de los primates anteriores. Particularmente, la autora intenta demostrar que el lenguaje impulsó el desarrollo de mecanismos de cooperación nunca antes vistos en la historia natural.

De acuerdo con la autora, existen dos factores cruciales para diferenciar a los homínidos de otros primates y que, sin embargo, no han recibido suficiente atención en la investigación evolutiva de nuestras capacidades cognitivas. De una parte, tradicionalmente los lingüistas han considerado que el lenguaje hablado apareció en la tierra con el *Homo sapiens*. De otra, los estudios de la evolución del cerebro y su relación con la evolución de las capacidades cognitivas se han centrado en las proporciones que ciertas secciones cerebrales guardan con el resto de sus partes, pero no se le ha dado mayor importancia a la función de su tamaño total. No obstante como veremos en lo que sigue, la autora piensa que, contrariamente a lo que comúnmente se ha hecho, se puede pensar una evolución gradual del lenguaje que haya tenido lugar de la mano del incremento en el tamaño del cerebro de los homínidos.

Siguiendo a William Foley (1997), Goody considera posible que antes de la aparición del homo habilis ya se haya utilizado cierto número de voces, de manera semejante a la de los chimpancés. Un primer léxico sería producto de combinaciones simples de palabras, y para el momento de la llegada del *Homo habilis* ya se podría decir que existía una suerte de protolenguaje. Ya que el *Homo erectus* contaba con un cerebro más grande que el de su antecesor, podemos pensar que entonces el protolenguaje ya contase con combinaciones más largas de palabras. Los primeros

Homo sapiens habrían sido capaces de combinar jerárquicamente unidades lingüísticas simples. Este mayor control sobre los componentes del lenguaje debería haber generado en sus usuarios conocimiento metalingüístico, específicamente respecto de *cómo significan los signos*. Ello constituiría el paso final hacia los lenguajes complejos de hoy (Goody, 1997).

El acceso al conocimiento sobre la construcción misma del significado parece haber enriquecido la interacción social de los homínidos antiguos en más de una forma. Por un lado, establecer una relación sistemáticamente reproducible entre palabras y objetos referidos por aquellas es una actividad necesariamente social. Por otro, las palabras así fijadas se convierten en artefactos o herramientas que pasan a existir «en el mundo», independientemente de sus usuarios. El empleo mismo de estos artefactos supone, además, el despliegue de una serie de prácticas cooperativas. La comunicación depende, en buena medida, de que los interlocutores compartan estrategias, generalmente tácitas, de adelanto y predicción de futuros actos de habla, lo cual se ve reflejado en secuencias de acciones coordinadas.

En suma, el lenguaje hablado, esta combinación de palabras-herramienta y secuencias de acciones socialmente instituidas, propicia la generación de representaciones colectivas. En la comunicación así concebida, los participantes deben ser capaces de adelantar pistas sobre sus propias intenciones y de reconocer las de los demás. Esto permite contextualizar las expresiones adecuadamente para hacer frente a su alta ambigüedad. Sin embargo, la ambigüedad y la posibilidad de reinterpretar retrospectivamente lo dicho cumplen un papel significativo con relación a las capacidades de engaño y contraengaño propias de la inteligencia maquiavélica. A medida que se desarrollan nuestros actos comunicativos, podemos, por ejemplo, reinterpretar significados previamente establecidos que ya no nos ayuden a cumplir nuestros objetivos o podemos alterar el sentido de lo que un interlocutor nos dice mediante las respuestas mismas que le brindamos. El tiempo es, entonces, también un elemento crucial del uso del lenguaje entendido de la manera aquí descrita.

La descripción del lenguaje que consideramos implica verlo como una herramienta que nos permite pensar y actuar con los demás. Goody afirma, además, que de todos los constructos sociales producidos gracias al uso del lenguaje, cabe destacar el papel de los *géneros comunicativos*, las *reglas* y los *roles*. Respecto de esto último, podemos decir que se trata de la idea general de que solemos esperar ciertas regularidades en el comportamiento (y, por tanto, también en lo que se significa con las palabras) de los agentes de acuerdo con la función que cumplen en el mundo social. También se observa uso de categorías en otros primates, pero el empleo del lenguaje nos permite referirnos a los objetos así clasificados incluso si estos se encuentran ausentes. Una vez establecidos los roles, podemos también hacer explícitas, gracias al lenguaje,

nuestras expectativas acerca de los individuos organizados en ellos. Las expectativas reciben el nombre de *reglas*. Tanto los roles como las reglas son existencias culturales, esto es, producidas por la interacción de los individuos, pero ahora independientes de ellos. Son artefactos que nos ayudan a determinar tanto el curso, en principio contingente, de las acciones, como nuestra forma de comprender a los demás.

Por lo dicho anteriormente, vemos que el lenguaje es una poderosa herramienta social que nos permite construir otras herramientas semejantes. Naturalmente, a mayor complejidad, también es mayor su capacidad de creación de herramientas. Goody cree que la sofisticación de los géneros conversacionales, como la ironía, las preguntas o la corrección en las formas, son muestras de ello, pues nos permiten hacer más de una cosa a la vez (como cumplir un objetivo ulterior mientras que, al mismo tiempo, hacemos escarnio de nuestro interlocutor).

Podemos decir que, de acuerdo con lo dicho, la aparición del lenguaje hablado en los homínidos no solamente habría incrementado las posibilidades de la inteligencia maquiavélica. También habría introducido nuevos modos de cooperación. Esto, a la larga, enriquecería enormemente la vida social de la especie. Como señala la autora, la cooperación supone operaciones bastante más complejas que la ganancia personal, pues no se trata únicamente de considerar y evaluar posibles escenarios, implica también hacerlo mientras se intentan sintonizar las propias expectativas con las de los interlocutores. Puede que para explicar la inteligencia maquiavélica de otros primates baste con demostrar algún grado de metarrepresentación de intenciones en ellos, pero el lenguaje implica mucho más; implica el desarrollo de mecanismos de constitución de significados compartidos.

2.3. Síntesis

En este capítulo hemos revisado algunas de las propuestas más resaltantes en torno a la evolución del lenguaje en nuestra especie. Dividimos el capítulo en dos secciones que presentan, por un lado, las principales propuestas sobre el origen y la evolución del lenguaje; y, por otro, los planteamientos más saltantes en torno a sus ventajas adaptativas para la comunicación y la cooperación en la especie.

El primer apartado estuvo dedicado a explorar las propuestas de los lingüistas Derek Bickerton (principalmente, 1990 y 2010) y Ray Jackendoff (principalmente, 1999 y 2002) sobre la evolución del lenguaje en etapas caracterizadas por la aparición de rasgos específicos de la capacidad lingüística. A partir de la revisión de datos provenientes del desarrollo lingüístico en el desarrollo y de los fenómenos de criollización, la propuesta de Bickerton sugiere un escenario discontinuo para la aparición del lenguaje en la especie. Primero, habríamos desarrollado la capacidad de representar

símbolos mediante gestos o vocalizaciones para, solamente después, poder combinarlos en cadenas significativas más complejas. La propuesta de Jackendoff extiende la de Bickerton dividiendo los escenarios planteados en nueve etapas: desde la concentración voluntaria de las vocalizaciones en un solo símbolo, hasta la combinatoria compleja propia del estado actual del lenguaje.

Concluimos este primer apartado con la presentación de un debate importante en torno a los componentes específicos del lenguaje y a su aparición y evolución en la especie. La discusión se inició a raíz de un artículo de Massimo Piattelli-Palmarini, publicado en 1989, y una serie de réplicas y contrapropuestas entre varios autores (Pinker & Bloom, 1990; Hauser, Chomsky & Fitch, 2002; Pinker & Jackendoff, 2005; Fitch, Hauser & Chomsky, 2005; Jackendoff & Pinker, 2005). El intercambio tuvo como tema central la constitución de la facultad del lenguaje y, sobre esta base, su aparición y desarrollo en la especie. Hauser, Chomsky y Fitch proponen distinguir dos componentes en la facultad del lenguaje: uno exclusivo de nuestra especie y de nuestra capacidad lingüística (FNL), y otro posiblemente compartido con otros animales y con otras capacidades cognitivas en nosotros (FLB). Los autores concluyen que, a partir de las características que FNL presenta, este no podría haber sido preferido por la selección natural, mucho menos para cumplir una función comunicativa. Por otro lado, Pinker, Bloom y Jackendoff, en diferentes momentos, insisten en que la complejidad exhibida por el lenguaje no podría, más bien, ser resultado de un mecanismo diferente al de la selección natural (esto es, una exaptación o mutación genética masiva) pues es justamente la sofisticación y complejidad del diseño de diferentes rasgos biológicos una característica del moldeamiento paulatino de la selección natural en la evolución de las especies.

No hay duda en que la aparición del lenguaje ofreció un conjunto de ventajas adaptativas a nuestra especie relacionadas a la comunicación y la cooperación en el grupo. En el segundo apartado del capítulo (§2.2.2.) presentamos dos propuestas acerca de las ventajas adaptativas que la aparición del lenguaje habría ofrecido para el desarrollo filogenético de la especie. Por un lado, Michael Tomasello (2008) sugiere que la colaboración en el grupo primate habría sido el lugar de nacimiento del campo conceptual común necesario para el surgimiento de la comunicación cooperativa inferencialmente rica como la humana. El estudio de comportamientos no lingüísticos, como el señalamiento o la pantomima, estarían así en la base del desarrollo comunicativo del grupo homínido. Por otra parte, Goody (1997) considera que un estadio inicial del lenguaje —un protolenguaje diferente al de Bickerton (1990)— habría aparecido inclusive antes que el *homo sapiens*, como puede inferirse a partir del registro cultural arqueológico, para la construcción de herramientas y el planeamiento estratégico.

Algunos investigadores sugieren que la capacidad del lenguaje como lo conocemos hoy data de aproximadamente 50 000 años atrás (Dunbar, 1996 y Fitch, 2010). En el capítulo a continuación presentaremos algunas de las propuestas más resaltantes sobre cómo un niño adquiere y desarrolla el conocimiento de una lengua que describimos previamente (ver §1.2.). Si bien el desarrollo del lenguaje involucra tanto aspectos biológicos (*nature*, o presente en el genotipo) como culturales y sociales (*nurture*, o aprendido a partir de la experiencia), las diferentes posturas teóricas se distinguen por el énfasis que advierten en cada cual.

3. EL DESARROLLO DEL LENGUAJE

El estudio del desarrollo del lenguaje analiza el proceso por el cual un niño pasa de no saber una lengua a dominar una con destreza sorprendente. Esta línea de investigación se concentra en las etapas que constituyen dicho proceso, así como en las habilidades o conocimientos que permiten el paso de un estado de conocimiento menos maduro o cercano a la lengua meta del entorno a uno más de tipo adulto y convencional.

Dentro de la Lingüística, el área específica que se ha ocupado tradicionalmente de este objeto de estudio es la Psicolingüística. La Psicolingüística del Desarrollo asume que es posible aproximarse científica y rigurosamente a los procesos de comprensión y producción lingüísticas de infantes y niños en edad preescolar con el objetivo de analizar su desarrollo. El estudio del lenguaje infantil parece revelar la presencia temprana de patrones o usos sistemáticos, lo que, además, cuestionaría seriamente aquellas posturas que consideran el habla infantil como esencialmente caótica y discontinua respecto de momentos posteriores del desarrollo (al estilo de Jakobson, 1968). El desplazamiento —aunque siempre parcial— de la atención desde la lengua meta adulta a la lengua infantil revela, pues, un uso mucho más sofisticado de lo que habríamos sospechado.

En esta línea, Tomasello y Bates (2001) sostienen que la función de la psicolingüística evolutiva consiste en analizar aquellos procesos por los que los niños aprovechan su herencia biológica —aquella capacidad genéticamente determinada que les permite adquirir la lengua empleada por las personas de su entorno— y cultural —la lengua en sí misma, tal como es usada para la interacción y comunicación en el medio en que se desarrollan— para convertirse, al cabo de unos años, en usuarios competentes de su lengua. En dicho proceso confluirían habilidades de tipo perceptivo, cognitivo, social, comunicativo y de aprendizaje.

El estudio del desarrollo del lenguaje en el niño permite, ante todo, responder a la pregunta por el origen del conocimiento lingüístico en el individuo adulto. Se trata de una pregunta cuya respuesta en el plano filogenético habrá que buscar en otra

parte; aun así, es una interrogante válida en sí misma, con una respuesta en principio distinta a la pregunta por los orígenes del lenguaje en la especie humana —aunque, en ocasiones, se tienda a considerar, erróneamente, que el desarrollo y la evolución del lenguaje no son sino dos versiones del mismo proceso, desarrollado en marcos temporales distintos.

Si bien, actualmente, la mayoría de enfoques teóricos consideran que la respuesta más adecuada involucra tanto el aspecto biológico (*nature*, o presente en el genotipo) como el cultural o social (*nurture*, o aprendido a partir de la experiencia), el peso asignado a cada una de estas fuerzas variará: ciertos enfoques teóricos destacarán el papel de lo biológico o innato, mientras que otros enfatizarán el papel del entorno y la interacción del infante con su cultura. Desde una postura que considera como igualmente importantes ambos factores, el conocimiento del lenguaje será resultado de un proceso epigenético, es decir, de la interacción entre la dotación biológica del ser humano y su desarrollo en un entorno sociocultural particular (Elman y otros, 1996). El estudio del desarrollo del lenguaje permitiría determinar el peso que cada uno de estos dos factores desempeña en el proceso de aprendizaje o adquisición de una lengua.

También, el estudio del desarrollo de la lengua en el niño permite evaluar si la capacidad lingüística es de dominio general o, más bien, de dominio específico. Esta pregunta busca determinar si la facultad que permite que, en condiciones normales, un infante que no habla o usa una lengua aún devenga, eventualmente, un adulto que es capaz de expresarse y comunicarse lingüísticamente sirve solo para aprender una lengua (dominio específico) o sirve también para otras tareas (dominio general), entre ellas para el aprendizaje de la lengua materna. Mientras las capacidades de dominio específico funcionan solo con un tipo particular de estímulos (como el input lingüístico provisto por el entorno) con un único objetivo (como la adquisición de la lengua del medio exterior), las capacidades de dominio general sirven a distintos procesos (como la habilidad humana para la detección de patrones regulares o para el establecimiento de categorías).

Finalmente, el estudio del desarrollo constituye una pieza clave para la comprensión de los comportamientos y conocimientos complejos de los que es capaz la mente adulta (Karmiloff-Smith, 1996). Ser capaces de reconstruir el camino por el cual ciertas predisposiciones innatas se encuentran en e interactúan con un determinado entorno social y cultural para dar lugar, al cabo de un tiempo, a un fenómeno tan impresionante como el conocimiento y dominio de una lengua natural redundará en una comprensión mayor de cómo funciona la mente adulta —y, evidentemente, el lenguaje. Así, el estudio del desarrollo del lenguaje puede brindar una mejor comprensión del desarrollo cognitivo general, y del estadio y funcionamiento de la mente humana.

En esta sección, ofreceremos una breve introducción a las etapas canónicas por las que pasa un niño cuando adquiere su lengua materna y presentaremos tres de los diferentes modelos que intentan explicar este proceso de adquisición.

3.1. Etapas del proceso de desarrollo de lenguas maternas

El desarrollo de una lengua materna sigue una secuencia de etapas identificables entre los seis y los treinta meses de edad. La evidencia proveniente del desarrollo infantil nos muestra que, si bien con cierta variación temporal entre individuos, todos los niños, en condiciones naturales y expuestos a experiencia lingüística, atraviesan estas etapas en el camino a convertirse hablantes eximios de la lengua de su comunidad.

Recordemos que, como señala Clark (2009), la meta en el desarrollo lingüístico del niño es convertirse en un miembro de la comunidad de hablantes. Es decir, adquirir, en la sucesión de estas etapas identificables, habilidades para poder comprender y producir enunciados en la variedad lingüística de sus pares. Por ejemplo, la habilidad para usar las formas de conjugación del castellano de Lima o las propias del voseo bonaerense surgirá durante un periodo de un año o más, comenzando por una etapa en que no se usaban flexiones verbales y finalizando con su dominio comparable al de un adulto.

Existen muchas maneras de caracterizar las etapas del desarrollo lingüístico infantil. Podemos centrarnos en la percepción (por ejemplo, cuál es el número de palabras que la niña comprende, qué características fonológicas sutiles parece diferenciar, etcétera) o en la producción (por ejemplo, cuántas palabras utiliza a diario, qué sonidos pronuncia con dificultad, etcétera). En esta presentación, tomaremos la perspectiva que se enfoca en la producción para establecer las diferentes etapas en la adquisición de una lengua. Sin embargo, a lo largo de la exposición, nos detendremos a revisar algunos experimentos y hallazgos importantes que involucran la asombrosamente temprana percepción lingüística de los niños.

La bibliografía especializada (Brown, 1973; De Villiers & De Villiers, 1985; Crain & Lillo-Martin, 1999) sugiere que las etapas de producción lingüística identificables en la adquisición infantil de una lengua son cinco: una *etapa inicial* de exploración de las cavidades del aparato fonador que va desde el nacimiento hasta los cinco o seis meses de vida; *el balbuceo*, que va de entre los seis a los ocho meses; *la etapa holofrástica*, que va de los nueve meses a los dieciocho meses; *la etapa de dos palabras*, entre los dieciocho y veinticuatro meses; y *la etapa telegráfica*, entre los veinticuatro y treinta meses de edad; dando lugar al habla similar a la de un adulto desde este momento en adelante. Recordemos que si bien todos los niños pasan por estas etapas, existe cierta variación interindividual así como la superposición parcial de las etapas en el tiempo de desarrollo. A lo largo de este apartado, revisaremos cada una

de estas etapas y ofreceremos algunos ejemplos que sirven para identificarlas. Antes de comenzar, sin embargo, examinaremos qué sucede en el momento preparatorio a la primera etapa, es decir, en ese que va desde el nacimiento a los seis meses de vida.

3.1.1. De los cero a los seis meses

Diferentes estudios (Brown, 1973; De Villiers & De Villiers, 1985; Crain & Lillo-Martin, 1999) muestran que, justo después del nacimiento, los niños son capaces de distinguir los sonidos del habla de otros sonidos en el entorno. Asimismo, son capaces de identificar diferencias mínimas o rasgos distintivos (ver §1.2.1.) entre los fonemas de la lengua a la que se encuentra expuesto. Por ejemplo, las diferencias de sonoridad entre los fonemas [t] y [d], o la del lugar de articulación entre [t] y [k] (Eimas y otros, 1971; Eimas, 1974 y 1975). No será hasta un par de meses después en que podrán diferenciar su lengua del resto, reduciendo su asombrosa capacidad perceptiva a las distinciones fonológicas características del input al que se encuentran expuestos.

Durante los primeros cuatro meses, los bebés producen chirridos, gruñidos y suspiros, característicos de la estabilidad anatómica del tracto vocal para la respiración y la deglución simultáneas. Un poco después se comienzan a producir arrullos de tipo vocálico y las producciones comienzan a darse separadas por oclusiones glotales. Hacia el final de los cuatro meses, los bebés comienzan a reír.

Entre los cuatro y aproximadamente seis o siete meses, los bebés comienzan a manipular el tono de sus producciones. Las risas, chirridos y rugidos se hacen sordos o estridentes. La manipulación del tracto vocal durante estos meses, genera la producción de sonidos nasales, labiales y labiodentales similares a bufidos; del mismo modo, se producen sonidos deslizantes, vibrantes y gorjeos que empiezan a tomar la forma de vocales y consonantes.

3.1.2. El balbuceo

El balbuceo se inicia, aproximadamente, a los seis meses de edad y continúa hasta los ocho. En esta etapa, el bebé comienza a manipular el sistema fonador con mayor destreza que en meses previos. Es durante este intervalo de tiempo que la laringe del bebé desciende y la capacidad de poder respirar y tragar alimentos al mismo tiempo se pierde.

Existen dos tipos de balbuceo. El primero consiste en la repetición de conjuntos de sonidos que exhiben un patrón regular de consonante y vocal, a manera de sílabas, de un solo tipo que se articulan; por ejemplo, «ba-ba-ba-ba», «gu-gu-gu-gu», «ne-ne-ne-ne». A este tipo de balbuceo suele denominársele *balbuceo reduplicado*. El segundo tipo, llamado *abigarrado*, empieza cuando se producen variaciones entre el tipo de sílabas que se emiten, así: ‘ka-du-du-ba-ba’, ‘ne-ne-ni-da’, ‘bu-ba-mi-do-do’ (Harley, 2008).

Se han propuesto diferentes explicaciones del balbuceo y su relación con la adquisición de una lengua. Algunos autores señalan que esta etapa constituye un paso determinante y es el inicial en la adquisición completa del lenguaje (Oller, Wieman, Doyle & Ross, 1976; Rees, 2011). En otras palabras, que el balbuceo constituye un tipo de comportamiento prelingüístico y da inicio a un proceso continuo en la adquisición de una lengua. Evidencia a favor de esta hipótesis la constituyen los niños sordos que exhiben el mismo comportamiento que un niño sano pero entre los doce y veinticuatro meses (Locke & Pearson, 1992). Asimismo, los bebés sordos que son expuestos a una lengua de signos comienzan a balbucear en modo manual entre los diez y trece meses de edad. Por otro lado, se encuentran los que consideran que el balbuceo no representa una etapa prelingüística sino, más bien, como un periodo natural en el desarrollo del tracto vocal, un estadio de «práctica» articularia previa al habla, con el fin de conseguir máximo control de los sistemas motores involucrados en la producción de sonidos lingüísticos. La argumentación se basa en casos de silencio marcado después de esta etapa y antes de adquirir el léxico, mostrando así una aparente discontinuidad en la adquisición de una lengua.

Los tipos de sonidos que producen los bebés en esta etapa (ver §1.2.1.) consisten, principalmente, en oclusivas, sordas y sonoras, y aproximantes. La evidencia muestra que las fricativas son extrañas, así como las africadas, y mucho más los grupos consonánticos. Las vocales canónicas son la [a] y otras abiertas, un poco más altas, como la [e] o la [o].

3.1.3. La etapa holofrástica

La etapa de una sola palabra u holofrástica da inicio a las tres etapas de desarrollo de la combinatoria morfológica y sintáctica: etapa holofrástica, de dos palabras y telegráfica. Este estadio, que va desde aproximadamente los nueve hasta los dieciocho meses de edad, está caracterizado por la producción de una sola palabra en intervalos de tiempo extendido. Se observan así pausas largas entre la emisión de una palabra y la siguiente.

Mientras que muchos niños se enfocan principalmente en la producción de una sola palabra, otros producen secuencias de balbuceo extendidas que cuentan como frases más largas. La noción de *holofrase* surge aquí como una noción para comprender la complejidad comunicativa que el niño pretende con su todavía escasa producción lingüística (Bavin, 2009). Scollon recuenta un experimento en el que Breda, una niña de dieciocho meses levanta un pie sobre la grabadora, pretendiendo que la está pisando, y dice las palabras *tape* ‘grabar/grabación’ y *step*, ‘pisar’, cada una con su respectivo contorno entonacional (1976, p. 70). La propuesta de Scollon avizora un desarrollo inicial ‘vertical’ de la sintaxis: las palabras se colocan como

estratos significativos, una tras de otra en un proceso de construcción holofrástica comparable al de una construcción con ladrillos. Solo posteriormente, el despliegue horizontal, caracterizado por una sola entonación e identificado con la oración gramatical, aparecerá en el desarrollo sintáctico (p. 219).

3.1.4. La etapa de dos palabras

La etapa de dos palabras se inicia alrededor de los dieciocho meses de edad. El niño comienza a producir enunciados de dos palabras. Entre estas no aparecen palabras funcionales ni tampoco afijos flexivos. Se ha observado que el proceso es bastante similar a lo largo de diferentes lenguas.

Ejemplos de la producción lingüística en la etapa de dos palabras son: «más leche», que indica un pedido; «carro yo», para indicar posesión, o «pega pelota», para describir un suceso o indicar una acción. Para algunos autores, estas construcciones pueden considerarse la base para otras innovadoras a partir de un *esquema de pivote* (Braine, 1963). Por ejemplo, a partir del enunciado «más leche», se considera posible el esquema de pivote «más _____», que podrá ser usado para conseguir otros pedidos: «más uvas», «más dulce», etcétera. El ordenamiento de los esquemas de pivote además de estar ordenado localmente, no cuenta con una sintaxis interna. Su aparición surge a partir de las formas de ordenamiento más comunes que el niño ha escuchado en el habla adulta (Bavin, 2009).

La etapa culmina con la aparición de la combinación telegráfica hacia los veinticuatro meses de edad.

3.1.5. La etapa telegráfica

La etapa telegráfica está caracterizada por tener una muy corta duración. Esta comienza aproximadamente a los veinticuatro meses y continúa hasta los treinta, momento en el cual el habla de un niño comienza a mostrar la combinación de varias palabras por enunciado.

La producción telegráfica carece de flexión irregular y se los sujetos y otros argumentos de la oración pueden aparecer elididos; por ejemplo, *mommy go store*, ‘mamá va tienda’, o *want go get it*, ‘quiero ir eso’. El orden en el que se combinan las palabras es esencial para la comprensión de los enunciados lingüísticos (ver §1.2.3.). Küntay y Slobin (1996) analizaron conversaciones entre padres e hijos turcos de edades entre los dos años y dos meses hasta los tres años y ocho meses en lengua turca. Los resultados mostraron que el porcentaje de uso de patrones de ordenamiento argumental (por ejemplo, SOV, OSV, SVO, OVS, etcétera) variaba en la misma proporción entre madres e hijos.

De acuerdo con Brown (1973), el rango de edad en el que se adquiere la recursión se da, aproximadamente, entre los dos años y once meses y los tres años y cinco meses. Es en este momento en el que se registra un límite superior de once palabras en la producción de enunciados del niño. Asimismo, alrededor de esta edad, los niños ya conocen un vocabulario aproximado de ciento cincuenta palabras (Nelson, 1973). Este dato es duplicado por las investigaciones de Fenson y otros (1993) quien, con una muestra de 1789 niños, señala el vocabulario a esta edad consta de aproximadamente 310 palabras.

A manera de conclusión podemos ofrecer el siguiente cuadro que sintetiza los principales hitos que aparecen en cada una de las etapas que hemos descrito en este apartado.

Etapas del desarrollo lingüístico	Edad aproximada	Características
De cero a seis meses	Cero a seis meses	<ul style="list-style-type: none"> - Chirridos - Gruñidos - Susurros
Balbuceo	Seis a ocho meses	<ul style="list-style-type: none"> - Patrones repetitivos de consonantes y vocales - Balbuceo reduplicativo y abigarrado
Holofrástica	Nueve a dieciocho meses	<ul style="list-style-type: none"> - Uso de clases de palabras abiertas - Ya tienen función comunicativa - 10 palabras (Fenson, 1993)
Dos palabras	Dieciocho a veinticuatro meses	<ul style="list-style-type: none"> - Pueden entenderse frases cortas - La curva entonacional indica la constitución de un enunciado - 50 palabras (Fenson, 1993, página)
Telegráfica	Veinticuatro a treinta meses	<ul style="list-style-type: none"> - Los enunciados carecen de elementos funcionales - La flexión irregular está ausente - 310 palabras (Fenson, 1993)

3.2. Modelos del proceso la adquisición de las lenguas maternas

En esta sección, presentamos tres modelos que intentan explicar la adquisición de las lenguas maternas. Nuestra selección se justifica en la relevancia y potencial explicativo de las teorías lingüísticas a partir de las cuales cada uno de estos modelos se ha desarrollado. El modelo innatista de la §2.3.2.1. se asienta en el paradigma de investigación de la gramática generativa delineado por Noam Chomsky a mediados de la década de 1950. Por otra parte, el modelo sociopragmático que presentamos

en §2.3.2.2 recoge las primeras teorizaciones de Bruner (1983) sobre la importancia de la cooperación en el desarrollo para el desarrollo lingüístico y comunicativo satisfactorio, hasta las propuestas más recientes de Tomasello (1999 y 2008) sobre esta base. Por último, los modelos conexionistas presentados en §2.3.2.2 se asientan en el desarrollo de la Teoría Computacional de la Mente de acuerdo con diferentes autores (Rumelhart & McClelland, 1986; Pylyshyn, 1984; Rumelhart, 1989 y McClelland, 1999).

3.2.1. Modelo innatista

La posición innatista (Chomsky, 1965, 1981, 1986 y 1995) propone que los seres humanos vienen dotados con la capacidad de adquirir una lengua, incluso a partir de estímulos relativamente pobres. Esa capacidad, que constituye el estado Inicial (E_0) del conocimiento lingüístico, tiene un contenido específico, que la Teoría Gramatical debe dilucidar. E_0 interactúa con los datos lingüísticos primarios (la experiencia lingüística) para producir un Estado Relativamente Estable (E_c), que constituye el conocimiento de cada hablante. La teoría sobre E_0 se denomina Gramática Universal (GU). La teoría sobre E_c se denomina Gramática Particular (GP) (1986).

Se ha observado que la idea de que el lenguaje es una forma de conocimiento está en conflicto con la concepción de la mayor parte de la tradición filosófica acerca del conocimiento (Ludow, 2011). Tradicionalmente, se considera que para que algo califique como conocimiento debe consistir de un conjunto de creencias verdaderas justificadas. Sin embargo, como observa Ludow (2011), no tiene sentido hablar de verdadero (o falso) en relación con el lenguaje, y tampoco es posible concebir la necesidad de justificación para las reglas gramaticales; por otra parte, si las creencias son actitudes proposicionales, es claro que no es posible identificar una actitud hacia una enorme porción de hechos gramaticales, que son inconscientes. En otras palabras, el lenguaje no se puede caracterizar como un conjunto de creencias verdaderas justificadas. Dado esto, si queremos mantener la idea del lenguaje como conocimiento, la única salida es postular la existencia de creencias innatas, es decir, de un cierto contenido prerregistrado en la dotación genética que permite el crecimiento de una lengua en la mente. Desde este punto de vista, la tarea del niño consiste en aprender cómo exteriorizar y poner en uso estas representaciones mentales innatas, que en conjunto forman GU; al hacerlo, construye una GP —en ese sentido, la tarea del lingüista quedará terminada solo cuando pueda explicitar completamente el contenido de GU y deducir todas las GP posibles.

Existen dos concepciones dominantes acerca del contenido de GU. La primera consiste en concebir GU como una combinación de dos elementos: (i) un conjunto de principios con valor especificado, responsables de las propiedades comunes entre

las lenguas y (ii) un conjunto de principios sin especificar, llamados parámetros, cuyo valor (concebido como la posición de un *switch*, es decir, de naturaleza binaria) es determinado por la experiencia lingüística (Chomsky, 1981 y 1986; Baker, 2001). En esta concepción, las GP se forman cuando los usuarios ajustan los *switches* para valorar los parámetros. Por ejemplo, las lenguas difieren en cuanto a la posibilidad de permitir sujetos tácitos: algunas lo permiten (como en español o quechua), y otras no (como en inglés o francés). Es interesante observar que los parámetros no tienen relación con las diferencias tipológicas: lenguas con muy distinto estatuto tipológico pueden compartir el mismo valor con respecto a algún parámetro (Newmeyer, 2005).

La segunda concepción mantiene que GU contiene una muy reducida cantidad de principios (Chomsky, 1995 y 2005; Hornstein, 2009), y que la variación no depende de parámetros prefigurados en GU sino que es simplemente una función del léxico, y en particular de la flexión (Borer, 1984 y Chomsky, 1995), o alternatively, que hay reglas específicas por cada GP (Newmeyer, 2005); algunos llaman a esto un punto de vista microparamétrico. Por ejemplo, la posibilidad de que una lengua tenga sujeto tácito no depende de una macrorregla que diferencia las lenguas sino de la posibilidad de adquirir una categoría que exprese el sujeto sin ser pronunciado (un pronombre tácito) y suficientes mecanismos que lo licencien (como una morfología flexiva rica). Una ventaja de este punto de vista es que puede explicar mejor casos en los que la variación no parece seguir un patrón binario, por ejemplo, hay lenguas (como el irlandés o el hebreo) en las cuales el sujeto puede ser tácito en ciertos paradigmas pero no en otros, y hay lenguas (como el alemán) en las que el núcleo del sintagma puede ir al comienzo en ciertas sintagmas pero va al final en otros. Si toda variación es léxica, entonces estos casos se explican con más naturalidad.

En este momento existe un debate entre los partidarios de ambas posiciones, que ha llevado, de un lado, a la postulación de ricos y detallados sistemas de categorías funcionales presentes en el léxico de manera innata (Cinque, 1999; Borer, 2005a y 2005b), así como a la formulación de macroparámetros muy abstractos que operan al nivel de diversos grupos de construcciones y operaciones (Baker, 2008). La discusión, que no ha terminado, ha sido particularmente fructífera pues ha producido un renovado interés en lo que hoy se llama sintaxis comparativa microparamétrica, que consiste en la exploración detenida de todos los rasgos que separan las lenguas (Kayne, 2012). Los estudios de adquisición del lenguaje, sin embargo, todavía muestran un sesgo microparamétrico —ver Biberauer (2008) para una revisión de la noción de parámetro en el marco de este debate—.

La naturaleza de la parte invariable de GU también ha sido discutida en las últimas décadas, en particular tras el advenimiento de llamado Programa Minimista (Chomsky, 1993 y 1995). En el modelo inicial (Chomsky, 1957), GU contenía

reglas de generación de sintagmas y un conjunto transformaciones. Estas fueron reducidas a una sola Teoría de Formación de Sintagmas (X-barra) y una sola transformación (Mueve-Alpha), más unos pocos principios que las regulaban y constreñían. Hoy se entiende que ambas son el resultado de la aplicación de una sola operación ('Ensamblar' o *Merge*): los sintagmas se crean por la iteración del ensamblado con nuevos elementos del léxico, y el movimiento consiste en el ensamblado del sintagma con una copia de una parte de sí mismo. La operación de Ensamblar no está restringida por la sintaxis, sino por dos tipos de fuerza externa a ella. En primer lugar, las especificaciones del léxico (que dan origen a la variación); en segundo lugar, las presiones de los sistemas de externalización (las interfaces con la pronunciación y con el sistema conceptual). Desde este punto de vista, la adquisición del lenguaje es concebida como un proceso de adquisición del léxico particular, y las reglas específicas de cada lengua surgen como resultado de la interacción de Ensamblar con el léxico, por una parte, y debido a las exigencias propias de la economía y eficiencia computacional, por la otra.

Desde este punto de vista, la meta teórica es que GU sea un componente mínimo del lenguaje, idealmente reducida solo a Ensamblar. A pesar de eso, operaciones adicionales a Ensamblar se han propuesto. Por ejemplo, Chomsky (2000) postula la operación Concordar, que cumple un rol crucial en el licenciamiento de sintagmas; Hornstein (2009), sin embargo, propone reducir Concordar a movimiento, es decir, a Ensamblar, aunque no deja de postular otros principios adicionales, de alcance más general. Por otra parte, el léxico no es una caja vacía sino que se forma a partir de primitivos innatos; es decir, el niño viene equipado con un inventario de nociones posibles y lo único que hace es combinarlas.

Si esta línea de trabajo resulta la correcta, los impactos sobre cómo funciona la adquisición del lenguaje son enormes, y todavía no completamente comprendidos. Aquello que es específico del lenguaje sería realmente mínimo, y el niño tiene que poner en uso habilidades no lingüísticas para el proceso de adquisición. Chomsky (2005), por ejemplo, sugiere la idea de que la habilidad del niño para procesar los datos lingüísticos primarios no es específica del lenguaje, sino que corresponde a principios generales de análisis que el niño despliega también en otros dominios; igualmente, los principios de economía y eficiencia no son exclusivos del lenguaje sino que son requisitos de todo sistema computacional —se trata de lo que Chomsky denomina el Tercer Factor en el diseño del lenguaje (siendo GU el primero y la experiencia el segundo)—.

Ante esto, nuevas preguntas de investigación surgen, renovando antiguos temas. Por ejemplo, la hipótesis del Periodo Crítico (Lennenberg, 1967) podría reformularse en términos del Tercer Factor: no un deterioro de GU sino de la capacidad

de procesar datos con la misma finura, quizá incluso un deterioro de la memoria especializada. Por otra parte, queda claro que los universales lingüísticos no pueden tener una forma como la siguiente «Todas las lenguas tienen verbos», puesto que incluso nociones como *verbo* son derivadas; de hecho, ni siquiera es necesario que Ensamblar alcance todo su potencial: es concebible una lengua que no emplee estructuras recursivas. En otras palabras, las generalizaciones a partir de datos externos no nos permiten descubrir universales, hay que generalizar a partir de reglas particulares (como ya había advertido Smith, 1989, p. 66) El problema de la pobreza del estímulo puede revisarse también a la luz de estas consideraciones: es verdad que el niño aprende una lengua sin instrucción directa y que le es posible comprender y producir enunciados que nunca ha oído (su conocimiento es más rico que su input), pero como algunos han mostrado (Lasnik & Uriagereka, 2007), una vez que se enriquece el algoritmo de adquisición con elementos propios del Tercer Factor, es posible *modelizar* el aprendizaje de estructuras nuevas, incluso en ausencia de datos directos.

Además de las consideraciones teóricas, ha llevado a la postulación de principios propios del proceso de adquisición mismo. Por ejemplo, Snyder (2007) propone, sobre la base de amplia evidencia, que el niño sigue un patrón conservador a la hora de decidir qué regla usar, y como resultado comete muy pocos errores de comisión (los errores del niño están limitados a errores de omisión y a empleos diferentes de reglas léxicas).

En suma, podemos decir que la meta innatista de explicar la adquisición del lenguaje a partir de conocimientos innatos en interacción con la experiencia se mantiene vigente como un paradigma de investigación, a pesar de los diferentes cambios en el detalle del modelo.

3.2.2. Modelo sociopragmático: de Jerome Bruner a Michael Tomasello

Michael Tomasello (1999 y 2008) parte de las ideas de Jerome Bruner (1983) para la formulación del modelo cooperativo del desarrollo del lenguaje. Ambos autores atribuyen un papel importante al uso lingüístico —específicamente, a las dinámicas o formatos de interacción compartidos por cuidadores e infantes— como motor del desarrollo comunicativo y lingüístico temprano. También, coinciden en su propuesta de otro tipo de universales como posibilitadores de la adquisición de lenguas humanas.

Así, en lugar de proponer una Gramática Universal —al estilo chomskiano—, Bruner, por un lado, propone la existencia del LASS (*Language Acquisition Support System*): una adaptación humana especial para la participación en interacciones de atención conjunta con adultos y, en ellas, para la comprensión de las intenciones y la atención adultas, lo que permitiría, eventualmente, el intercambio de roles con los adultos participantes en estas interacciones, incluido el uso de convenciones

lingüísticas particulares. Dichas rutinas de socialización progresiva serían especialmente sensibles al nivel de competencia comunicativa y lingüística del niño. A partir de estas —que constituyen el elemento central del LASS—, el infante humano iría descubriendo las distintas funciones del lenguaje, así como la forma adecuada y convencional de expresarlas.

Tomasello (2003), por otro lado, propone que la base de la comunicación humana serían adaptaciones biológicas para la cooperación y la interacción social en general (universales de la cognición humana), no una base biológica de tipo gramatical o representacional (universales lingüísticos). En el marco de su modelo del desarrollo del lenguaje humano, las innegables dimensiones más formales o estrictamente lingüísticas serían construidas culturalmente y transmitidas en el marco de comunidades específicas. En este sentido, retoma el valor de los formatos de interacción en el planteamiento de su modelo cooperativo de la comunicación (y del lenguaje). Dichos formatos de interacción social, diseñados *ad hoc* y modelados en función de las competencias en desarrollo del niño, constituyen formas de cultura en las que las habilidades cognitivas y sociocognitivas infantiles permitirían una comprensión suficiente de lo que está ocurriendo, lo que posibilitaría la adquisición de las convenciones lingüísticas pertinentes a la dinámica en cuestión. De acuerdo con su perspectiva, el desarrollo lingüístico del infante se basaría en dos habilidades cognitivas generales: lectura de intenciones (dimensión funcional) y detección de patrones (dimensión gramatical).

Ambos autores destacan, pues, la adquisición del lenguaje como instrumento para regular la actividad y la atención conjunta en el contexto de patrones de interacción ritualizada entre infante y cuidador —no tanto como sistema formal de unidades y reglas de combinación. En este sentido, Tomasello se refiere a la comunicación —y al lenguaje— como una extensión especializada y convencional de acción cooperativa.

3.2.2.1. Modelo cooperativo de la comunicación

Según Tomasello, la especial habilidad humana para cooperar con otros se halla estructurada a partir de la intencionalidad compartida, es decir, la habilidad de identificar las intenciones de los demás y compartirlas con el fin de mantener metas conjuntas. Esta habilidad de lectura de intenciones emergería, en el sujeto humano, entre los nueve y los doce meses de edad.

Dado que el conocimiento de la lengua descansaría sobre una infraestructura no específicamente lingüística —comprensión intencional y campo conceptual común—, Tomasello (2008) postula la conveniencia de emprender el estudio de la comunicación humana no a partir del lenguaje sino de los gestos tempranos, en los que se encontrarían los orígenes del desarrollo lingüístico en los seres humanos.

La comunicación lingüística y la gestual compartirían una serie de aspectos: el establecimiento de un terreno conceptual común entre los interlocutores, las mismas intenciones comunicativas y la misma búsqueda de relevancia por parte del receptor, guiada por asunciones mutuas de cooperatividad. En el caso de las convenciones comunicativas arbitrarias —como las que componen las lenguas humanas—, intervendrían también habilidades para el aprendizaje cultural, es decir, el tipo de aprendizaje imitativo orientado a acciones intencionales, por el cual el sujeto invierte los roles con su interlocutor con el fin de imitar su acción (*role reversal imitation*).

De este modo, aun antes del desarrollo del lenguaje, los infantes humanos estructuran su comunicación —primero gestual, luego lingüística— de forma cooperativa, conforme se van desarrollando sus habilidades más generales de intencionalidad compartida. Los motivos tempranos exhibidos en la comunicación lingüística serán, en realidad, los mismos presentes en el caso del gesto de señalar, es decir, informar, requerir o demandar, compartir opiniones o actitudes. Se observa, además, una sincronía evolutiva estrecha entre la comunicación gestual y la lingüística, debido a que tanto los gestos como el lenguaje se aprenden y se usan con el mismo nexo interpersonal de intencionalidad compartida. De este modo, la comunicación prelingüística serviría de base para la posterior adquisición de las convenciones lingüísticas arbitrarias.

3.2.2.2. *Las primeras convenciones lingüísticas*

Hacia su primer cumpleaños, el niño comprende las intenciones comunicativas de su interlocutor en el contexto de formatos de interacción como los antes comentados; comprende las metas de los demás, lo que le permite participar en rutinas de acción conjunta de forma más sofisticada; comprende, aunque de forma básica, la importancia del conocimiento compartido con el interlocutor; y cuenta con la habilidad y la motivación de ayudar a los demás a alcanzar una meta común. La presencia de estas habilidades en edades tan tempranas ha sido tomada como evidencia de una teoría de la mente mucho más robusta de lo antes sospechado (Carpenter, 2009).

También hacia su primer cumpleaños, el niño empieza a comprender y producir convenciones lingüísticas. La adaptación humana para la comunicación simbólica emerge más claramente conforme se van desarrollando nuevas habilidades sociocognitivas: atención conjunta, lectura de intenciones y el tipo de aprendizaje cultural caracterizado por la imitación con inversión de roles.

En cuanto al marco de atención conjunta, los periodos en que infante y adulto prestan atención al mismo objeto o situación de forma conjunta establecen el terreno común en el que se desarrollan sus intercambios comunicativos. A partir de este, el niño podrá comprender las intenciones comunicativas adultas cuando se use una

nueva palabra. En segundo lugar, la comprensión de las intenciones comunicativas se desarrolla en los marcos de atención conjunta anteriormente descritos, los cuales crean un terreno intersubjetivo en el que niño y adulto ven facilitada la tarea de comprender los intentos comunicativos de su interlocutor. Aunque es probable que, en los meses previos al inicio del lenguaje, el niño haya empezado ya a comprender las acciones intencionales de los demás, las comunicativas constituyen un tipo particular de intenciones: no solo se dirigen hacia un objeto, sobre el que se ejecuta cierta acción, sino hacia los estados intencionales de otro sujeto. En tercer lugar, una vez que ha comprendido que los demás establecen relaciones intencionales con el mundo —tal como él mismo hace—, puede enfocarse en los medios conductuales de los que se valen para lograr sus objetivos e imitarlos. El aprendizaje de un símbolo comunicativo implica la habilidad de imitación con inversión de roles: el niño aprende a usar un símbolo con el adulto de la misma manera que el adulto lo usa con él. Para ello, sustituye al adulto como actor y, además, se sustituye por el adulto como objetivo del acto comunicativo intencional, es decir, reemplaza su propio estado atencional como meta por el del adulto. El resultado es un símbolo lingüístico: un medio comunicativo comprendido intersubjetivamente por ambas partes de la interacción.

En suma, poco después de su primer cumpleaños, los infantes aprenden a comunicarse simbólicamente acerca de las escenas importantes de sus vidas por medio de expresiones lingüísticas convencionales. Sus enunciados suelen reflejar los componentes más saltantes de los enunciados completos que los adultos usan en esas situaciones. Finalmente, funcionalmente hablando, los enunciados infantiles de una unidad constituyen bloques semántico-pragmáticos completos (holofrasas) que expresan una intención comunicativa relativamente coherente, aunque indiferenciada aún y dependiente del contexto (Tomasello, 2003).

3.2.2.3. Teoría sociopragmática de la adquisición léxica

El aprendizaje de palabras estaría guiado, por un lado, por el estructurado mundo social en que el niño nace (lleno de rutinas, juegos sociales, etcétera) y, por otro lado, por sus capacidades sociocognitivas para participar en dicho entorno (atención conjunta, lectura de intenciones y aprendizaje cultural). Así, el niño aprendería la mayor parte de su lenguaje más temprano en estas rutinas culturales, gracias a sus habilidades sociocognitivas, que le permitirían participar intersubjetivamente. Dentro de estos marcos de atención conjunta, los adultos suelen usar enunciados lingüísticos para atraer la atención del niño hacia determinados aspectos de la situación. En su intento por comprender la intención comunicativa del adulto, el niño se valdría de estrategias interpretativas basadas en la asunción pragmática de que los enunciados son relevantes para la situación social en curso.

Desde esta perspectiva sociopragmática, el aprendizaje léxico no involucraría tanto una tarea cognitiva reflexiva de asociación de palabras con ciertos aspectos del mundo, sino que este emergería naturalmente a partir de la interpretación, en el marco de ciertas interacciones sociales, de las intenciones comunicativas adultas expresadas en enunciados. Primero, se desentrañaría la intención comunicativa global de un enunciado y, después, se identificaría el papel funcional específico que desempeña una palabra en dicha intención (*blame-assignment*, Tomasello, 2009). El terreno común creado por el marco de atención conjunta restringiría el espacio de hipótesis posibles respecto de la intención del interlocutor, sin que sea necesario postular restricciones o principios específicamente lingüísticos.

3.2.2.4. *Enfoque basado en el uso del desarrollo sintáctico*

Esta postura propone un proceso de aprendizaje lingüístico infantil más gradual y dependiente de lo léxico, y a la vez menos sistemático. Así, a diferencia de aquellas teorías que conciben el lenguaje como un conjunto de reglas formales, más bien algebraicas y vacías, Tomasello (2003) entiende la competencia lingüística como el dominio de un repertorio estructurado de construcciones lingüísticas significativas, en el que tanto estas como las categorías implicadas constituirían símbolos lingüísticos significativos empleados en la comunicación.

De acuerdo con este enfoque, el niño aprende, a partir de los enunciados adultos, estructuras lingüísticas con significado y, luego, produce sus propios enunciados para expresar sus intenciones comunicativas inmediatas. Las construcciones resultantes no serían el producto de reglas vacías, sino de patrones de uso, portadores de un significado o función comunicativa.

Así, según Tomasello (2009), entre los dieciocho y los veinticuatro meses, empezaría la producción de enunciados multipalabra, a partir de la comprensión de las escenas recurrentes de interacción. Conforme progresa su desarrollo, el niño (1) secciona estas escenas específicas en sus elementos componentes, cada uno indicado por un símbolo lingüístico diferente, y (2) se vale de símbolos sintácticos, como el orden de palabras o los marcadores de caso, para identificar los roles que estos componentes desempeñan en la escena global.

Las combinaciones de dos palabras (holofrases) aparecen en situaciones en que ambas son relevantes («perro patio», ante un perro que está en el patio). Se trata de enunciados totalmente compuestos por piezas concretas de lenguaje, no categorías. Hacia la misma edad, sin embargo, algunas de las combinaciones empiezan a exhibir un patrón más sistemático: una palabra o frase estructura el enunciado en tanto determina la función del acto de habla global, mientras que el otro ítem lingüístico (o los otros) simplemente llena la posición con valor variable. Este primer tipo de abstracción lingüística (como la observada en «más leche», «más papa»,

«más pan») se conoce como esquema de pivote (*pivot schema*, Braine, 1963, citado por Tomasello, 2003). La constitución de ciertas categorías lingüísticas se basaría en los tipos de ítems lingüísticos que pueden ocupar la misma posición en un esquema particular. Este tipo de construcciones no exhiben aún generalizaciones basadas en varios esquemas. Además, aunque presentan ya un orden consistente, no tienen sintaxis: no hay uso contrastivo de símbolos sintácticos productivos para indicar qué papel desempeña una palabra en un enunciado mayor.

Hacia los veinticuatro meses, aparecen las construcciones basadas en ejemplares, que sí exhiben ya señales sintácticas en su construcción (morfología, orden de palabras, etcétera) para marcar los roles que los participantes desempeñan en los eventos, incluyendo posiciones (*slots*) generalizadas que incluyen categorías completas de entidades como participantes. Dicho marcado sintáctico se relaciona estrechamente con el uso de un verbo específico en el input lingüístico recibido por el niño. Por ello, en el mismo periodo evolutivo, algunos verbos pueden ser usados solo en un tipo de construcción bastante simple («cortar + ____»), y otros en marcos más complejos y variados («dibujar + ____», «dibujar + ____ + en ____», «dibujar + ____ + para + ____», «____ + dibujar en + ____»). Este marcado sintáctico temprano sería local, es decir, aprendido para cada verbo de forma individual, por lo que semejarían islas construccionales y no darían lugar a procesos más amplios de generalización.

En suma, las representaciones lingüísticas subyacentes a la producción son, en un inicio, bastante concretas, basadas en palabras individuales con algunas posiciones tipo variable. Este patrón se derivaría del input provisto por el entorno, basado a menudo en palabras y frases específicas —a lo que se añade que, en esta etapa, el infante no ha oído aún suficiente lenguaje como para hacer generalizaciones profundas. Las generalizaciones en este punto del desarrollo son producto de la esquematización —proceso de descubrimiento de patrones que elabora construcciones a partir de piezas lingüísticas concretas. Recién a los treinta y seis meses, las construcciones devendrán más abstractas, por medio del marcado sintáctico de los participantes de acuerdo con clases generales de verbos y constituyentes gramaticales.

Finalmente, si concebimos la competencia gramatical temprana en términos de patrones construccionales asociados con un determinado contenido semántico (o con una intención comunicativa particular), deviene innecesario postular procesos adquisicionales distintos que los propuestos para el aprendizaje de palabras. De esta forma, Tomasello (2009) rechaza, también, el argumento de la pobreza del estímulo, propio de aquellas corrientes que proponen una Gramática Universal de carácter innato: una gramática infantil del tipo aquí propuesto es perfectamente explicable por las sofisticadas habilidades de aprendizaje del infante (categorización, analogía y aprendizaje distribucional) y, además, evidencia efectos de la frecuencia, consistencia y complejidad con que las construcciones ocurren en el input del entorno.

3.2.3. Modelos conexionistas

Uno de los principales argumentos esgrimidos a favor de la explicación innatista del desarrollo del lenguaje es el de la pobreza del estímulo (Chomsky, 1957 y 1967, ver §3.2.1.). Se sostiene, así, que las estructuras alcanzadas al término del desarrollo lingüístico son demasiado complejas y homogéneas, que el tiempo es demasiado corto y que el estímulo es demasiado degradado y fragmentario, como para que una parte muy importante de lo que está presente en el estadio final de dicho desarrollo no haya estado ya presente en el estadio inicial del mismo: es lo que se conoce como Gramática Universal. Algo muy significativo en este argumento es su carácter negativo, pues el hipotético innatismo de esas estructuras no es propiamente un hallazgo positivo de la investigación, sino que es más bien una consecuencia de la incapacidad de las teorías del aprendizaje para explicar cómo estas estructuras podrían adquirirse mediante los mecanismos que postulan. Es importante, en este sentido, tener en cuenta que cuando Chomsky propuso este argumento tenía en mente fundamentalmente dos teorías del aprendizaje, por un lado la teoría conductista de Skinner y, por el otro, la teoría constructivista del desarrollo de la inteligencia de Piaget. Es por ello muy significativo que desde mediados de los años ochenta haya surgido una nueva teoría del aprendizaje que sí pretende estar en condiciones de explicar el desarrollo del lenguaje sin postular estructuras innatas de conocimiento (en el sentido de una Gramática Universal), cuestionando así el argumento de la pobreza del estímulo. Nos referimos al emergentismo y a su implementación computacional, el conexionismo.

3.2.3.1. *La concepción emergentista del desarrollo*

Los emergentistas consideran que mecanismos de aprendizaje simples, que operan a lo largo de los sistemas humanos para la percepción, la acción motora y la cognición, en la medida en que son expuestos a los datos lingüísticos como parte de un entorno social humano comunicativamente rico, por un organismo deseoso de explotar la funcionalidad del lenguaje, son suficientes para desencadenar la emergencia de representaciones lingüísticas complejas (Ellis, 1998, p. 657). Los representantes de esta corriente cuestionan de este modo los supuestos fundamentales de la teoría de la adquisición del lenguaje propuesta por Chomsky (1965, 1981 y 1986), en particular en el marco de su teoría de Principios y Parámetros, como la idea de que gramática es un sistema de manipulación de símbolos, el argumento de la pobreza del estímulo mencionado anteriormente, la noción de un instinto del lenguaje, las tesis de que las restricciones de la Gramática Universal están representadas en el cerebro de manera innata, de que existen universales lingüísticos hereditarios, y de que la facultad del lenguaje es modular, y en general la concepción de la adquisición del lenguaje como

fijación de parámetros (Ellis, 1998, pp. 632-633). Desde la perspectiva emergentista, interacciones que tienen lugar en todos los niveles, desde los genes hasta el entorno, dan lugar a formas y comportamientos emergentes. Estos pueden estar altamente restringidos y ser universales, pero ellos mismos no están contenidos directamente en los genes de ninguna forma específica para ningún dominio (1998, p. 638). Es por ello que los emergentistas se interesan por el conexionismo, en la medida en que este les ofrece un conjunto de herramientas computacionales para explorar las condiciones bajo las cuales surgen las propiedades emergentes. Los enfoques conexionistas al lenguaje investigan, así, las representaciones que pueden surgir cuando mecanismos de aprendizaje simples son expuestos a la compleja evidencia lingüística (p. 645).

3.2.3.2. *¿Qué es el conexionismo?*

Para entender qué es el conexionismo es necesario situarse en el marco más amplio del desarrollo de la ciencia cognitiva en general. Al surgir en oposición al paradigma conductista, la ciencia cognitiva retoma como objeto de estudio a la mente, pero considerándola un mecanismo computacional. La primera forma en que se concibió dicho mecanismo fue comparándolo a las computadoras digitales, dando lugar a la llamada Teoría Computacional de la Mente (TCM), para la cual la mente es un mecanismo de estados discretos (*discrete-state device*) que almacena representaciones simbólicas y las manipula de acuerdo a reglas sintácticas (Horst, 1999, pp. 170-171). La TCM está asociada además con lo que se conoce con una arquitectura simbólica de la mente, que está básicamente tomada de las computadoras digitales que poseen una memoria de trabajo, una memoria a largo plazo, un procesador central que toma decisiones y mecanismos de input y output. Se distingue además tres niveles (Pylyshyn, 1984), que son el semántico-intencional, el simbólico y el sustrato físico: las propiedades sintácticas son los determinantes causales del razonamiento, la sintaxis (nivel simbólico) registra (*tracks*) la semántica, y las propiedades sintácticas pueden ser implementadas en un sistema físico (Horst, 1999). Un aspecto muy importante de la TCM es lo que se conoce como funcionalismo: así como en las computadoras se puede describir el nivel simbólico independientemente del sustrato físico de su realización, se asumió que la descripción de las computaciones en el nivel simbólico de la cognición humana, que sería el más importante, era independiente también de su realización física en un cerebro humano.

El conexionismo surge precisamente por una insatisfacción con el funcionalismo de la TCM. Manteniendo la idea inicial de la ciencia cognitiva de que la mente es un mecanismo computacional, se propone modelos computacionales inspirados neurológicamente, es decir, modelos que tomen en cuenta el tipo de sustrato físico en el cual dichas computaciones deben tener lugar. Los modelos conexionistas,

llamados también modelos de procesamiento distribuido paralelamente (*parallel distributed processing models*) o redes neuronales (*neural networks*), parten del supuesto de que la cognición natural tiene lugar mediante las interacciones de grandes cantidades de simples unidades de procesamiento. Estos modelos se inspiran en el hecho de que el cerebro parece estar compuesto por grandes cantidades de tales unidades, las neuronas. En los sistemas conexionistas, una representación mental activa es un patrón de activación a través (*over*) del conjunto de unidades de procesamiento del modelo. El procesamiento tiene lugar mediante la propagación de activación entre las unidades, a través de conexiones que tienen diferentes pesos. El ‘saber’ que gobierna el procesamiento consiste en los valores de los pesos de conexión, y el aprendizaje ocurre a través de la adaptación gradual de los pesos de conexión (McClelland, 1999 y Rumelhart, 1989). Como vemos, hay dos diferencias muy importantes entre las arquitecturas conexionistas y las simbólicas: en las primeras no hay un único procesador central, sino muchas unidades simples de procesamiento que interactúan mediante las conexiones que poseen en una red, y el saber no está almacenado en una memoria separada del sistema, sino que está almacenado directamente en las conexiones entre las unidades, o en los pesos de las mismas.

Si bien en los modelos conexionistas tempranos los pesos de las conexiones eran asignados manualmente, más adelante se desarrollaron algoritmos mediante los cuales las redes podían aprender por su cuenta los valores para sus pesos (es decir, podían auto-programarse). Se exponía a las redes a ejemplos del comportamiento buscado (*target behavior*), y mediante el aprendizaje, la red era capaz de aprender a ajustar los pesos, en pequeños pasos incrementales, de modo que con el tiempo la precisión de la respuesta de la red mejoraba (Elman, 2001, p. 298).

3.2.3.3. Conexionismo y lenguaje

En 1986, Rumelhart y McClelland publicaron un influyente artículo en el que describían una red conexionista capaz de aprender el *past tense* del inglés. Al recibir varios ejemplos del tipo [*walk*→*walked*], la red era capaz no solo de producir correctamente el pretérito de aquellos verbos a los que había sido expuesta, sino, además, también el de nuevos verbos, incluyendo hasta verbos irregulares [*sing*→*sang*]. Lo más importante es que no se había instruido reglas explícitas a la red, sino que esta había aprendido el patrón de formación del *past tense* mediante un proceso inductivo basado en muchos ejemplos. La conclusión de Rumelhart y McClelland fue que si bien la noción de *regla* podía servir para describir el comportamiento tanto de los niños como de las redes, esta no jugaba ningún papel en la producción de dicho comportamiento. Esta afirmación controversial generó un gran debate, que llevó a refinamientos ulteriores del modelo (Elman, 2001, p. 298).

3.2.3.4. *Conexionismo, emergentismo y la crítica del innatismo*

En un influyente libro publicado en 1996, Elman y otros proponen una revisión profunda del concepto del innatismo en el desarrollo desde un enfoque emergentista, que trata de integrar resultados de las neurociencias, de la genética y de las teorías conexionistas en la ciencia cognitiva. Frente a una primera concepción del desarrollo como el despliegue de conductas predeterminadas y un saber innato, y a una segunda concepción del desarrollo como aprendizaje inductivo (*inductive learning*), el que implicaría copiar o internalizar conductas presentes en el entorno, Elman y sus colaboradores proponen un enfoque emergentista (propuesto también en su momento por Piaget), que concibe al desarrollo como el resultado de la *interacción* entre factores madurativos que están bajo control genético por un lado, y el entorno, por el otro. Para este enfoque, los caminos que llevan del genotipo al fenotipo pueden ser altamente indirectos y no obvios (Elman y otros, 1996, p. 1). Estos investigadores proponen además volver a pensar la noción de innatismo desde esta perspectiva, distinguiendo tres tipos de innatismo.

El primero postula restricciones representacionales innatas; el segundo, restricciones arquitecturales innatas; y el tercero, restricciones temporales innatas. En relación al primero, observan que una de las hipótesis más fuertes que se puede plantear es que el saber y los comportamientos son innatos debido a que las representaciones en las que estos se apoyan estarían implementadas neurológicamente (*hard-wired*) de antemano. Se ha planteado, así, que los niños nacen con un conocimiento innato de los principios de la gramática (Lightfoot, 1989 y Pinker, 1994). Elman y otros plantean un cuestionamiento radical a esta tesis, señalando que la implementación neural de un tal saber en el cerebro tendría que tener la forma de patrones muy específicos de conectividad sináptica a nivel cortical, y que un innatismo representacional de este tipo es muy poco común en organismos más elevados, al menos en el nivel cortical, en que lo que encontramos (en particular en los seres humanos) es que el córtex ha evolucionado como un 'órgano de plasticidad' capaz de codificar una amplia variedad de tipos representacionales (1996, p. 26). Señalan, sin embargo, que esto no significa que no puede haber otras maneras en las cuales los genes pueden operar para asegurar formas de organización cerebral propias de cada especie, así como los pensamientos y conductas mediadas por esas formas de organización cerebral, que son las arquitecturales (características de las unidades, de sus conexiones locales y de las conexiones en el sistema) y las temporales (secuencia de los eventos del desarrollo). Estos autores subrayan así que si bien desde un punto de vista teórico las restricciones representacionales (que son la forma más fuerte de innatismo) pueden ser plausibles, la investigación realizada en las últimas dos décadas acerca del desarrollo cerebral de los vertebrados nos obliga a concluir que es muy poco probable

que existan especificaciones innatas de conectividad sináptica en el nivel cortical: «[a]rgüimos, por tanto, que el innatismo representacional es con poca frecuencia, si alguna vez, una postura sostenible» (1996, p. 361).

Elman y sus colaboradores rechazan asimismo la noción de un instinto del lenguaje. Luego de observar que la evidencia más persuasiva acerca de que un comportamiento es innato la encontramos, por ejemplo, cuando una araña teje una telaraña perfecta en su primer intento, sin haber tenido una oportunidad previa de observar a otra araña tejer su tela, pues tales ejemplos constituyen evidencias de performance en ausencia de experiencia y aprendizaje, sostienen que el lenguaje humano no cae directamente dentro de esta clase de comportamientos, pues los niños requieren por lo menos tres años para alcanzar una gramática completa y que funcione (p. 384). En relación al argumento de la pobreza del estímulo, finalmente, comentan que las simulaciones conexionistas del aprendizaje del lenguaje pueden ser vistas como pruebas empíricas de estas aseveraciones acerca de la posibilidad de la gramática de ser aprendida (*learnability*), que están basados en supuestos diferentes acerca de la naturaleza del saber gramatical y de la naturaleza del mecanismo que aprende (p. 385).

En síntesis, el enfoque defendido por los emergentistas y conexionistas concibe al lenguaje como un producto, específico para un dominio, que emerge mediante la interacción de múltiples restricciones, aunque ninguna de estas es específica para el lenguaje, como puede apreciarse en la siguiente cita de Elman: «Aquí, el lenguaje es entendido como un producto (*outcome*) de dominio específico que emerge a partir de la interacción de diversas restricciones, ninguna específica al lenguaje» (2005, p. 114; ver además Ellis, 1998; Elman, 2001; Elman y otros, 1996).

3.3. Síntesis

En este capítulo hemos presentado algunas de las propuestas más resaltantes sobre la adquisición infantil de una primera lengua. Notamos la importancia del estudio del desarrollo del lenguaje en el niño pues permite responder a la pregunta por el origen del conocimiento lingüístico en el individuo adulto. Pregunta diferente de aquella que indaga por los orígenes filogenéticos de la facultad y que revisamos en el capítulo anterior.

La primera sección estuvo dedicada a la presentación de las diferentes etapas del desarrollo lingüístico por las que un niño atraviesa. La revisión bibliográfica menciona cinco más resaltantes: el inicio del desarrollo de cero a seis meses, el balbuceo (seis a ocho meses), la etapa holofrástica (ocho a dieciocho meses), la etapa de dos palabras (dieciocho a veinticuatro meses), la etapa telegráfica (veinticuatro a treinta meses). Si bien existe variación interindividual y, en algunos casos las etapas pueden sobreponerse en el tiempo, todos los niños las atraviesan camino a convertirse en eximios hablantes de una lengua natural.

Las propuestas en torno al desarrollo lingüístico del niño varían en el énfasis que colocan sea en los aspectos biológicos, o en los socioculturales, ambos determinantes para la óptima adquisición de una lengua. En la segunda sección del capítulo, revisamos tres modelos teóricos que suponen diferentes mecanismos cognitivos para el desarrollo lingüístico temprano: el modelo innatista, que encuentra su base en la gramática generativa delineada por Noam Chomsky; el modelo sociopragmático, propuesto por Michael Tomasello en consecuencia a las propuestas de Jerome Bruner y que enfatiza la función comunicativa identificables en el proceso; y finalmente, los modelos conexionistas, que conciben la *emergencia* de una gramática como el resultado de la interacción de múltiples restricciones cognitivas, que no son específicas del lenguaje, desde el primer contacto entre el niño y su entorno.

REFERENCIAS BIBLIOGRÁFICAS

- Aitchison, Jean (1994). *Words in the Mind. An Introduction to the Mental Lexicon*. Segunda edición. Oxford: Blackwell.
- Baker, Mark C. (2001). *The Atoms of Language: The Mind's Hidden Rules of Grammar*. Nueva York: Basic Books.
- Baker, Mark C. (2008). *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.
- Barkow, Jerome, Leda Cosmides & John Tooby (1992). Introduction. Evolutionary Psychology and conceptual integration. En Jerome Barkow, Leda Cosmides y John Tooby (eds.), *The Adapted Mind. Evolutionary Psychology and the Generation of Culture* (pp. 3-18). Nueva York: Oxford University Press.
- Barton, Robert (2007). Evolution of the Social Brain as a Distributed Neural System. En Robin Dunbar y Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 129-144). Nueva York: Oxford University Press.
- Bavin, Edith L. (ed.) (2009). *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press.
- Beebe, Beatrice, Doriene Sorter, Judith Rustin & Steven Knoblauch (2003). A Comparison of Meltzoff, Trevarthen, and Stern. *Psychoanalytic Dialogues*, 13(6), 777-804.
- Biberauer, Theresa (2008). Introduction. En Theresa Biberauer (ed.), *The Limits of Syntactic Variation* (pp. 7-72). Ámsterdam: John Benjamins.
- Bickerton, Derek (1990). *Language & Species*. Chicago: University of Chicago Press.
- Bickerton, Derek (2007). Language Evolution: A Brief Guide for Linguists. *Lingua*, 117(3), 510-526.

- Bickerton, Derek (2009). *Adam's Tongue: How Humans Made Language, How Language Made Humans*. Nueva York: Hill and Wang.
- Bickerton, Derek (2010). *How Language Began* [manuscrito]. Conferencia magistral. Lima: PUCP.
- Blakemore, Diane (1992). *Understanding Utterances. An Introduction to Pragmatics*. Oxford: Blackwell.
- Borer, Hagit (1984). *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris Publications.
- Borer, Hagit (2005a). *In Name Only. Structuring Sense*. Volumen I. Oxford: Oxford University Press.
- Borer, Hagit (2005b). *The Normal Course of Events. Structuring Sense*. Volumen II. Oxford: Oxford University Press.
- Bouchard, Marc-André & Serge Lecours (2008). Contemporary approaches to mentalization in the light of Freud's project. En Fredric Busch (ed.), *Mentalization. Theoretical Considerations, Research Findings and Clinical Implications* (pp. 103-132). Nueva York: The Analytic Press.
- Braine, Martin D. S. (1963). The Ontogeny of English Phrase Structure: The First Phrase. *Language*, 39, 1-13.
- Brown, Roger (1973). *A First Language: The Early Stages*. Londres: George Allen & Unwin.
- Brown, Keith & Sarah Ogilvie (2009). *Concise Encyclopedia of Languages of the World*. Oxford: Elsevier.
- Bruner, Jerome (1973). Organization of Early Skilled Action. *Child Development*, 44, 1-11.
- Bruner, Jerome (1975). From Communication to Language. A Psychological Perspective. *Cognition*, 3, 225-287.
- Bruner, Jerome (1983). *Child's Talk. Learning to Use Language*. Nueva York-Londres: W. W. Norton & Company.
- Buller, David (2005). *Adapting Minds. Evolutionary Psychology and The Persistent Quest for Human Nature*. Londres-Cambridge: The MIT Press.
- Busch, Fredric (2008). *Mentalization. Theoretical Considerations Research Findings and Clinical Implications*. Nueva York: The Analytic Press.
- Byrne, Richard (1997). The Technical Intelligence Hypothesis. En Andrew Whiten y Richard Byrne (eds.), *Machiavellian Intelligence II. Extensions and Evaluations* (pp. 289-311). Cambridge: University Press.
- Cairns, Robert, Glen H. Elder & Jane Costello (1996). *Developmental Science*. Nueva York: Cambridge University Press.

- Calvin, William & Derek Bickerton (2001). *Lingua ex Machina. Reconciling Darwin and Chomsky with the Human Brain*. Segunda edición. Cambridge: The MIT Press.
- Carpendale, Jeremy & Charlie Lewis (2010). The Development of Social Understanding. A Relational Perspective. En Willis Overton (ed.), *The Handbook of Life-Span Development. Cognition, Biology and Methods*. Volumen I (pp. 584-627). Nueva Jersey: John Wiley & Sons.
- Carpenter, Malinda (2009). Just How Joint is Joint Action in Infancy? *Topics in Cognitive Science*, 1, 380-392.
- Carruthers, Peter (2009). How We Know Our Own Minds: The Relationship between Mindreading and Metacognition. *Behavioral and Brain Sciences*, 32(2), 121-182.
- Chierchia, Gennaro & Sally McConnell-Ginet (2000). *Meaning and Grammar. An Introduction to Semantics*. Segunda edición. Cambridge: The MIT Press.
- Chomsky, Noam (1957). *Syntactic Structures*. Berlín: Mouton de Gruyter.
- Chomsky, Noam (1967[1959]). Review of Skinner's Verbal Behavior. En Leon A. Jakobovits y Murray S. Miron (eds.), *Readings in the Psychology of Language* (pp. 142-143). Englewood Cliffs: Prentice Hall.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.
- Chomsky, Noam (1971). *El lenguaje y el entendimiento*. Barcelona: Seix Barral.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1986). *Knowledge of Language: its Nature, Origins and Use*. Nueva York: Praeger.
- Chomsky, Noam (1989). *El conocimiento del lenguaje. Su naturaleza, origen y uso*. Madrid: Alianza.
- Chomsky, Noam (1993). A Minimalist Program for Linguistic Theory. En Ken Hale y S. Jay Keyser (eds.), *The View from Building 20* (pp. 1-52). Cambridge: The MIT Press.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge: The MIT Press.
- Chomsky, Noam (2000). Minimalist Inquiries: The Framework. En Roger Martin, David Michaels y Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* (pp. 89-155). Cambridge: The MIT Press.
- Chomsky, Noam (2005). Three Factors in Language Design. *Linguistic Inquiry*, 36(1), 1-22.
- Chomsky, Noam (2010). Some Simple Evo Devo Theses: How True might They Be for Language? En Richard Larson, Viviane Déprez y Hiroko Yamakido (eds.), *The Evolution of Human Language* (pp. 45-62). *Biolinguistic Perspectives*. Cambridge: Cambridge University Press.
- Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistics Perspective*. Oxford: Oxford University Press.

- Clark, Eve (2009). *First Language Acquisition*. Segunda edición. Cambridge: Cambridge University Press.
- Comrie, Bernard (ed.) (2009). *The World's Major Languages*. Segunda edición. Nueva York: Routledge.
- Coral, Karen & Jorge Iván Pérez Silva (2004). *Manual de gramática del castellano. Variedad estándar y usos regionales*. Lima: Proeduca-GTZ.
- Crain, Stephen & Diane Lillo-Martin (1999). *An Introduction to Linguistic Theory and Language Acquisition*. Malden: Blackwell Publishing.
- Cruse, Alan (2000). *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.
- Darwin, Charles (1969[1871]). *El origen del hombre y la selección natural en relación al sexo*. México DF: Diana.
- Davidson, Donald (2001a). Actions, Reasons, and Causes. En *Essays on Actions and Events* (pp. 3-20). Segunda edición. Oxford: Oxford University Press.
- Davidson, Donald (2001b). Mental Events. En *Essays on Actions and Events* (pp. 207-224). Segunda edición. Oxford: Oxford University Press.
- De Saussure, Ferdinand (1974[1916]). *Curso de lingüística general*. Decimotercera edición. Buenos Aires: Losada.
- De Vega, Manuel & Fernando Cuetos (eds.) (1999). *Psicolingüística del español*. Madrid: Trotta.
- De Villiers, Jill G. & Peter A. de Villiers (1985). The Acquisition of English. En Dan Isaac Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*. Volumen I: *The Data* (pp. 27-140). Hillsdale: Lawrence Erlbaum Associates.
- Deacon, Terrence (1992). Impressions of Ancestral Brains. En Steve Jones, Robert Martin y David Pilbeam (eds.), *The Cambridge Encyclopedia of Human Evolution* (pp. 116-117). Cambridge: Cambridge University Press.
- Delson, Eric, Ian Tattersall, John A. van Couvering & Alison Brooks (eds.) (2000). *Encyclopedia of Human Evolution and Prehistory*. Segunda edición. Nueva York-Londres: Garland.
- Deutscher, Guy (2010). Does Your Language Shape what You Think? *The New York Times*. <http://www.nytimes.com/2010/08/29/magazine/29language-t.html?page-wanted=2&r=1>
- Dunbar, Robin (1988). *Primate Social Systems*. Londres: Chapman & Hall.
- Dunbar, Robin (1996). *Grooming, Gossip and the Evolution of Language*. Londres: Faber & Faber.
- Dunbar, Robin (1998). The Social Brain Hypothesis. *Evolutionary Anthropology*, 6, 178-190.

- Dunbar, Robin (2000). On the Origin of the Human Mind. En Peter Carruthers (ed.), *Evolution and the Human Mind. Modularity, Language and Meta-Cognition* (pp. 238-239). Cambridge: Cambridge University Press.
- Dunbar, Robin & Louise Barrett (2007). Evolutionary Psychology in the Round. En Robin Dunbar y Louise Barrett (eds.), *The Oxford Handbook of Evolutionary Psychology* (pp. 3-10). Nueva York: Oxford University Press.
- Eimas, Peter D. (1974). Auditory and Linguistic Processing of Cues for Place of Articulation by Infants. *Perception Psychophysiology*, 16, 513-521.
- Eimas, Peter D. (1975). Auditory and Phonetic Coding of Cues for Speech: Discrimination of the [r-l] Distinction by Young Infants. *Perception & Psychophysics*, 18(5), 341-347.
- Eimas, Peter D., Einer R. Siqueland, Peter Jusczyk & James Vigorito (1971). Speech Perception in Infants. *Science*, 171(3968), 303-306.
- El-Hani, Charbel & Antonio M. Pereira (1999). Understanding Biological Causation. En Valerie G. Hardcastle (ed.), *Where Biology Meets Psychology* (pp. 333-356). Massachusetts: The MIT Press.
- Ellis, Nick C. (1998). Emergentism, Connectionism and Language Learning. *Language Learning*, 48(4), 631-664.
- Elman, Jeffrey L. (2001). Connectionism and Language Acquisition». En Michael Tomasello y Elizabeth Bates (editores), *Language Development. The Essential Readings* (pp. 295-306). Oxford: Blackwell.
- Elman, Jeffrey L. (2005). Connectionist Models of Cognitive Development: Where Next? *Trends in Cognitive Sciences*, 9(3), 111-117.
- Elman, Jeffrey L. y otros (1996). *Rethinking Innateness. A Connectionist Perspective on Development*. Cambridge: The MIT Press.
- Escandell, María Victoria (1993). *Introducción a la pragmática*. Barcelona: Antropos.
- Fenson, Larry y otros (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. San Diego: Singular.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky (2005). The Evolution of the Language Faculty: Clarifications and Implications. *Cognition*, 97(2), 179-210.
- Fitch, W. Tecumseh (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.
- Flinn, Mark & Carol Ward (2005). Ontogeny and Evolution of the Social Child. En Bruce Ellis y David Bjorklund (eds.), *The Origins of the Social Mind. Evolutionary Psychology and Child Development* (pp. 19-44). Nueva York: Guilford.
- Foley, William A. (1997). *Anthropological Linguistics: An Introduction*. Oxford: Blackwell Scientific Publications.

- Fonagy, Peter & Mary Target (1997). Attachment and Reflective Function: Their Role in Self-Organization. *Development and Psychopathology*, 9(4), 679-700.
- Fonagy, Peter (2008). The Mentalization Focused Approach to Social Development. En Fredric Busch (ed.), *Mentalization. Theoretical Considerations, Research Findings and Clinical Implications* (pp. 3-56). Nueva York: The Analytic Press.
- Gergely, György (2007). The Social Construction of the Subjective Self: the Role of Affect Mirroring, Markedness, and Ostensive Communication in Self Development. En Linda Mayes, Peter Fonagy y Mary Target (eds.), *Developmental Science and Psychoanalysis. Integration and Innovation* (pp. 45-82). Londres: Karnac Books.
- Gilbert, Daniel (1999). Social Cognition. En Robert Wilson y Frank Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 777-778). Cambridge: The MIT Press.
- Goldman, Alvin (2006). *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. Nueva York: Oxford University Press.
- Goody, Esther (1997). Social Intelligence and Language: Another Rubicon? En Andrew Whiten y Richard Byrne (eds.), *Machiavellian Intelligence II. Extensions and Evaluations* (pp. 365-396). Cambridge: Cambridge University Press.
- Gould, Steven J. & Elisabeth Vrba (1982). Exaptation—A Missing Term in the Science of Form. *Paleobiology*, 8(1), 4-15.
- Gould, Steven J. & Richard C. Lewontin (1984[1979]). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. En Elliot Sober (ed.), *Conceptual issues in Evolutionary Biology* (pp. 252-270). Cambridge: Bradford Books-MIT.
- Greenberg, Joseph (1987). *Language in the Americas*. Stanford: Stanford University Press.
- Hamilton, David L. (2005). Social Cognition: An Introductory Overview. En David Hamilton (ed.), *Social Cognition* (pp. 1-26). Nueva York: Psychology Press.
- Harley, Trevor (2008). *The Psychology of Language. From Data to Theory*. Tercera edición. Nueva York: Taylor and Francis Group, Psychology Press.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598), 1569-1579.
- Heritage, John (2001). Ethno-Sciences and Their Significance for Conversation Linguistics. En Klaus Brinker, Gerd Antos, Wolfgang Heinemann y Sven F. Sager (eds.), *Text- und Gesprächslinguistik. Linguistics of Text and Conversation* (pp. 908-919). Berlín: De Gruyter.
- Hernández, Carlos & David Bjorklund (2008). When Development Matters: From Evolutionary Psychology to Evolutionary Developmental Psychology. *Anuario de Psicología*, 39(2), 177-192.

- Herrera, Marcos (2002). ¿Al pan, 'pan' y al vino, 'vino'? Un enfoque representacionista no esencialista del significado de las palabras. *Lexis*, XXVI(2), 345-393.
- Hornstein, Norbert (2009). *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.
- Horst, Steven (1999). Computational Theory of Mind. En Robert A. Wilson y Frank C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 170-171). Cambridge: The MIT Press.
- Jackendoff, Ray (1999). Possible Stages in the Evolution of the Language Capacity. *Trends in Cognitive Sciences*, 3(7), 272-279.
- Jackendoff, Ray (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, Ray & Steven Pinker (2005). The Nature of the Language Faculty and its Implications for Evolution of Language (Reply to Fitch, Hauser and Chomsky). *Cognition*, 97, 211-225.
- Jackendoff, Ray (2010). Your Theory of Language Evolution Depends on Your Theory of Language. En Richard Larson, Viviane Déprez y Hiroko Yamakido (eds.), *The Evolution of Human Language. Biolinguistic Perspectives* (pp. 63-72). Cambridge: Cambridge University Press.
- Jakobson, Roman (1984). *Ensayos de lingüística general*. Pensilvania: Ariel.
- Jakobson, Roman (1968). *Child Language, Aphasia and Phonological Universals*. La Haya: Mouton.
- Jellema, Tjeerd & David Perrett (2007). Neural Pathways of Social Cognition. En Robin Dunbar y Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 163-178). Nueva York: Oxford University Press.
- Jones, Steve, Robert Martin & David Pilbeam (eds.) (1992). *The Cambridge Encyclopedia of Human Evolution*. Cambridge: Cambridge University Press.
- Karmiloff-Smith, Annette (1996) *Beyond Modularity. A Developmental Perspective on Cognitive Science*. Cambridge: The MIT Press.
- Kirby, Simon (2007). The Evolution of Language. En Robin Dunbar y Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 669-681). Nueva York: Oxford University Press.
- Küntay, Aylin & Dan I. Slobin (1996). Listening to a Turkish Mother: Some Puzzles for Acquisition. En Dan Isaac Slobin, Julie Gerhardt, Amy Kyratzis y Jiansheng Guo (eds.), *Social Interaction, Social Context, and Language: Essays in Honor of Susan Ervin-Tripp* (pp. 265-286). Hillsdale: Erlbaum.

- Laitman, Jeffrey T. (2000). Origins of Speech. En Eric Delson, Ian Tattersall, John A. van Couvering y Alison Brooks (eds.), *Encyclopedia of Human Evolution and Prehistory* (pp. 1358-1362). Segunda edición. Nueva York-Londres: Garland.
- Larson, Richard K., Vivianne Déprez & Hiroko Yamakido (2010). *The Evolution of Human Language. Bilingual Perspectives*. Cambridge: Cambridge University Press.
- Lasnik, Howard & Juan Uriagereka (2007). *Structure Dependence, the Rational Learner, and Putnam's «Sane Person»* [charla]. En *Where Does Syntax Come From? Have We All Been Wrong?* [taller del MIT], Cambridge, octubre.
- Lenneberg, Eric H. (1967). *Biological foundations of language*. Nueva York: Wiley.
- Lenneberg, Eric H. (1975). *Fundamentos biológicos del lenguaje*. Madrid: Alianza.
- Levinson, Stephen C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Lewin, Roger (1999). *Human Evolution. An Illustrated Introduction*. Cuarta edición. Malden: Blackwell Science.
- Lieberman, Philip (2007). The Evolution of Human Speech. Its Anatomical and Neural Bases. *Current Anthropology*, 48(1), 39-66.
- Lightfoot, David (1989). The Child's Trigger Experience: Degree-0 Learnability. *Behavioral and Brain Sciences*, 12(2), 321-334.
- Locke, John L. & Dawn M. Pearson (1992). Vocal Learning and the Emergence of Phonological Capacity. A Neurobiological Approach. En Charles Albert Ferguson, Lise Menn y Carol Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications* (pp. 91-129). Timonium: York Press.
- Ludlow, Peter (2011). *The Philosophy of Generative Linguistics*. Oxford: Oxford University Press.
- Menzies, Peter (2010). Reasons and Causes Revisited. En Mario de Caro y David Macarthur (eds.), *Normativity and Naturalism* (pp. 142-170). Nueva York: Columbia University Press.
- McClelland, James L. (1999). Cognitive Modeling, Connectionist. En Robert A. Wilson y Frank C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 137-139). Cambridge: The MIT Press.
- Nelson, Katherine (1973). Structure and Strategy in Learning to Talk. *Monographs of the Society for Research in Child Development*, 38(1-2) 1-135.
- Newmeyer, Frederick J. (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.
- Olarrea, Antxon (2007). *Orígenes del lenguaje y selección natural*. Madrid: Sirius.
- Oller, D. Kimbrough, Leslie A. Wieman, William J. Doyle & Carol Ross (1976). Infant Babbling and Speech. *Journal of Child Language*, 3(1), 1-11.

- Panksepp, Jaak (2007). The Neuroevolutionary and Neuroaffective Psychobiology of the Prosocial Brain. En Robin Dunbar y Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 145-162). Nueva York: Oxford University Press.
- Piattelli-Palmarini, Massimo (1989). Evolution, Selection and Cognition: From «Learning» to Parameter Setting in Biology and in the Study of Language. *Cognition*, 31(1), 1-44.
- Pinker, Steven (1994). *The Language Instinct*. Nueva York: Harper Collins.
- Pinker, Steven & Paul Bloom (1990). Natural Language and Natural Selection. *Behavioral and Brain Sciences*, 13(4), 707-784.
- Pinker, Steven & Ray Jackendoff (2005). The Faculty of Language: What's Special About It? *Cognition*, 95(2), 201-236.
- Premack, David & Guy Woodruff (1978). Does the Chimpanzee Have a Theory of Mind? *Behavioral and Brain Sciences*, 1(4), 515-526.
- Premack, David (1988). Does the Chimpanzee Have a Theory of Mind? Revisited. En Richard W. Byrne y Andrew Whiten (eds.), *Machiavellian Intelligence. Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans* (pp. 160-179). Oxford: Clarendon Press.
- Pylyshyn, Zenon (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge: The MIT Press.
- Quintanilla, Pablo (2009). La evolución de la mente y el comportamiento moral. *Acta biológica colombiana. Número especial: Darwin 200 años: evolución, diversificación y ramificación permanente* 14(S), 425-440.
- Rees, Norma S. (2011). The Role of Babbling in the Child's Acquisition of Language. *International Journal of Language & Communication Disorders*, 7(1), 17-23.
- Renkema, Jan (1993). *Discourse Studies. An Introductory Textbook*. Ámsterdam-Filadelfia: John Benjamins.
- Rizzolatti, Giacomo & Leonardo Fogassi (2007). Mirror Neurons and Social Cognition. En Robin Dunbar y Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 179-196). Nueva York: Oxford University Press.
- Ruhlen, Merritt (1994). *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.
- Rumelhart, David E. (1989). The Architecture of Mind: A Connectionist Approach. En Michael I. Posner (ed.), *Foundations of Cognitive Science* (pp. 133-159). Cambridge: The MIT Press.
- Rumelhart, David E. & James L. McClelland (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: The MIT Press.
- Saeed, John I. (1997). *Semantics*. Oxford: Blackwell.

- Scollon, Ron (1976). *Conversations with a One Year Old*. Honolulu: Universidad de Hawái.
- Siegel, Daniel (1999). *La mente en desarrollo. Cómo interactúan las relaciones y el cerebro para moldear nuestro ser*. Bilbao: Desclée de Brouwer.
- Skinner, Burrhus F. (1957). *Verbal Behavior*. Nueva York: Appleton-Century-Crofts.
- Smith, Neil (1989). *The Twitter Machine: Reflections on Language*. Oxford: Blackwell.
- Snyder, William (2007). *Child Language: The Parametric Approach*. Oxford: Oxford University Press.
- Stassen, Kathleen (2008). *The Developing Person Through the Life Span*. Nueva York: Worth Publishers.
- Sterelny, Kim (1999). Situated Agency and the Descent of Desire. En Valerie G. Hardcastle (ed.), *Where Biology Meets Psychology* (pp. 203-219). Cambridge: The MIT Press.
- Thompson, Ross (2006). The Development of the Person: Social Understanding, Relationships, Conscience, Self. En Nancy Eisenberg (ed.), *The Handbook of Child Psychology*. Volumen 3: *Social, Emotional and Personality Development* (pp. 26-41). Nueva Jersey: John Wiley and Sons.
- Tomasello, Michael (1999). *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.
- Tomasello, Michael (2001). Bruner on Language Acquisition. En David Backhurst y Stuart Shanker (eds.), *Jerome Bruner: Language, Culture and Self* (pp. 31-48). Londres: Sage Publications.
- Tomasello, Michael (2003). *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Tomasello, Michael (2008). *Origins of Human Communication*. Cambridge: The MIT Press.
- Tomasello, Michael (2009). The Usage-Based Theory of Language Acquisition. En Edith L. Bavin (ed.), *The Cambridge Handbook of Child Language* (pp. 69-88). Nueva York: Cambridge University Press.
- Tomasello, Michael & Elizabeth Bates (2001). *Language Development. The Essential Readings*. Oxford: Blackwell Publishers.
- Whiten, Andrew (1997). The Machiavellian Mind Reader. En Andrew Whiten y Richard W. Byrne (eds.), *Machiavellian Intelligence II. Extensions and Evaluations* (pp. 144-173). Cambridge: Cambridge University Press.
- Whorf, Benjamin. L. (1971). *Lenguaje, pensamiento y realidad*. Barcelona: Seix Barral.
- Wimmer, Heinz & Josef Perner (1983). Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. *Cognition*, 13(1), 103-128.

PARTE II
ENSAYOS SOBRE LA COGNICIÓN SOCIAL Y EL LENGUAJE

SECCIÓN I
ENSAYOS SOBRE LA EVOLUCIÓN DE LA COGNICIÓN SOCIAL

THE NONLINEAR EVOLUTION OF HUMAN COGNITION

H. Clark Barrett

University of California - Los Angeles (UCLA)

Humans are special. We tell ourselves this all the time, and maybe we are right. Over the last several hundred thousand years we have populated the globe, mastering both our physical environments and our social ones. It is widely held that these achievements have something to do with how smart we are, and in particular, with our *social* intelligence. Compared to other primates, humans appear to be social geniuses: we are able to live peacefully, and even cooperate, in groups both large and small. We are experts at reading each other's thoughts, feelings, intentions, and desires. We transmit information using elaborate systems of symbols, leveraging the accumulated knowledge of those who have gone before us to transform our landscapes, domesticate our food, govern ourselves, and land on the moon.

All of this cries out for explanation, and plenty of explanations have been offered. Just as scientists in the early twentieth century searched for the «missing link» that separated humans from other apes, scientists in the twenty-first century have searched for the evolutionary «it factor» —the psychological skill or capacity that separates human cognition from that of our fellow primates. Is it social learning? Symbols? Concepts? Language? Theory of mind? Prosociality? Morality? Neural plasticity? Bigger brains? Insert your favorite candidate here.

All of these, of course, are things that humans have, make, or do —that is not in question. What is hotly debated in today's evolutionary literature is the question of which of these factors —or more likely, which combination of them— explains human psychological uniqueness. And since all evolutionists are materialists who accept that our minds are made of physical stuff —neurons, neurotransmitters, genes, and other biochemical ingredients— this question obliges us to ask what changes in brains have occurred in our evolutionary lineage to give rise to the new abilities that humans have. And here is where, despite decades if not centuries of accumulating data, there still seems to be no consensus in sight.

Why not? Part of the answer, to be sure, has to do with missing pieces of the puzzle: although we've accumulated lots of data about the brains and minds of humans and other primates, we don't yet have enough for a definitive and complete answer to the question of what makes them different. But a glance around the scholarly landscape suggests that this isn't all: our problems are also, almost certainly, theoretical. In this essay I would like to briefly consider one such theoretical problem: the problem of how to theorize psychological adaptations, and how they evolve in a mind composed of multiple, interacting adaptations. First I will describe what I think the problem is. Then I will explain how I think it is standing in the way of progress in understanding the evolution of the mind. Finally, I will close with some thoughts about how this problem can be solved.

1. THE PROBLEM

In 1983, philosopher Jerry Fodor published his landmark book *The Modularity of Mind*, in which he formalized the concept of a psychological «module»: a sort of cognitive «reflex», in his terms, which he contrasted with what he called «central» (flexible and domain-general) systems of thought. In some ways, Fodor's book revolutionized the field of cognitive psychology by helping to make explicit what is meant by a mental mechanism, and how we would go about studying one empirically. But in another way, the book has had a damaging consequence on the field by promulgating the widespread view that the mind is composed of two fundamentally different kinds of mechanisms: specialized mechanisms, or «modules», and unspecialized «central» systems. This is the view that I will call «psychological dualism» (Barrett, 2012). I will contrast it with what one might call a psychological *diversity* view: that the mind is composed of diverse adaptations, which in turn have diverse functions and take diverse forms, not just two. Importantly, these diverse forms can include properties that are thought to be characteristic of «central» systems —such as flexibility and interactivity— despite being specialized to carry out particular functions (Barrett & Kurzban, 2006).

Fodor originally proposed a set of nine features which he held to be characteristic of modules, and which distinguished them from central, non-modular systems: these included domain specificity, mandatory operation (automaticity), limited central access, speed, informational encapsulation, shallow outputs, fixed neural architecture, characteristic and specific breakdown patterns, and characteristic pace and sequencing of ontogeny (1983). At the time, he was clear that he didn't intend modularity to be a binary property but a «matter of degree». In the literature, however, this subtlety has been lost, and modules are now seen as uniformly possessing this

Fodorian package of traits. They are, in essence, like bits of psychological Lego that are rigid, reflex-like, automatic, innate, and independently operating. Non-modular systems, in contrast, have the opposite set of properties: they are flexible, interactive, and shaped by experience. This contrasting set of features defines a dualist taxonomy of mental mechanisms. In the «dual systems» literature, these are sometimes called *System 1* (modular) and *System 2* (non-modular), and System 1 is often explicitly associated with «specialized» psychological adaptations (Chiappe & Gardner, 2012; Kahneman, 2011; Stanovich, 2004).

There are many problems with psychological dualism, about which I have written in detail elsewhere (Barrett, 2005, 2006, 2012; Barrett & Kurzban, 2006, 2012). Here I would like to focus on one set of properties of putative «modules» which, if it were truly characteristic of psychological adaptations in general, would have extremely important implications for the study of mental evolution, and in particular, for at least some conceptualizations of the «it factor». I will refer to this set of properties as *autonomy*. Other suitable words might include *independence* and *isolability*. There are several ways in which «modules» or «adaptations» are widely held to be autonomous, independent, or isolable, all of which are probably, in a strict sense, wrong.

2. THE NON-AUTONOMY OF MENTAL EVOLUTION

What I mean by autonomy is a mix of two related components —*architectural* autonomy, implying that modules can and do operate independently of other systems and can be removed while leaving the operation of other systems intact— and *evolutionary* autonomy, implying that modules are shaped by evolution independently of other systems, such that their functions are relatively isolable.

In the evolutionary developmental literature, modularity is often seen as related to the question of «evolvability»: the degree to which evolutionary forces, including mutation, drift, and natural selection, can independently shape different aspects of organismal phenotype, as a function the «decomposability» of the phenotype into relatively independent parts or subunits (Wagner & Altenberg, 1996). Modeling work, as well as empirical studies of the modularity of existing development systems, has shown that the answers are neither obvious nor entirely intuitive (Wagner and others, 2007). Importantly, it is likely that no organismic traits are *entirely* autonomous, either architecturally or evolutionarily. Degrees and aspects of autonomy must and do play a role in shaping the course of evolution—but so do degrees and aspects of interactivity and non-independence. And perhaps most importantly for thinking about the mind, this doesn't mean that everything inside

the organism is one big mishmash: there is functional specialization and modularity, but the important evolutionary questions lie just as much if not more in how the specializations *interact*, as in their relative autonomy or independence.

There are several basic facts about how evolution occurs that make strict autonomy a poor diagnostic for psychological adaptations. The first is that new adaptations evolve from old ones. Therefore, «new» psychological design inherits much «old» design. This means that multiple psychological mechanisms might, indeed must, share aspects of their information-processing design because they inherited it from older mechanisms they both evolved from (Barrett, 2012). As a consequence, these mechanisms might not appear «autonomous» along dimensions of shared design, and indeed, would be empirically impossible to disentangle if investigated solely along those dimensions.

For example, it has been proposed that human minds might contain specialized subsystems for recognizing and processing different kinds of physical objects, including faces, artifacts, and places (Kanwisher, 2010). Critics of this proposal point to aspects of object processing that can be shared across object domains, such as «configural» processing, as evidence against such sub-specialization (Gauthier & Nelson, 2001). But if these subsystems evolved from a shared common ancestor and are now to some degree phenotypically distinct, configural processing might be a characteristic of multiple object-processing systems that are, in fact, functionally and phenotypically distinct. Other properties widely shared across mental mechanisms might include, for example, statistical learning and Bayesian updating. These are widely regarded as diagnostic of «general purpose» systems, but there is no reason to think they could not be characteristic of specialized systems as well.

Second, there is what can be called evolutionary «feedback»: changes in one system can lead to changes in another, because of how these systems interact to impact fitness. Again, this might lead to apparent signatures of non-autonomy, even in what might be phenotypically distinct psychological specializations (Barrett, 2012). For example, the evolution of language may have involved evolutionary changes in multiple areas of the brain, including areas involved in language production, language comprehension, and the white matter pathways connecting them (Buxhoeveden and others, 2001; Friederici, 2009; Hickok & Poeppel, 2007). Clearly this speaks against autonomy in the sense that changes in these distinct areas are not evolutionarily independent, but does it does not necessarily speak against modularity, or psychological specialization, if it turns out that areas involved in language production, language comprehension, and communication between them each carry out different functions.

Thus, even when selection has favored what from one point of view might seem to be a single ability, such as language, multiple systems can change as a consequence. This can include systems that execute «subfunctions» of the ability, such as language production and comprehension, and also systems causally related to the ability in question, such as, for example, mechanisms of social learning whose evolution might have altered, and been altered by, changes in language abilities (see below). In such causal webs, the «it factor» can't necessarily be equated with a single module—but neither is it the result solely of changes to a single «general-purpose» system. The answer is more complicated. When evolution is nonlinear, such that changes within a complex system of interacting parts cause other, non-additive changes, the search for an «it factor» may lead to just the kind of situation we see today: different parties backing their favorite candidate, and all partially right.

3. TOWARDS A SOLUTION

Faced with the interactive complexity of the mind, some psychologists are tempted by a seemingly simple solution: let's toss out ideas like specialization and adaptation altogether and model the mind as a big dynamic system that adjusts itself in light of experience.

Of course, the mind *is* a big dynamic system that adjusts itself in light of experience (just as most biological systems are). But that doesn't mean it isn't composed of specializations shaped by the evolutionary process. Instead, it means that whatever those adaptations are, they have been shaped by how they *interact*, both with the world and with each other. This is at odds with the Fodorian view of a module as entirely dissociable. Removing a specialized component might have profound effects on the rest of the system, but that doesn't speak against specialization (imagine, for example, that you removed a component in your car engine, and the car failed to start; would you conclude that the part wasn't specialized?). It is also at odds with the dualist view of the mind as composed of two kinds of things: the rigid, inflexible, autonomous adaptations in System 1, and the flexible, interactive, undifferentiated System 2.

What if we were to stop thinking about adaptations as rigid, autonomous, reflex-like psychological Lego, and to start thinking of them as neural information-processing mechanisms that evolve through descent with modification from older systems, and that are shaped by how they interact with other systems and with the external world? On its surface this proposal would appear uncontroversial, and yet it is clearly not yet taken seriously by the psychology community because if it were, it would have several important effects on how we study psychological adaptations.

First, Fodor's list of modular properties would stop being used as a kind of diagnostic checklist for psychological adaptations. For example, techniques designed to activate System 2 mechanisms, such as placing subjects under «cognitive load», are sometimes taken to demonstrate that a particular process can't be due to a specialized adaptation, because if it were (according to dualist logic), the operation of that adaptation would be «automatic» —autonomous from other systems— and therefore not influenceable by putative System 2 processes like working memory (De Steno and others, 2002; for a critique see Barrett and others, 2006). On the view proposed here, properties like automaticity and speed are not *necessary* features of any psychological adaptation, but instead depend on the evolutionary history of the putative mechanism and what it has been selected to do. For example, we might expect many specialized adaptations in the mind, such as adaptations for learning and long-term decision-making (including partner choice), to be relatively slow in their operation, to use (potentially) lots of information in their operations, and to be influenced by the operation of other systems.

Second, a non-dualist view of adaptations would drastically change the landscape of how cross-cultural and developmental evidence is interpreted. Currently, cross-cultural variation in a skill is widely seen as evidence against the hypothesis that the skill represents an evolved adaptation, and in favor of the hypothesis that it is acquired through general-purpose learning. Conversely, uniformity in development, and in particular early development, are seen as counting for a specialization view. Rather than having either one count for or against the operation of adaptations, it makes more sense to consider each adaptation in its own terms, and ask whether features like cross-cultural uniformity or early development are likely to be an aspect of its evolved *design*. For example, some adaptations, such as adaptations causing mate competition and sexual jealousy, might be expected to evolve late in childhood; others, such as psychological adaptations against stressful circumstances, might develop at different times or rates in different individuals or environments, and perhaps sometimes not at all —despite being evolved adaptations (Frankenhuis & Panchanathan, 2011).

Third, in the case of debates about human uniqueness, a more biologically realistic view of psychological adaptations would temper the search for a specific «it factor» in human psychological evolution. For reasons given above, it's highly likely that multiple mechanisms were modified in the course of human evolution. Even if it were the case that there is in some sense a «prime mover», like the evolution of culture, it is quite possible that this could exert selection on multiple aspects of human brains, and / or that it could beget an evolutionary cascade in which culture selects for other kinds of abilities, such as cooperation and morality

(Boyd & Richerson, 2006). And empirically, it might be difficult to isolate the causal contributions of multiple factors that have co-evolved in the human lineage. For example, the comparative method, perhaps the foremost technique in evolutionary biology for inferring evolutionary causation, cannot be used to make inferences about evolutionary changes that have occurred uniquely within a single taxon (that is, autapomorphies; Brooks & McLennan, 1991).

Finally, accepting that specialization and interactivity are simultaneously possible in brain evolution should lead us to reconsider what are often considered as mutually exclusive evolutionary scenarios. For example, increasing brain size is widely considered a mutually exclusive alternative to «new modules» in human evolution (Deacon, 1997; Donald, 2001). However, there are reasons to suspect that increasing brain size might be positively, not negatively, correlated with degree of modularity. At minimum, purely architectural considerations suggest that as brain volume increases, network connectivity between different regions within the volume must necessarily decrease —implying increased modularity in the sense that the term is used in network theory (Kaas, 2000; Streidter, 2005). This makes several rarely considered scenarios possible (Barrett, 2012). First, increasing brain size may have *led* to increasing modularity, with increasing functional specialization as a consequence. However, the opposite is also possible: namely, that brain size increased *because of* selection for increased modularity and specialization —one way computational systems are known to gain increased flexibility (Baldwin & Clark, 2000). These are, of course, empirical questions. But either way, changes in modularity and changes in brain size / «general-purpose» cognition no longer should be seen as necessarily exclusive alternatives for the evolution of uniquely human cognition.

4. THE FUTURE OF THE SEARCH FOR HUMAN UNIQUENESS

The view I have been advocating here is that we stop thinking about psychological adaptations in terms of properties like rigidity, autonomy, and innateness, and start thinking about them in terms of *design*. By «design», of course, I mean evolved design: how the adaptation has been shaped by the evolutionary process to interact with the rest of the system in which it is situated, including both the brain and the world. Part of evolutionary design thinking, then, involves theorizing how changes occur within systems of interacting parts —parts that can have distinct functions, but that do not necessarily evolve autonomously. This will require the development of new theoretical tools and the abandonment of outdated positions, but it is likely get us much closer to the biological reality of what has changed in brains since the chimp-human common ancestor than the «psychological Lego» view of adaptations

as purely autonomous and *sui generis*, as opposed to hierarchically evolving via descent with modification from other mechanisms (Barrett, 2012).

Let me conclude with a hypothetical example of how we might think about nonlinear evolutionary interactions within a specialized brain: in particular, *synergistic* interactions between parts that are not, themselves, entirely new, but alter each others' design through an evolutionary feedback process. The example I have in mind is the potential synergistic relationship between three abilities that have each been proposed as distinct «it factors» of human uniqueness: (1) theory of mind, or the ability to infer others' intentions, motivations, and beliefs; (2) cultural learning and transmission; and (3) cooperation. Nobody would claim that these three abilities are the *same* ability; in all likelihood, they involve at least partly distinct psychological mechanisms. And yet, it is quite plausible, if not likely, that synergistic interactions occur between each of the abilities within this triad, meaning that evolutionary alterations in one have led to, and / or have been caused by, alterations in the other.

For example, evolutionary models show that mechanisms stabilizing cooperation, such as group-level norms, can result from cultural evolution —meaning that the evolution of cultural transmission mechanisms can lead to enhanced cooperation, and can even select for psychological mechanisms stabilizing cooperation, such as conformity biases (Boyd & Richerson, 2009). Similarly, the fundamental problem in achieving cooperation is avoiding the free-rider problem, which can favor mechanisms for predicting whether or not a cooperation partner will cooperate or defect. Improved abilities to detect the intention or motivation to defect —or to cooperate— could therefore stabilize cooperation. Thus, the benefits of cooperation could have been a selective factor favoring the evolution of theory of mind (Barrett and others, 2010). This would help explain why much of human morality seems deeply intertwined with reasoning about the causes of others' actions: for example, whether the action was intentional or accidental, done knowingly or in ignorance, by commission or omission, and so on (Mikhail, 2007; Young & Saxe, 2009).

Finally, there appears to be a synergistic relationship between theory of mind and cultural transmission. While social learning is certainly possible in the absence of theory of mind, increasing evidence suggests that human cultural learning is greatly boosted by our capacity to go past the surface features of others' actions to infer the goals and intentions underlying what they are doing (Csibra & Gergely, 2009; Lyons and others, 2007; Meltzoff, 1995; Whiten and others, 2009). In contrast with other apes, humans, even infants, view pedagogical demonstrations as demonstrations not of arbitrary sets of actions, but of ways to achieve a particular goal (Csibra & Gergely, 2009). And in linguistic communication —arguably, one of the most important means of cultural transmission of humans— transfer of meaning is greatly aided

by our ability to use the surface content of speech acts to infer the speaker's intended meaning (Sperber & Wilson, 1995). Thus, theory of mind and cultural transmission could have coevolved synergistically, at least in part, in the human lineage.

Clearly, humans differ from other apes in each of these ways: we cooperate more, have better theory of mind, engage in much more cultural transmission, and use language to do so. None of these, then, is the «it factor» in the sense of being *the* ability that sets us apart from other apes. But is one of them the «prime mover?». Did one of them serve as the main selective factor that drove the evolution of the others? It's possible, but difficult to tell just from the fact that all three differ in humans. To find out the answer, we need to develop more sophisticated models that allow for the kinds of synergies described above. Some such models exist, but more will be needed—perhaps with greater detail, and exploring more parameters in the space of possible brain variation, including ones we haven't yet considered—before the question of human uniqueness can be settled.

Importantly, when interactions become nonlinear, as the interactions between theory of mind, cultural transmission, and cooperation probably are, then verbal arguments rapidly lose force, suggesting the need for formal models using the tools of evolutionary, developmental and cognitive modeling (McElreath & Boyd, 2007). Arguably, one of the reasons for the impasse that confronts us in debates about human cognition is that our dualist models of the mind are many orders of magnitude simpler than the phenomena we are invoking them to explain. Unlike the recipe for french fries, the recipe for brain evolution is likely to require far more than two ingredients. Hopefully our theories will become more complex, rather than less so, in order to rise to the challenge.

REFERENCES

- Baldwin, Carliss Young & Kim Clark (2000). *Design Rules: The Power of Modularity*. Volume I. Cambridge: The MIT Press.
- Barrett, Clark (2005). Enzymatic Computation and Cognitive Modularity. *Mind and Language*, 20(3), 259-287.
- Barrett, Clark (2006). Modularity and Design Reincarnation. In Peter Carruthers, Stephen Laurence and Stephen Stich (eds.), *The Innate Mind: Culture and Cognition* (pp. 199-271). New York: Oxford University Press.
- Barrett, Clark (2012). A Hierarchical Model of the Evolution of Human Brain Specializations. *Proceedings of the National Academy of Sciences*, 109(1), 10733-10740.
- Barrett, Clark, Leda Cosmides & John Tooby (2010). Coevolution of Cooperation, Causal Cognition and Mindreading. *Communicative & Integrative Biology*, 3(6), 522-524.

- Barrett, Clark, David A. Fredrick, Martie Haselton & Robert Kurzban (2006). Can Manipulations of Cognitive Load Be Used to Test Evolutionary Hypotheses? *Journal of Personality and Social Psychology*, 91(3), 513-518.
- Barrett, Clark & Robert Kurzban (2006). Modularity in Cognition: Framing the Debate. *Psychological Review*, 113(3), 628-647.
- Barrett, Clark & Robert Kurzban (2012). What Are the Functions of System 2 Modules? A Reply to Chiappe and Gardner. *Theory and Psychology*, 22(5), 683-688.
- Boyd, Robert & Peter Richerson (2006). Culture and the Evolution of the Human Social Instincts. In Stephen Levinson and Nicholas J. Enfield (eds.), *Roots of Human Sociality. Culture Cognition and Interaction* (pp. 453-477). Oxford: Berg.
- Boyd, Robert & Peter Richerson (2009). Culture and the Evolution of Human Cooperation. *Philosophical Transactions of the Royal Society Biological Sciences*, 364, 3281-3288.
- Brooks, Daniel & Deborah McLennan (1991). *Phylogeny, Ecology, and Behavior: A Research Program in Comparative Biology*. Chicago: University of Chicago Press.
- Buxhoeveden, Daniel, Andrew E. Switala, Mark Litaker, Emil Roy & Manuel F. Cassanova (2001). Lateralization of Minicolumns in Human Planum Temporale is Absent in Nonhuman Primate Cortex. *Brain, Behavior and Evolution*, 57(6), 349-358.
- Chiappe, Dan & Ricca Gardner (2012). The Modularity Debate in Evolutionary Psychology. *Theory & Psychology*, 22(5), 669-682.
- Csibra, Gergerly & György Gergely (2009). Natural Pedagogy. *Trends in Cognitive Sciences*, 13(4), 148-153.
- Deacon, Terrence (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: Norton & Company.
- De Steno, David, Monica Y. Barlett, Julia Braveman & Peter Salovey (2002). Sex Differences in Jealousy: Evolutionary Mechanism or Artifact of Measurement? *Journal of Personality and Social Psychology*, 83(5), 1103-1116.
- Donald, Merlin (2001). *A Mind so Rare: The Evolution of Human Consciousness*. New York: Norton & Company.
- Fodor, Jerry (1983). *The Modularity of Mind*. Cambridge: The MIT Press.
- Frankenhuis, Willem & Karthik Panchanathan (2011). Balancing Sampling and Specialization: An Adaptationist Model of Incremental Development. *Proceedings of the Royal Society B: Biological Sciences*, 278(1724), 3558-3565.
- Friederici, Angela (2009). Pathways to Language: Fiber Tracts in the Human Brain. *Trends in Cognitive Sciences*, 13(3), 175-181.
- Gauthier, Isabel & Charles Nelson (2001). The Development of Face Expertise. *Current Opinion in Neurobiology*, 11(2), 219-224.

- Hickok, Gregory & David Poeppel (2007). The Cortical Organization of Speech Processing. *Nature Reviews Neuroscience*, 8(5), 93-402.
- Kaas, Jon (2000). Why is Brain Size so Important: Design Problems and Solutions as Neocortex Gets Bigger or Smaller. *Brain and Mind*, 1(1), 7-23.
- Kahneman, Daniel (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus & Giroux.
- Kanwisher, Nancy (2010). Functional Specificity in the Human Brain: A Window into the Functional Architecture of the Mind. *Proceedings of the National Academy of Sciences*, 107(25), 11163-11170.
- Lyons, Derek, Andrew Young & Frank Keil (2007). The Hidden Structure of Overimitation. *Proceedings of the National Academy of Sciences*, 104(25), 19751-19756.
- McElreath, Richard & Robert Boyd (2007). *Mathematical Models of Social Evolution: A Guide for the Perplexed*. Chicago: University of Chicago Press.
- Meltzoff, Andrew (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-old Children. *Developmental Psychology*, 31(5), 838-850.
- Mikhail, John (2007). Universal Moral Grammar: Theory, Evidence and the Future. *Trends in Cognitive Sciences*, 11(4), 143-152.
- Sperber, Dan & Deirdre Wilson (1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Stanovich, Keith (2004). *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. Chicago: University of Chicago Press.
- Striedter, Georg (2005). *Principles of Brain Evolution*. Sunderland: Sinauer Associates.
- Wagner, Gunter & Lee Altenberg. (1996). Complex Adaptations and the Evolution of Evolvability. *Evolution*, 50(3), 967-976.
- Wagner, Gunter, Mihaela Pavlicev & James Cheverud (2007). The Road to Modularity. *Nature Reviews Genetics*, 8(12), 921-931.
- Whiten, Andrew, Nicola McGuigan, Sarah Marshall-Pescini & Lydia Hopper (2009). Emulation, Imitation, Over-Imitation and the Scope of Culture for Child and Chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(7), 2417-2428.
- Young, Lianne & Rebecca Saxe (2009). An fMRI Investigation of Spontaneous Mental State Inference for Moral Judgment. *Journal of Cognitive Neuroscience*, 21(364), 1396-1405.

THE FRAGMENTATION OF REASONING

Peter Carruthers

University of Maryland

1. INTRODUCTION

Scientists who study human reasoning across a range of cognitive domains have increasingly converged on the idea that there are two distinct systems (or types of system) involved. These domains include learning (Berry & Dienes, 1993; Reber, 1993), conditional and probabilistic reasoning (Evans & Over, 1996; Sloman, 1996, 2002; Stanovich, 1999), decision making (Kahneman & Frederick, 2002; Kahneman, 2003), and social cognition of various sorts (Petty & Cacioppo, 1986; Chaiken and others, 1989; Wilson and others, 2000). Although terminology has differed, many now use the labels *System 1* and *System 2* to mark the intended distinction. System 1 is supposed to be fast and unconscious in its operations, issuing in intuitively compelling answers to learning or reasoning problems in ways that subjects themselves have no access to. System 2, in contrast, is supposed to be slow and conscious in its operations, and is engaged whenever we are induced to tackle reasoning tasks in a reflective manner. Many theorists now accept that System 1 is really a *set* of systems, arranged in parallel, while believing that System 2 is a single serially-operating ability.

Moreover, System 1 is generally thought to be unchangeable in its basic operations, to be universal amongst humans, and to be shared (at least in significant part) with other species of animal. It cannot be directly influenced by verbal instruction, and it operates independently of the subject's explicit beliefs. In addition, its operations are thought to be either associative or heuristic in character (or both), epitomized in the phrase «quick and dirty». System 2, in contrast, is thought to be uniquely human, to be malleable, and to differ significantly between people, varying both by individual and by culture. It can be influenced by verbal instruction and can be guided, at least in part, by the subject's beliefs. Furthermore, System 2 is supposed to embody some sort of normatively correct competence (at least to some significant degree, and

presumably following appropriate training). The properties generally attributed to the two systems can be seen summarized in figure 1¹.

Figure 1. The properties of the two proposed systems

System 1	System 2
Unconscious / intuitive	Conscious / reflective
Fast	Slow
Parallel, high capacity	Serial, limited capacity
Automatic	Controlled
A set of systems	A single system
Not easily altered	Malleable
Universal amongst humans	Variable (by culture and by individual)
Mostly shared with other animals	Uniquely human
Impervious to verbal instruction	Responsive to verbal instruction
Independent of normative beliefs	Influenced by normative beliefs
Associative and/or heuristic-based	Rule-based and/or embodying normative competence

I shall argue that there is, indeed, a real, scientifically valid, distinction between a set of intuitive, unconsciously operating reasoning systems, on the one hand, and a reflective system whose operations are partly conscious, on the other. But I shall argue that this division fails to line up with many of the other properties generally associated with Systems 1 and 2. In particular, some intuitive systems can be slow, some can be controlled, and some can approach the highest normative standards; so the moniker «quick and dirty» is certainly inappropriate when applied to intuitive reasoning as such. Nor is it true that reflective reasoning always leads to improvement. On the contrary, in some contexts reflection leads to *worse* performance, and there are some tasks where reliance on intuitive reasoning is best. Moreover, reflective reasoning can also employ heuristics. Indeed, although there is (I shall argue) a fixed architecture underlying our reflective capacities, the latter can employ a hodge-podge of different procedures and abilities. In consequence, then, the distinction between System 1 and System 2 processes is not scientifically valid, failing to mark out any natural division within the mind.

¹ The properties listed in figure 1 have been culled from many different sources. Perhaps only a few researchers would accept all of them. But all (until very recently) would accept most. (Some fresh proposals for characterizing the System 1 / System 2 distinction will be considered in Section 4)

Here is how the paper will proceed. Section 2 will outline and motivate an architecture for reflective reasoning, for the most part following Carruthers (2006). On this account, reflective reasoning is real, but it is largely realized in the cyclical operations of unconscious, intuitive, processes (including the set of systems normally regarded as belonging to System 1), rather than existing alongside of the latter. I shall explain how we have independent evidence of the various components of the reflective system, and how most of the components either already exist in non-human animals or are developments of precursors that exist in animals. I shall then go on, in Section 3, to argue that despite the reality of the distinction between intuitive and reflective reasoning processes, this does *not* amount to a vindication of anything resembling the full System 1 / System 2 distinction as outlined above. On the contrary, I shall show that the latter distinction fails to mark out anything that deserves to be regarded as a *natural kind*. Finally, in Section 4, I shall compare the thesis defended in this article with some other recent proposals by leading figures in the field for reconceptualizing the System 1 / System 2 distinction, put forward by Evans (2009), Frankish (2009) and Stanovich (2009), in particular.

2. AN ARCHITECTURE FOR REFLECTIVE REASONING

This section provides an explanation of the distinction between intuitive and reflective processes, and shows how the latter can be partly realized in cyclical operations of the former.

2.1. The puzzle of reflective reason

We begin with some *prima facie* puzzles about the postulated reflective reasoning system. The latter is presumably meant to be a pan-human capacity of some kind; and yet at the same time its operations are dependent upon learning and vary a great deal by culture and individual. How can this be? One possible answer is that there is an innately channeled reasoning system of some sort which is then molded and completed on the basis of experience. Compare the human motor system. This, too, is pan-human in its architecture, and presumably contains a significant amount of innate structure (think of sneezing and yawning, for the clearest examples, but innately channeled behaviors probably also include walking and talking). Yet it is also built to be a learning system, in which new motor modules are constructed each time someone acquires a novel skill (Ghahramani & Wolpert, 1997; Manoel and others, 2002). Hence the mature state of the motor system will likewise vary greatly by culture and by individual, depending on the skills that the people in question possess.

What is somewhat more problematic is to understand how the reflective reasoning system can be guided by verbal instruction, and how it can be partly dependent upon people's *beliefs* about how they should reason. For the system in question is a cognitive one, issuing in transitions between one set of belief states and another, or generating decisions from a given set of beliefs and goals. And these cognitive transitions are presumably realized in some computational process or other. But how can verbal instruction and normative beliefs have the power to reach into the «innards» of this system, altering the state-transitions that are employed? How can a subject's beliefs have the capacity to re-write the computational algorithms that are used? Since verbal instructions and normative beliefs have their paradigmatic effects on *action*, it looks as if the reflective reasoning system must somehow be action-dependent; and it must likewise be capable of being intentionally controlled. Indeed, this is what I shall shortly argue.

The deepest difficulty for anyone wishing to defend the reality of a distinction between intuitive and reflective reasoning processes, however, is to understand how the two are related to one another, especially when viewed from an evolutionary perspective. For each seems to replicate, to a significant degree, the functionality of the other, and is concerned with the same types of task. Each issues in new beliefs on the basis of evidence, and each issues in decisions about what to do when faced with choices. How can this be? How could the (relatively short) period of hominid evolution have issued in such wholesale change and duplication? And if there were pressures for improvements in reasoning and decision-making (such as an increased need for cognitive flexibility), then why didn't they lead to modifications in the existing systems, rather than to the construction of an entirely different sort of system *de novo*? For that is the way in which evolutionary processes generally work.

How one answers these questions will obviously depend quite a bit on what one thinks about the character of non-human cognitive processes. For the more the latter resemble distinctively-human forms of thinking and reasoning, the deeper the problem. And in contrast, if one thought that all non-human animal behavior could be explained as resulting from various forms of associative conditioning, then the puzzle would largely disappear. For the benefits of rule-governed, systematically structured, forms of thinking and reasoning are obvious, especially in respect of flexibility and opportunities for one-shot learning. And it might be argued that these evolved in the hominid lineage subsequent to the evolution of the human language faculty, which provided the sort of structured representational system necessary for rule-governed reasoning to make its appearance (Bickerton, 2009). These questions will be addressed in Section 2.2.

2.2. Unreflective reasoning in non-human animals

Contrary to the proposed explanation of the evolutionary need for System 2 reasoning, cognitive processes involving one-shot learning and structured representational states are *rife* within the animal kingdom. Bees, for example, can extract information about the direction and distance of a food source from a single presentation of a waggle dance. And experiments involving bee navigation demonstrate that they use both vector information (deriving either from their own dead reckoning computations or the dances of other bees) and landmark information gained during their own exploratory flights from the nest before they begin their lives as foragers (Menzel and others, 2005; De Marco & Menzel, 2008). They can use this information flexibly in the service of multiple goals (searching for nectar, pollen, water, etcetera, returning to the hive, or dancing for other bees), suggesting very strongly that their decision making involves computations over structured representational states (Carruthers, 2006).

Moreover, Gallistel has demonstrated that conditioning behavior itself is best explained in rule-governed computational terms, rather than in terms of associative strengths (Gallistel & Gibbon, 2001; Gallistel & King, 2009). He points out there are many well-known conditioning phenomena that are extremely puzzling from an associationist perspective, but that fall out quite naturally from a computational account. To give just a single example: the number of reinforcements that are necessary for an animal to acquire an intended behavior is unaffected by mixing *unreinforced* trials into the learning process. One set of animals might be trained on a 1:1 schedule: these animals receive a reward every time that they respond when the stimulus is present. But another set of animals might be trained on a 10:1 schedule: here the animals only receive a reward once in every ten trials that they respond when the stimulus is present. Still it will, on average, take both sets of animals the same number of rewarded trials to acquire the behavior. It will take the second set of animals *longer* to acquire the behavior, of course. If it takes both sets of animals 40 rewarded trials to acquire the behavior, then the first set might learn it in 40 trials, whereas the second set will take 400. But the number of *reinforcements* to acquisition is the same. This is extremely puzzling from the standpoint of an associationist. One would expect that all those times when the stimulus *isn't* paired with a reward ought to weaken the association between stimulus and reward, and hence make learning the intended behavior harder. But it doesn't, just as Gallistel's computational model predicts.

Indeed, Gallistel and others (2001) demonstrate that animals in conditioning experiments who are required to respond to randomly changing rates of reward are able to track changes in the rate of reward about as closely as it is theoretically possible to do. Thus both pigeons and rats on a variable reward schedule from two different alcoves will match their behavior to the changing rates of reward.

There is a lever in each alcove, each set on a random reward schedule of a given probability. But these probabilities themselves change at random intervals. It turns out that the animals respond to these changes *very* rapidly, closely tracking the random variations in the immediately preceding rates. They certainly aren't averaging over previous reinforcements, as associationist models would predict. On the contrary, the animals' performance comes very close to that of an ideal Bayesian reasoner. And the only model that can predict the animals' behavior is one that assumes that they are capable of calculating the ratio of the two most recent intervals between rewards from the two alcoves.

I conclude, therefore, that the evolutionary challenge to the distinction between System 1 and System 2 is sustained: since non-human animals engage in unreflective processes that can be both flexible and rule-governed, it is puzzling how a distinct system for reflective reasoning could ever have evolved. What I shall suggest in Section 2.3 is that there is a sense in which it did not. Rather, reflective reasoning is realized in cycles of operation of unreflective forms of cognition, building upon and exapting the resources provided by the latter.

2.3. The mental rehearsal of action

In outline, the proposal is that reflection operates like this: action-schemata are selected and activated, and are mentally rehearsed (with overt action suppressed); this gives rise to conscious images which are globally broadcast (in the manner of Baars, 1988) and thus made available as input to the full suite of intuitive systems; the latter draw inferences from them, activate relevant memories, and issue in emotional reactions; during decision making the somatic consequences of the latter are monitored (in the manner of Damasio, 1994) and motivations are adjusted up or down accordingly; and the result is a whole new cognitive and affective environment influencing the selection of the next mentally rehearsed action (or in some cases, issuing in overt action). We have robust evidence of each of the components appealed to in this account, and each is very likely present in non-human animals, as I shall now briefly explain.

The creation of «forward models» of the expected sensory consequences of activated action-schemata is now quite well understood, as is its original function in facilitating fast on-line correction of action as it unfolds (Wolpert & Ghahramani, 2000; Wolpert & Flanagan, 2001; Wolpert and others, 2003). When attended to, these forward models issue in conscious motor imagery, as well as imagery of other sorts; and they can also do so when actions are mentally rehearsed with overt performance suppressed (Jeannerod, 2006). The result is sequences of motor images, visual images, or auditory images (often in the form of so-called «inner speech», when the actions rehearsed are speech actions), which serve as the conscious components of reflective thought.

It is also quite widely agreed that «global broadcasting» underlies the conscious status of conscious experiences and images (Baars, 2002; Dehaene & Naccache, 2001; Dehaene and others, 2003; Baars and others, 2003). Attended perceptual and imagistic representations are made accessible as input to a wide range of other cognitive systems for drawing inferences, for evoking and forming memories, for decision making, and for generating emotional and motivational responses. Moreover, there is evidence that motor cortex is active in the creation and transformation of visual images (Kosslyn, 1994; Ganis and others, 2000; Richter and others, 2000; Kosslyn and others, 2001; Lamm and others, 2001), and we also know that during episodes of inner speech not only the language comprehension areas of the brain but also language *production* areas and associated regions of motor cortex are active (Paulescu and others, 1993; Shergill and others, 2002). So we have evidence that activations of motor schemata are used to drive the sequences of conscious images found in reflective thinking. In addition, Damasio and colleagues have amassed a great deal of evidence of the crucial role that monitoring of emotional reactions to globally broadcast images plays in normal human decision-making (Damasio, 1994, 2003).

There is no reason to think that any of the components appealed to in this account is uniquely human (with the exception of speech and inner speech, of course). Indeed, there is some reason to think that apes might occasionally engage in reflective decision-making involving mental rehearsals of action, issuing in so-called «insight» phenomena (Carruthers, 2006). Moreover, we can be confident that by the time of *Homo ergaster*, some 1.4 million years ago, our ancestors were regularly using mental rehearsal of action in their reasoning and decision making (2006). For we know that there is no other way of producing the symmetrical hand-axes and blades of that era, using variable and unpredictable materials (Gowlett, 1984; Pelegrin, 1993; Mithen, 1996; Schlinger, 1996; Wynn, 2000). Stone knappers need to be able to plan several strikes ahead, in each case visualizing both the intended blow and its anticipated effects.

2.4. Reflection explained

Note that the above account has the resources to explain a significant number of the properties generally attributed to System 2, while also avoiding the puzzles about its existence raised in Section 2.1. Because globally broadcast images are conscious, that element of each cycle of mentally rehearsed action will also be conscious (while the cognitive activity that precedes and follows the broadcast image will generally be *unconscious*). And because mental rehearsal activates and co-opts the resources of the various intuitive reasoning systems, its overall operations are likely to be significantly slower than most of the latter. Nor is there any special difficulty in explaining

how reflective reasoning could have evolved. For rather than existing alongside of intuitive reasoning systems while performing many of the same functions, reflection is partly realized in cycles of operation of the latter, utilizing pre-existing mechanisms and capacities. All that had to evolve were (a) a language system and (b) a disposition to engage in mental rehearsals of action on a routine basis.

Moreover, because action selection in general is under intentional control and can be influenced by normative belief and verbal instruction, so can the operations of the described reflective system. We can *choose* to engage in mental rehearsal (often unconsciously), just as we choose to engage in any other form of action. And just as with other forms of action, some sequences of rehearsal can be produced smoothly and automatically, resulting from previous practice (think, here, of doing an addition sum in your head). Others can be guided by beliefs about how one *should* reason, sometimes by activating a stored memory of a previous instruction (when faced with a version of the Wason selection task, for example, I might rehearse the sentence, «In order to evaluate a conditional, I should look for cases where the antecedent is true and the consequent false», or I might form a mental picture the standard truth-table for the conditional). And of course with each iteration of mentally rehearsed action the various System 1 systems that «consume» the globally broadcast images become active, sometimes producing an output that contains or contributes towards the solution.

3. AGAINST THE SYSTEM 1 / SYSTEM 2 DISTINCTION

The ideas presented in Section 2 (when suitably worked out and developed; see Carruthers, 2006, 2009, for more in this direction) appear sufficient to establish the reality of the distinction between intuitive and reflective reasoning, while explaining how the latter is realized in cycles of mental rehearsal. The present section will argue that this is not, however, a vindication of the System 1 / System 2 distinction as it has normally been understood. Quite the contrary. This will be demonstrated, in part, via examination of some recent data that fall outside the normal range that reasoning researchers might typically consider.

3.1. Heuristics can be «rational»

One powerful critique of the System 1 / System 2 distinction is implicit in the «simple heuristics» research program of Gigerenzer and others (1999)². For they

² It should be noted, however, that some reasoning researchers have always understood the System 1 / System 2 distinction in such a way as to leave room for the point developed in this section. Thus, Evans and Over (1996), in particular, distinguish between two different kinds of rationality that are

challenge the assumption that intuitive processes are always «irrational». Although these processes might fail to comply with the normative standards laid down by philosophers and logicians, they argue that standards of rationality need to be relativized to the cognitive powers and life-demands of those who employ them. Human beings have to form beliefs and take decisions in real time (often seconds or minutes; rarely longer), and they have to do so given limited and noisy information, and with severely limited computational powers. In such circumstances compliance with rational norms might be impossible. And indeed, when people *try* to take decisions in a way that complies with normative standards —as did Darwin when attempting to decide whether or not to get married, forming a weighted list of the pros and cons— they often fail, frequently overturning the results of their own calculations in favor of a «gut feeling» (Gigerenzer, 2007). Moreover, when heuristic methods are pitched against «normatively correct» ones, they often prove remarkably successful, and are frequently more robust than the latter in generalizing successfully to new environments (Gigerenzer and others, 1999).

Everyone allows, of course, that heuristic reasoning methods can lead to errors, and that humans often think and act in ways that are downright foolish. But instead of focusing on the distinction between «error-prone» heuristics and a supposed set of ideally rational norms, it is more fruitful to investigate the circumstances in which a given heuristic works well and those in which it doesn't. This has issued in the concept of «ecological rationality» (Gigerenzer, 2000), which is the idea that particular reasoning heuristics will be well-adapted to certain cognitive environments but not others.

It is worth emphasizing, however, that proponents of the simple heuristics program need not deny the distinction between intuitive and reflective reasoning. And they had better *not* do so, for there is no doubt that humans often engage in conscious forms of reflection and that this can sometimes lead to improved outcomes. What they insist on (and rightly so) is just that the intuitive / reflective distinction fails to line up with the irrational / rational distinction (at least when the latter is properly naturalized).

3.2. Intuitive reasoning can be «slow»

Recall that System 1 processes are supposed to be *fast* by comparison with reflective forms of reasoning and decision making. No doubt this is true of many intuitive systems. But it surely isn't true of all. Consider the processes that issue in romantic love, or romantic attachment.

applicable to the two systems, one of which is ecological rationality (leading to successful outcomes in many circumstances) whereas the other is normative rationality.

Although some people believe in love at first sight, there is no doubt that in general the process takes considerably longer —often days or weeks, and sometimes months or years. Moreover, although people sometimes reflectively weigh up the good and bad points of a prospective partner, this certainly doesn't always happen, and it is doubtful how effective it is when it does. Yet there is good reason to regard the process in question as one of reasoning. For psychologists have documented the multiple cues that influence romantic attachment, which are gathered over time and appear to be integrated with one another somehow to issue in an eventual intuition of the form, «This person is for me». These range from sensitivity to cues of kindness (such as being nice to dogs and children) and intelligence (such as displaying a sense of humor), to the influence of physical attractiveness (including indicators of youthfulness in women and bodily symmetry in both sexes), to the unconscious detection of pheromones, and chemical information about the other person's immune system obtained from saliva through kissing (Barrett and others, 2002; Buss, 2005). And in addition, of course (as readers of Jane Austen will be well aware), our own level of attachment is strongly influenced by signals of interest from the prospective partner.

A quite different form of intuitive reasoning that is slow rather than fast is popularly known as «sleeping on it». It is a familiar occurrence in daily life that one's reflective reasoning about some problem has become «stuck» —one is unable to see a way to a solution—. So, one lays the problem aside, either literally going to sleep for the night or occupying oneself with other tasks. But then sometimes after an interval the solution suddenly and unexpectedly emerges into consciousness. It is natural to think that one must have continued reasoning about the problem, unconsciously, during the interim. Admittedly, this isn't the only interpretation possible. For it may be that the reason one had got stuck with the problem in the first place resulted from adopting an inappropriate mental «set», which framed and constrained one's conscious reflection. What happened during the time when one was asleep or doing other things, then, may have been that this initial set was *forgotten*. With that constraint out of the way, one is then able to approach the problem afresh, often resulting in success.

No doubt this sort of «set shifting» explanation is sometimes appropriate (Schooler & Melcher, 1995). But it can't explain those cases where a solution emerges fully-formed into consciousness without any prior attention to the problem. For if all that happens in the interval following one's previous attempts is that one forgets the initial line of approach to the problem that had led one astray, then, one would expect that one would thereafter have to think *reflectively* about the problem once again in order to achieve a solution. But this isn't always what happens. And indeed, we will see towards the end of Section 3.3 that there is direct experimental evidence that unconscious reasoning often takes place in the intervals during which one's conscious mind is otherwise occupied.

3.3. Intuitive reasoning is better than reflection for some tasks

One set of data supporting the claim made in the title of this section is provided by Wilson and others (1993). They offered their subjects a choice between a number of different posters to take home to display on their walls. All subjects were allowed some time to study the posters on offer before making their choice. But one group were asked to articulate the positive and negative features of the various options, thus forcing them to be reflective in making their selection. The experimenters then followed up with each of the subjects a week later, asking them how happy they were with their new poster. Those who had chosen unreflectively expressed significantly greater satisfaction; and when asked how much money would be necessary to persuade them to sell their choice of poster back to the experimenters, those in the unreflective condition asked for double the amount demanded by those who had made their choice following explicit reflection.

This seems to be a case where unreflective choice is better than reflection, at least in the sense of leading to greater subjective satisfaction with the chosen object. Wilson and others (1993) hypothesize that the effect occurs because reflection will inevitably focus most attention on those good-making and bad-making features that are easily expressible in language, as well as perhaps narrowing attention to fewer features more generally. And the moral seems to be that when faced with a complex multi-faceted choice one should pay attention to as many features as one can, but one shouldn't attempt to make one's decision via conscious reflection and reasoning; rather, following study of the choices on offer one should «trust one's gut feeling».

Even more decisive evidence of the benefits of unconscious over conscious forms of decision making has been provided by Dijksterhuis and colleagues (Dijksterhuis, 2004; Dijksterhuis and others, 2006). In one set of experiments subjects were asked to make a complex choice under one of three conditions. For example, they might be asked to choose the best of four cars when provided with ten items of information about each (cost, gas mileage, and so on). The information was arranged so that one of the items was best (with say eight positive features and two negative ones) and another was worst (with eight negative features and two positive ones), whereas the other two choices were of intermediate quality (in some experiments the good-making and bad-making features were calibrated against subjects' own evaluations made at another time).

In one condition subjects were asked to make an immediate choice having been presented with all the information; in a second condition they were told to reflect for a few minutes before making their choice; and in the third condition they were provided with a demanding distracter task for the same amount of time, which would have made conscious reflection impossible. What the experimenters found

is that intuitive reasoning is significantly better than immediate choice, suggesting that the process in question is fairly slow. But they also found that intuitive reasoning is significantly better than conscious reflection (there was no qualitative difference between immediate and reflective choosing, suggesting that in these circumstances reflection brings no benefit).

Indeed, Dijksterhuis and colleagues have been able to demonstrate that successful intuitive reasoning is goal-dependent (Bos and others, 2008)³. They have thereby demonstrated that such reasoning isn't always automatic, in contrast with almost every theorist who writes about the System 1 / System 2 distinction (Evans, 2009; Stanovich, 2009). In one experiment subjects studied complex information about four different makes of car, as before. One set of subjects was asked to make a choice after a period of reflection, while another group was told that they would make their choice after completing an attentionally-demanding task, as previously. But now the third group was told that the relevant phase of the experiment was over, hence blocking the formation of any goal of choosing between the four cars. But after they had completed the same attention-occupying task as the second group, they were in fact presented with the same choice. There were no significant differences between the first and third groups. Only the subjects who had the opportunity to reason intuitively about their choice and *also* had the goal of choosing a car did better than those who chose reflectively. (Note that this result is inconsistent with a «set shifting» explanation, since the third group would have had the same opportunity to forget their initial characterization of the problem as did the second group.)

Even more remarkably, Bos and others (2008) were able to demonstrate that intuitive reasoning is guided by quite specific goals. They asked another set of subjects to study complex information about two kinds of object (cars and potential room-mates). One group of subjects was then told, before undertaking an attentionally-demanding task, that they would be choosing between the cars, whereas the other group was told that they would be choosing between room-mates. But then both groups were actually thereafter asked to choose the best car *and* the best roommate. Members of both groups showed the benefits of unconscious reflection, but *only* with respect to the items that they had been given the goal of selecting between.

3.4. Intuitive reason can approximate to the highest normative standards

Not only can intuitive reasoning be markedly better than reflection, in some circumstances, but some unconscious processes can be highly optimal, approximating

³ They have also shown that intuitive reasoning issues in better organization of memory than does conscious reflection. See Dijksterhuis and Nordgren (2006).

to the most stringent normative standards. One example of this has already been discussed in Section 2.2. This is the data from Gallistel and others (2001) concerning rate estimation in rats and pigeons, showing that the animals were tracking the randomly changing rates about as closely as it is theoretically possible to do (I presume that these animals were making their estimates intuitively, and weren't engaging in conscious reflection!). But an even more dramatic example is provided by Balci and others (2009). They tested swift and intuitive assessments of risk, using similar experiments in both humans and mice, with very similar results (I shall discuss only the human data here, however, since our main topic is dual systems of reasoning in human beings).

The human subjects were set the task of «capturing» an object in one of two positions on a computer screen for a reward. There were two types of trial, short latency and long latency, the probability of each of which varied from one series of trials to the next. If the trial was a short one, the target could be captured in the left-hand position within two seconds of the stimulus onset; if the trial was a long one, the target could be captured in the right-hand position during the third second. Subjects were therefore required to estimate the optimum time to switch from the short-latency strategy to the long-latency strategy. This estimate depends upon two factors. One is the objective chance that the interval would be either short or long (this was set by the experimenters in each series of trials). The other is the accuracy of each subject's own estimate of elapsed time (which varies from person to person, but which is normally in the region of $\pm 15\%$). Balci and others (2009) were able to compute the optimum switch time for each subject, combining both sets of probabilities. This was then compared with actual performance. Subjects came within 98% of optimum performance. Moreover, very little learning was involved (either for humans or for the mice). In most series of trials, subjects were just as successful during the first tenth or the first quarter of the series as they were during the final tenth or the final quarter. This is a remarkable result, especially given the repeated finding that humans are so poor at reasoning about probabilities in explicit tasks.

A very different kind of normatively correct intuitive reasoning system is proposed by Sperber and Mercier (2009), which is designed for public argumentation. They initially motivate the need for such a system on evolutionary grounds, having to do with the epistemic vigilance that is necessary to sustain successful systems of communication. They point out that speech is a cooperative activity, and like all forms of cooperation it is vulnerable to being parasitized and undermined by cheaters and free-riders. People therefore need to be alert to the possibility of being misled, and should have a healthy skepticism about what others tell them, especially in circumstances where it is plain that conflicts of interest are involved.

This creates pressure, in turn, for would-be communicators to lay out reasons why the communication should be accepted by their audience; which in turn creates pressure on hearers to evaluate those reasons and to formulate counter-reasons where necessary. So we should predict the emergence of a specialized argumentation system to facilitate these tasks⁴.

Sperber and Mercier (2009) then amass a broad body of evidence in support of their proposal. Part of this consists in an extensive set of studies by a variety of researchers demonstrating that people are much better at reasoning in argumentative contexts (and this isn't just a product of greater motivation, since paying subjects to reason well in individual reasoning tasks has little effect). For example, although subjects are notoriously bad at employing modus tollens arguments in standard paper-and-pencil tasks, when people want to attack the views of others they are actually very good at employing such arguments (Pennington & Hastie, 1993). Even more striking, Moshman and Geil (1998) had one group of subjects attempt Wason conditional reasoning tasks individually. Only 9% succeeded. Another set of subjects were put into groups and asked to solve the same tasks, and a massive 70% of the groups succeeded. It seems that subjects have the capacity to recognize correct solutions when proposed by others. Indeed, it seems that such solutions can emerge out of a process of discussion. For when groups were formed using only subjects who had previously failed at the task on an individual basis, 30% of those groups were nevertheless able to come up with the correct solution. Moreover, discussion involving dissent has been shown to be a crucial determinant of group performance (Schulz-Hardt and others, 2006).

In addition, and in contrast with the extensive studies that demonstrate poor reasoning in subjects when working on tasks individually, naturalistic studies of persuasion have found that people in argumentative contexts are remarkably good at distinguishing good arguments from bad ones and at generating good arguments for themselves (Petty & Wegener, 1998; Neuman and others, 2006). Especially striking is the fact that even young children can engage successfully in argumentation (Stein & Albro, 2001) and can spot argumentative fallacies (Baum and others, 2007).

⁴ Confusingly, Sperber and Mercier themselves describe the argumentation system as a *reflective* one, on the grounds that it is often engaged when we explicitly and consciously evaluate the arguments of others or propose counter-arguments ourselves. This may be so. But since the internal operations of the system are always unconscious it is plain, I think, that it should be classified as an intuitive system. For *all* intuitive systems are capable of issuing in conscious outputs, and many may help to guide cycles of conscious reflection. Indeed, when presented with someone's argument, the argumentation system will swiftly and unconsciously generate an intuition about the strength of that argument, together with intuitions about its likely weak spots. This fits the normal characterization of an intuitive system.

3.5. Reflection can employ heuristics

So far this section has built a cumulative case that it is a mistake to describe intuitive reasoning systems as «quick, automatic, and dirty». For some of them can be slow; some can be controlled; some are more successful than reflective reasoning when dealing with the same types of task; and some can approximate to the highest normative standards. I now turn attention to reflective kinds of reasoning, questioning the extent to which such reasoning processes possess the properties generally attributed to System 2. The first point I want to make is that reflective reasoning, just as well as intuitive reasoning, can have a heuristic character.

While reflection can sometimes involve appeal to valid norms of reasoning, it can also involve the use of heuristics (this point is now fully accepted by some two-systems theorists, especially recently; see Frankish [2004], Evans [2009], and Stanovich [2009]). Some of these can be good and useful, and some can be bad. The «sleep on it» heuristic, which is often consciously and reflectively employed by people in our culture, might be a good example of the first sort. Moreover, if Sperber and Mercier (2009) are right about the existence of an intuitive argumentation system, then one might predict that a useful heuristic to employ when engaged in private reflective reasoning about some problem would be to imagine that the context is an argumentative one (for example, by engaging an imaginary opponent). For by simulating argument, one will provide the cues necessary for the argumentation system to become active. And then each cycle of mental rehearsal will be evaluated by that system, thus facilitating the transition to the next step in one's reasoning in a normatively correct manner.

No doubt some conscious reasoning heuristics may result from individual learning and theorizing. But many will be culturally acquired (the «sleep on it» heuristic is surely an example of the latter sort). And a moment's reflection is sufficient to demonstrate that many culturally sanctioned reasoning heuristics have been quite bad. Consider the common practice in many cultures of consulting an oracle, such as the entrails of a newly killed chicken, when making an important decision. It is plain to us that decisions made on this basis will be at chance. And it is equally obvious that people in such a culture would be much better off not reflecting about the decision at all, given that reflection will call to mind the «look at the entrails» heuristic, which is then likely to be followed⁵.

⁵ Note there is also a quite different sense in which reflection can involve heuristics. For insofar as reflection often activates and depends upon intuitive reasoning processes, then to that extent it will involve any heuristics that are employed within those processes. What I have in mind in Section 3.5., in contrast, are heuristics that are consciously activated and deployed during reflection.

3.6. The fragmentation of reflective reasoning

The present section argues that reflective reasoning processes are by no means unitary in character. On the contrary, some can be based upon normative beliefs and can lead to good performance if the beliefs in question are correct (or to poor performance if the beliefs are incorrect). Some are skill-based, mediated by learned and internalized action sequences. And some depend upon the modular systems that mediate transitions within the architecture. I shall discuss each in turn. But I should stress that I don't mean to imply that these three types of process are mutually exclusive. On the contrary, some instances of reflective reasoning will involve a combination of at least two of them, and some will involve all three.

Recall that reflection, on the account provided in Section 2, is action-based. It can therefore be influenced in any of the ways that action can. And in particular, it can be guided by the subject's beliefs (especially normative beliefs about how one *should* reason). One effect of taking a course in logic or probability theory, for example, is that one forms a number of beliefs about the ways in which reasoning should be conducted. As we noted earlier, when asked to evaluate a conditional one might pause to reflect, perhaps asking oneself the explicit question, «How does one evaluate a conditional?». When globally broadcast, this might call to mind (whether consciously or unconsciously) the acquired belief, «A conditional is only false when the antecedent is true and the consequent false». This belief provides the information needed to guide the next stage of reflection, leading one to think, in the context of a Wason selection task, for example, «So I should turn over the *P* and the $\sim Q$ cards». Likewise when asked to assign probabilities to a series of statements, pausing to reflect might call to mind the belief, «A conjunction cannot be more probable than one of its conjuncts». This might then be sufficient to override some of one's initial intuitions with respect to the cases.

These are examples where reflection will lead to improved performance, because the beliefs accessed via reflection are true. But of course training can also issue in beliefs that are false, either because of errors in the transmission process, or because the teachers themselves have false beliefs about how one should reason. The history of the philosophy of science, for example, is littered with cases where philosophers have endorsed normative standards for the conduct of scientific reasoning—including naïve inductivism and Popperian falsificationism—which would have seriously impeded scientific progress had they been adopted. Fortunately for us, most scientists received their training from other scientists, and not from philosophers of science, so the normative beliefs in question did little damage (for an argument that significant damage was done in psychology, however, especially in studies of children's scientific reasoning, see Koslowski, 1996).

Since reflection is action-based, it can also involve habitual sequences that have been synergized together, in the same way that physical skills can involve such sequences. These can be acquired through training and practice, often in the context of formal schooling. This is surely true of the sequences of mental rehearsal involved when doing multiplication or long division sums, for example. But reasoning skills can also be acquired through imitation of the overt reasoning of other people. Much of the learning that goes on in universities is probably of this sort. Philosophy instructors exhibit, in their reasoning during class, patterns of thought that they hope their students will copy. And likewise faculty involved in scientific lab meetings will be exhibiting patterns of reasoning and problem solving to the students who are present.

The most straightforward way for reflection to operate, however, is by harnessing and engaging the resources of the intuitive systems that are the consumers of globally broadcast events (Shanahan & Baars, 2005). For example, by asking myself a question I might evoke information contained in one of the intuitive systems that I didn't know that I had, and that might not otherwise have become activated (Frankish, 2004). Another instance of the same phenomenon was discussed in Section 3.5, on the assumption that Sperber and Mercier (2009) are correct about the existence of an intuitive argumentation system. Reflectively simulating an argument with an imagined opponent will cause this system will become active, thereby leading to an improvement in one's reasoning.

Moreover, the deployment of intuitive systems through cycles of reflection is central to Damasio's (1994, 2003) account of the operations of reflective practical reason. By mentally rehearsing a proposed action, images of that action become globally broadcast to the full suite of intuitive systems. Some of these elaborate the image to include its immediate consequences. But the results are also received as input by the emotional and motivational systems which respond accordingly, issuing in a set of bodily / emotional reactions. These are monitored by the subject, adjusting the motivation towards the proposed action upwards or downwards. We know from cases of people with frontal brain damage who have lost this monitoring capacity (but whose *theoretical* reasoning about action and the costs and benefits of different options can be fully intact) that this sort of practical reflection is crucial for leading a successful life (Damasio, 1994). But it is by no means infallible. On the contrary, it is subject to a number of fallacies and biases, such as overestimating the pain that we will feel through the loss of some major goal, for instance failure to be granted tenure (Gilbert, 2005).

4. COMPARISONS AND CONCLUSION

I conclude that there is no pair of reasoning systems (nor pair of system types) that lines up with all or most of the properties traditionally ascribed to System 1 and System 2. What really exists is a distinction between a set of intuitive systems and a reflective (mental rehearsal involving) architecture. This secures *some* of the properties of the original distinction (unconscious versus consciousness-involving, impervious to instruction versus malleable, and impenetrable by normative belief versus capable of being controlled by such beliefs). But some of the subsystems or processes within each can display normatively correct reasoning, each often makes use of heuristics, some intuitive systems can be as slow or slower than reflection, and some are goal-dependent rather than automatic. Moreover, some intuitive systems are better than reflective processes on the same task. In addition, although all reflective reasoning employs a common *architecture* of mental rehearsal and global broadcast, very different reasoning strategies and processes can be employed within that architecture. Some of these can be heuristic based, and some can be quite unreliable. Indeed, the very same architecture is also employed for *non*-reasoning processes, such as personal fantasy and mere «idle thinking».

I suggest, therefore, that while the distinction between intuitive and reflective processes is real, the distinction between System 1 and System 2 reasoning systems is not. Reasoning researchers should abandon the latter conceptual framework, and should instead focus on investigating the nature and variety of processes of both intuitive and reflective kinds. However, others, too, have recently suggested a reconceptualization of the System 1 / System 2 distinction. I shall briefly consider the main two alternatives, contrasting them with my own proposal before concluding.

Evans (2009) now argues that System 1 processes fall into two distinct types. There are those that *compete with* System 2 processes, operating alongside them and issuing in intuitions, judgments, or decisions. And there are those that *influence* the operations of System 2, or *provide input* to System 2, while not generating the same sorts of outcomes in their own right. He also argues that we need to recognize a third system, which receives input from Systems 1 and 2 and makes decisions between them⁶. I have no quarrel with these suggestions. But in respect of System 2 itself, Evans argues that it can be identified with the set of processes that involve working memory (which can otherwise differ from one another in many of the respects outlined in Section 3). Although this suggestion contains an important element of truth, it doesn't quite capture the distinction that we need. For anyone answering

⁶ Stanovich (2009), too, draws a related but different distinction within System 2 processes, between the computational powers of the conscious system and its reflective beliefs and dispositions.

a question in an intuitive, *unreflective*, way will nevertheless be employing working memory to process the task instructions and maintain the ensuing representations long enough for the intuitive systems to generate an answer. What is distinctive of reflective processes is that *actions* should be mentally rehearsed, thereby issuing in appropriate sequences of imagery within working memory. In short, reflection is an *activity*, whereas working memory, as such, is not. But in any case (since processes that involve working memory are a heterogeneous lot) Evans' recent position is consistent with one of the main claims argued for here, namely that there is no unified character to the set of processes labeled «System 2».

Frankish (2009), in contrast, characterizes the key difference between System 1 and System 2 in terms of *control*, as does Stanovich (2009). But we have already had occasion to note in Section 3.3 that some *intuitive* processes can be controlled as well, at least in the sense that they are goal-dependent. Moreover, not all reflective processes are controlled by the agent's beliefs and decisions, surely. For recall that reflection is *action*-based. Yet we know that actions can be triggered and guided independently of conceptual thought and decision-making. People will thoughtlessly mirror one another's behavior in social situations, for example (Jeannerod, 2006), and the mere sight of object affordances (such as the «to-be-grasped» shape of a hammer or coffee mug) will trigger activation of the appropriate motor schemata, which then need to be actively *suppressed* to prevent them from becoming acted upon (Shallice and others, 1989; Jacob & Jeannerod, 2003). Moreover, the work of Bargh and colleagues demonstrates that action schemata (such as walking in the manner of an old man) can be activated and executed by suitable conceptual priming without giving rise to any *belief* that the subject is an old man or to any corresponding intention (Bargh and others, 1996, 2001; Bargh, 2005). We should predict, then, that reflective processes can be just as automatic as overt actions can be. Moreover (and just as one might expect, on my account), in individuals suffering from obsessive-compulsive disorder, not only behavior but also reflective thought processes are undertaken compulsively, outside of the control of the agent (Leckman and others, 1997).

I conclude that human reasoning is best characterized in terms of a distinction between a varied set of intuitive systems, on the one hand, and a cognitive architecture that enables us to engage in reflective thinking and reasoning, on the other (with the latter operating in heterogeneous ways). If one of the goals of science is to discover what natural kinds there are in the world—in the sense of homeostatic property clusters with unifying causal etiologies (Boyd, 1989, 1991)—then cognitive scientists would be well-advised to abandon the System 1 / System 2 conceptual framework. The human mind is messier and more fine-grained than that⁷.

⁷ I am grateful to Keith Frankish and Hugo Mercier for their comments on an earlier draft of this paper.

REFERENCES

- Baars, Bernard (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Baars, Bernard (2002). The Conscious Access Hypothesis: Origins and Recent Evidence. *Trends in Cognitive Science*, 6, 47-52.
- Baars, Bernard, Thomas Ramsoy & Steven Laureys (2003). Brain, Consciousness, and the Observing Self. *Trends in Neurosciences*, 26, 671-675.
- Balci, Fuat, David Freestone & Charles Gallistel (2009). Risk assessment in Man and Mouse. *Proceedings of the National Academy of Sciences*, 106, 2459-2463.
- Bargh, John (2005). Bypassing the Will: Toward Demystifying the Nonconscious Control of Social Behavior. In Ran R. Hassin, James S. Uleman and John A. Bargh (eds.), *The New Unconscious* (pp. 37-58). Oxford: Oxford University Press.
- Bargh, John, Mark Chen & Lara Burrows (1996). Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Priming on Action. *Journal of Personality and Social Psychology*, 71, 230-244.
- Bargh, John, Peter Gollwitzer, Annette Lee-Chai, Kimberly Barndollar & Roman Troetschel (2001). The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals. *Journal of Personality and Social Psychology*, 81, 1014-1027.
- Barrett, Louise, Robin Dunbar & John Lycett (2002). *Human Evolutionary Psychology*. New Jersey: Princeton University Press.
- Baum, Laura, Judith Danovitch & Frank Keil (2007). Children's Sensitivity to Circular Explanations. *Journal of Experimental Child Psychology*, 100, 146-155.
- Berry, Diane & Zoltan Dienes (1993). *Implicit Learning*. New Jersey: Erlbaum.
- Bickerton, Derek (2009). *Adam's Tongue: How Humans Made Language, How Language Made Humans*. New York: Hill and Wang.
- Maarten Bos, Ap Dijksterhuis & Rick van Baaren (2008). On the Goal-dependency of Unconscious Thought. *Journal of Experimental Social Psychology*, 44, 1114-1120.
- Boyd, Richard (1989). What Realism Implies and What it Does Not. *Dialectica*, 43, 5-29.
- Boyd, Richard (1991). Realism, Anti-foundationalism, and the Enthusiasm for Natural Kinds. *Philosophical Studies*, 61, 127-148.
- Buss, David (ed.) (2005). *The Handbook of Evolutionary Psychology*. New Jersey: Wiley.
- Carruthers, Peter (2006). *The Architecture of the Mind*. Oxford: Oxford University Press.
- Carruthers, Peter (2009). An Architecture for Dual Reasoning. In Jonathan Evans and Keith Frankish (eds.), *Two Minds: Dual Processes and Beyond* (pp. 109-127). Oxford: Oxford University Press.

- Chaiken, Shelly, Akiva Liberman & Alice H. Eagly (1989). Heuristic and Systematic Processing within and Beyond the Persuasion Context. In James S. Uleman and John Bargh (eds.), *Unintended Thought* (pp. 212-252). New York: Guilford Press.
- Damasio, Antonio (1994). *Descartes' Error*. London: Papermac.
- Damasio, Antonio (2003). *Looking for Spinoza*. San Diego: Harcourt.
- De Marco, Rodrigo & Randolph Menzel (2008). Learning and Memory in Communication and Navigation in Insects. In Randolph Menzel and John H. Byrne (eds.), *Learning Theory and Behavior* (pp. 477-498). Amsterdam: Elsevier.
- Dehaene, Stanislas & Lionel Naccache (2001). Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework. *Cognition*, 79, 1-37.
- Dehaene, Stanislas, Claire Sergent & Jean-Pierre Changeux (2003). A Neuronal Network Model Linking Subjective Reports and Objective Physiological Data during Conscious Perception. *Proceedings of the National Academy of Science*, 100, 8520-8525.
- Dijksterhuis, Ap (2004). Think Different: The Merits of Unconscious thought in Preference Development and Decision Making. *Journal of Personality and Social Psychology*, 87, 586-598.
- Dijksterhuis, Ap & Loran Nordgren (2006). A Theory of Unconscious Thought. *Perspectives on Psychological Science*, 1, 95-109.
- Dijksterhuis, Ap, Maarten Bos, Loran Nordgren & Rick van Baaren (2006). On Making the Right Choice: The Deliberation-Without-Attention Effect. *Science*, 311(5763), 1005-1007.
- Evans, Jonathan & David Over (1996). *Rationality and Reasoning*. Hove: Psychology Press.
- Evans, Jonathan (2009). How Many Dual-Processing Theories do We Need? One, Two, or Many? In Jonathan Evans and Keith Frankish (eds.), *In Two Minds* (pp. 33-88). New York: Oxford University Press.
- Frankish, Keith (2004). *Mind and Supermind*. Cambridge: Cambridge University Press.
- Frankish, Keith (2009). Systems and Levels: Dual-system Theories and the Personal-Subpersonal Distinction. In Jonathan Evans and Keith Frankish (eds.), *In Two Minds* (pp. 89-107). New York: Oxford University Press.
- Gallistel, Charles & John Gibbon (2001). Time, Rate and Conditioning. *Psychological Review*, 108, 289-344.
- Gallistel, Charles, Terence Mark, Adam King & Peter Latham (2001). The Rat Approximates an Ideal Detector of Rates of Reward: Implications for the Law of Effect. *Journal of Experimental Psychology: Animal Behavior Processes*, 27, 354-372.
- Gallistel, Charles & Adam King (2009). *Memory and the Computational Brain*. Singapore: Wiley-Blackwell.

- Ganis, Giorgio, Julian Keenan, Stephen Kosslyn & Alvaro Pascual-Leone (2000). Transcranial Magnetic Stimulation of Primary Motor Cortex Affects Mental Rotation. *Cerebral Cortex*, 10, 175-180.
- Ghahramani, Zoubin & Daniel Wolpert (1997). Modular Decomposition in Visuomotor Learning. *Nature*, 386, 165-193.
- Gigerenzer, Gerd (2000). *Adaptive Thinking*. New York: Oxford University Press.
- Gigerenzer, Gerd (2007). *Gut Feelings*. New York: Viking Press.
- Gigerenzer, Gerd, Peter Todd & The ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Gilbert, Daniel (2005). *Stumbling on Happiness*. New York: Vintage Books.
- Gowlett, John (1984). Mental Abilities of Early Man: A Look at Some Hard Evidence. In Robert Foley (ed.), *Hominid Evolution and Community Ecology* (pp. 167-192). Nueva York: Academic Press.
- Jacob, Pierre & Marc Jeannerod (2003). *Ways of Seeing*. New York: Oxford University Press.
- Jeannerod, Marc (2006). *Motor Cognition*. New York: Oxford University Press.
- Kahneman, Daniel (2003). A Perspective on Judgment and Choice: Mapping Bounded Rationality. *American Psychologist*, 58, 697-720.
- Kahneman, Daniel & Shane Frederick (2002). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In Thomas Gilovich, Dale Griffin and Daniel Kahneman (eds.), *Heuristics and Biases* (pp. 49-81). Cambridge: Cambridge University Press.
- Koslowski, Barbara (1996). *Theory and Evidence*. Cambridge: The MIT Press.
- Kosslyn, Stephen (1994). *Image and Brain*. Cambridge: The MIT Press.
- Kosslyn, Stephen, William Thompson, Mary Wraga & Nathaniel Alpert (2001). Imagining Rotation by Endogenous versus Exogenous Forces: Distinct Neural Mechanisms. *NeuroReport*, 12, 2519-2525.
- Lamm, Claus, Christian Windtschberger, Ulrich Leodolter, Ewald Moser & Herbert Bauer (2001). Evidence for Premotor Cortex Activity during Dynamic Visuospatial Imagery from Single Trial Functional Magnetic Resonance Imaging and Event-Related Slow Cortical Potentials. *Neuroimage*, 14, 268-283.
- Leckman, James and others (1997). Symptoms of Obsessive-Compulsive Disorder. *American Journal of Psychiatry*, 154, 911-917.
- Manoel, Edison, Luciano Basso, Umberto Correa & Go Tani (2002). Modularity and Hierarchical Organization of Action Programs in Human Acquisition of Graphic Skills. *Neuroscience Letters*, 335, 83-86.

- Menzel, Randolph and others (2005). Honey Bees Navigate According to a Map-Like Spatial Memory. *Proceedings of the National Academy of Sciences*, 102, 3040-3045.
- Mithen, Steven (1996). *The Pre-History of the Mind*. London: Thames and Hudson.
- Moshman, David & Molly Geil (1998). Collaborative Reasoning: Evidence for Collective Rationality. *Thinking and Reasoning*, 4, 231-248.
- Neuman, Yair, Michael Weinstock & Amnon Glasner (2006). The Effect of Contextual Factors on the Judgment of Informal Reasoning Fallacies. *Quarterly Journal of Experimental Psychology*, 59, 411-425.
- Paulescu, Eraldo, Chris D. Frith & Richard Frackowiak (1993). The Neural Correlates of the Verbal Component of Working Memory. *Nature*, 362, 342-345.
- Pelegrin, Jacques (1993). A Framework for Analyzing Prehistoric Stone Tool Manufacture and a Tentative Application of Some Early Stone Industries. In Arlette Berthelet and Jean Chavaillon (eds.), *The Use of Tools by Human and Non-Human Primates* (pp. 302-314). New York: Oxford University Press.
- Pennington, Nancy & Reid Hastie (1993). Reasoning in Explanation-Based Decision-Making. *Cognition*, 49(1), 123-163.
- Petty, Richard & John Cacioppo (1986). The Elaboration Likelihood Model of Persuasion. Leonard Berkowitz (ed.), *Advances in Experimental Social Psychology*. Volume 19. New York: Academic Press.
- Petty, Richard & Duane Wegener (1998). Attitude Change: Multiple Roles for Persuasion Variables. In Daniel Gilbert, Susan Fiske and Gardner Lindzey (eds.), *The Handbook of Social Psychology* (pp. 323-390). New York: McGraw-Hill.
- Reber, Arthur (1993). *Implicit Learning and Tacit Knowledge*. New York: Oxford University Press.
- Richter, Wolfgang and others (2000). Motor Area Activity during Mental Rotation Studied by Time-Resolved Single-Trial (fMRI). *Journal of Cognitive Neuroscience*, 12(2), 310-320.
- Schlanger, Nathan (1996). Understanding Levallois: Lithic Technology and Cognitive Archaeology. *Cambridge Archaeological Journal*, 6(2), 75-92.
- Schooler, Jonathan & Joseph Melcher (1995). The Ineffability of Insight. In Steven Smith, Thomas Ward and Ronald Finke (eds.), *The Creative Cognition Approach* (pp. 97-133). Cambridge: The MIT Press.
- Schulz-Hardt, Stefan, Felix Brodbeck, Andreas Mojzisch, Rudolph Kerschreiter & Dieter Frey (2006). Group Decision Making in Hidden Profile Situations: Dissent as a Facilitator for Decision Quality. *Journal of Personality and Social Psychology*, 91(6), 1080-1093.

- Shallice, Tim, Paul Burgess, Frederick Schon & Doreen Baxter (1989). The Origins of Utilization Behavior. *Brain*, 112(6), 1587-1598.
- Shanahan, Murray & Bernard Baars (2005). Applying Global Workspace Theory to the Frame Problem. *Cognition*, 98(2), 157-176.
- Shergill, Suckwinder, Michael Brammer, Rimmei Fukuda, Ed Bullmore, Edson Amaro Robin Murray & Philip McGuire (2002). Modulation of Activity in Temporal Cortex during Generation of Inner Speech. *Human Brain Mapping*, 16(4), 219-227.
- Slooman, Steven (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin*, 119(1), 3-22.
- Slooman, Steven (2002). Two Systems of Reasoning. In Thomas Gilovich, Dale Griffin and Daniel Kahneman (eds.), *Heuristics and Biases* (pp. 379-396). Cambridge: Cambridge University Press.
- Sperber, Dan & Hugo Mercier (2009). Reasoning as a Social Competence. In Jon Elster and Hélène Landemore (eds.), *Collective Wisdom* (pp. 368-392) Cambridge: The MIT Press.
- Stanovich, Keith (1999). *Who is Rational? Studies of Individual Differences in Reasoning*. New Jersey: Erlbaum.
- Stanovich, Keith (2009). Distinguishing the Reflective, Algorithmic, and Autonomous Minds: Is it Time for a Tri-Process Theory? In Jonathan Evans and Keith Frankish (eds.), *In Two Minds* (pp. 55-88). Oxford: Oxford University Press.
- Stein, Nancy & Elizabeth Albro (2001). The Origins and Nature of Arguments. *Discourse Processes*, 32(2-3), 113-133.
- Wilson, Timothy, Samuel Lindsey & Tonya Schooler (2000). A Model of Dual Attitudes. *Psychological Review*, 107(1), 101-126.
- Wilson, Timothy, Douglas Lisle, Jonathan Schooler, Sara Hodges, Kristen Klaaren & Suzanne LaFleur (1993). Introspecting About Reasons Can Reduce Post-Choice Satisfaction. *Personality and Social Psychology Bulletin*, 19(3), 331-339.
- Wolpert, Daniel & Randall Flanagan (2001). Motor Prediction. *Current Biology*, 11(18), 729-732.
- Wolpert, Daniel & Zoubin Ghahramani (2000). Computational Principles of Movement Neuroscience. *Nature Neuroscience*, 3, 1212-1217.
- Wolpert, Daniel, Kenji Doya & Mitsuo Kawato (2003). A Unifying Computational Framework for Motor Control and Social Interaction. *Philosophical Transactions of the Royal Society of London*, B 358, 593-602.
- Wynn, Thomas (2000). Symmetry and the Evolution of the Modular Linguistic Mind. In Peter Carruthers and Andrew Chamberlain (eds.), *Evolution and the Human Mind* (pp. 113-139). Cambridge: Cambridge University Press.

LA BRECHA EN MENTE O POR QUÉ LOS HUMANOS NO SON SOLO GRANDES SIMIOS¹

Robin Dunbar

Miembro de la Academia Británica
Universidad de Oxford

1. INTRODUCCIÓN

La evolución ha tenido una mala acogida por parte de la prensa en las ciencias sociales y humanidades durante el siglo pasado, aunque no siempre ha sido así: el siglo XIX fue testigo de un considerable interés por parte de las ciencias sociales emergentes en las ideas propuestas por los nuevos evolucionistas dentro de la biología. Este no es el momento ni el lugar para profundizar en la historia respecto a por qué los científicos sociales posteriormente rehuyeron con tanta vehemencia a las ideas evolucionistas. Más bien, mi objetivo es simplemente subrayar la afirmación de que una perspectiva evolutiva no es, como a menudo se supone, un paradigma que compite por las explicaciones convencionales de las ciencias sociales. Por el contrario, una perspectiva evolutiva debería funcionar, como lo hace en la biología, como una teoría marco que permite a todas las diferentes subdisciplinas integrarse de forma que puedan conversar entre sí en igualdad de condiciones. La biología se ha beneficiado enormemente de una perspectiva evolutiva durante el último medio siglo en particular, ya que un marco evolutivo ha permitido que los etólogos y ecologistas integren su trabajo con los fisiólogos y genetistas moleculares (aunque los primeros todavía se quejan de la molecularización de la biología de los últimos). Yo sostengo que la psicología —una disciplina notoriamente fraccionada en el mejor de los casos— podría beneficiarse de la misma manera, puesto que un enfoque evolutivo permitiría a los neurocientíficos dialogar con psicólogos sociales y del desarrollo de una manera que, hasta ahora, evidentemente no han podido hacer.

¹ La transcripción de «Mind the Gap or Why Humans Aren't just Great Apes» de Robin Dunbar se publicó por primera vez en el año 2007 (2007b). Esta es la primera traducción al castellano.

No quiero defender esta propuesta en detalle aquí. Simplemente se presenta como un marco para mi charla. Más bien, trataré de demostrar el valor de un marco evolutivo mediante un ejemplo. El propósito de mi charla es esbozar un argumento sobre por qué los humanos son tan diferentes de otros simios y monos, a pesar de que compartimos gran parte de nuestra historia evolutiva con ellos. Mi punto consiste en que, al hacerlo, voy a tener que recurrir a muchas diferentes subdisciplinas de la psicología (así como otras disciplinas como la antropología) cuya integración a esta historia solo es posible porque la teoría evolutiva nos brinda un marco global dentro del cual se deben combinar.

Compartimos una larga historia evolutiva con los grandes simios, en particular: el linaje humano (incluyendo muchas de tantas especies fósiles que han existido durante los últimos 6 millones de años aproximadamente, desde que nuestro linaje se separó del de los chimpancés) está firmemente arraigado en la familia del gran simio africano y comparte con ellos muchos aspectos de su biología, genética, psicología y conducta. Más trivialmente, también compartimos un porcentaje muy alto de nuestro ADN con los chimpancés (aunque *gran parte* de lo que esto significa es discutible). No obstante, seguramente para todos es obvio que no somos «solo grandes simios». Somos muy diferentes de muchas maneras evidentes. Esa diferencia realmente no parece corresponder con los tipos de diferencias anatómicas o cognitivas —nuestro bipedalismo y las habilidades para fabricar herramientas— que los antropólogos han tendido a enfatizar en el pasado. Más bien, quiero discutir que la verdadera diferencia radica en un conjunto mucho más intangible de competencias —la capacidad de vivir en el mundo virtual de la mente. En una palabra, este es el mundo de la cultura—. Voy a centrarme en dos aspectos de la conducta humana que son, en muchos aspectos, arquetípicos de la cultura humana: la narración de historias y la religión. Ambas requieren que podamos imaginar mundos que físicamente no existen. Debo alegar que, a pesar de la tan alardeada capacidad de aprendizaje cultural de los simios, ninguna otra especie viva está siquiera en la misma página que los seres humanos en este aspecto —porque todas las otras especies animales no tienen el poder computacional neuronal necesario para hacerlo posible—. La clave para entender por qué esto es así se encuentra en las razones por las cuales nuestros cerebros han evolucionado.

2. LA EVOLUCIÓN DEL CEREBRO SOCIAL

Byrne y Whiten (1988) plantearon por primera vez a fines de los años ochenta la hipótesis del cerebro social como una explicación al hecho de que los primates tengan cerebros mucho más grandes respecto al tamaño de sus cuerpos que cualquier

otro grupo taxonómico (Jerison hizo esta observación primero en 1973). Casi todo el aumento de tamaño del cerebro se debe a un neocórtex más grande (Finlay & Darlington, 1990; Finlay, Darlington & Nicastro, 2000), y especialmente a un lóbulo frontal más grande (aunque algunas otras áreas subcorticales, como la del cerebelo, están también diferencialmente aumentadas en los humanos: MacLeod y otros, 2003). Los primates dependen de la sociabilidad como herramienta para resolver los problemas ecológicos cotidianos de supervivencia y reproducción exitosa antes que resolver dichos problemas individualmente por ensayo y error. Así que la propuesta fue que las vidas sociales más complejas de los primates les impusieron demandas cognitivas más grandes que en el caso de otras especies no primates.

Desde que esta propuesta fue originalmente sugerida, se han aducido importantes evidencias en su respaldo (para resúmenes recientes, véase a Whiten & Byrne, 1997; Emery, Clayton & Frith, 2007; Dunbar & Schultz, 2007a y 2007b). Uno de los principales hallazgos fue la relación cuantitativa entre el tamaño del grupo social y el tamaño relativo del neocórtex en los primates (gráfico 1). En un diagrama de doble algoritmo, el tamaño promedio del grupo por especie es directamente proporcional al tamaño relativo del neocórtex (indicado como la proporción del volumen del neocórtex respecto al volumen del resto del cerebro). Esto ha sido interpretado como implicando que algún aspecto de la cognición impone una restricción al número de relaciones que un individuo puede mantener como una red social coherente.

Aunque la esencia de la hipótesis del cerebro social en realidad se trata de la complejidad social y sus demandas cognitivas (y ahora hay evidencia importante disponible para mostrar que diversos índices de complejidad de la conducta se correlacionan con el tamaño del neocórtex: véase Dunbar & Schultz, 2007b), es la relación cuantitativa con el tamaño del grupo lo que ha llamado principalmente la atención. Ciertamente, parece que la hipótesis del cerebro social es a menudo vista como sinónimo de este hallazgo clave. Sin embargo, es importante no perder de vista el hecho de que este efecto del tamaño del grupo es realmente solo una propiedad emergente de la relación subyacente que se centra en las demandas cognitivas de la complejidad de la conducta.

Recientemente, esto ha recibido un mayor énfasis mediante nuevos hallazgos surgidos de los intentos por probar la hipótesis del cerebro social en especies no primates. Schultz y Dunbar (2007a y 2007b) han mostrado que, frente a una amplia gama de especies de aves y mamíferos (específicamente carnívoros, ungulados y murciélagos), la hipótesis del cerebro social toma una forma cualitativa más que una cuantitativa. En todos estos grupos taxonómicos, las especies con vínculos de pareja (es decir, monógamos reproductivamente) son las que tienen cerebros desproporcionadamente grandes. Los primates antropoides contrastan de modo impresionante

con este patrón en el que ellos, y solamente ellos, muestran una relación cuantitativa entre el tamaño del grupo y el tamaño del cerebro.

Estos resultados sugieren que las demandas cognitivas de los vínculos de pareja son las que apuntalan el cerebro social (produciendo así el alargamiento inicial del tamaño del cerebro en diversas familias animales), y que los primates antropoides (y quizás uno o dos grupos numéricamente pequeños, como los elefantes y la familia de los caballos) han expandido el uso de estas relaciones con vínculos parecidos a otros miembros de grupos no reproductivos —creando de ese modo la relación cuantitativa que observamos en su caso— (Schultz & Dunbar, 2007). En efecto, han desarrollado lo que bien podemos referir como «amistades» (Smuts, 1985 y Silk, 2002), relaciones que tienen todas las características de vínculos de pareja, pero sin la connotación sexual —aunque quizás esto puede explicar por qué las amistades pueden adquirir tan fácilmente matices sexuales en los humanos—. En primates antropoides, las amistades funcionan como alianzas —coaliciones recíprocas cuyos miembros van al auxilio del otro de manera confiable— (véase por ejemplo, Dunbar 1980 y 1989). Estas cumplen la importante función de mediar entre los individuos y los costos de la vida en grupo —la competencia durante la búsqueda de alimentos, y las formas persistentes de un bajo nivel de acoso, como consecuencia de vivir muy cerca de otros, lo cual produce infertilidad en las hembras tanto en cautiverio como en libertad— (Abbott, Keverne, Moore & Yodyingud, 1986; Dunbar, 1980).

Debido a que el vínculo de pareja a menudo se asocia al cuidado biparental, no es intuitivamente obvio si la carga cognitiva es impuesta por las demandas del vínculo de pareja como tal o por las del cuidado biparental. Las aves nos permiten resolver este asunto porque algunas especies tienen una sin la otra: esta disociación demuestra inequívocamente que es una especie con vínculos de pareja, independientemente de si aplica el cuidado biparental o no, que tiene cerebros grandes. Las aves también enfatizan un segundo punto: son las especies que tienen parejas *de por vida* las que tienen cerebros inusualmente grandes, no aquellas que tienen parejas anuales (es decir, aquellas que encuentran nuevas parejas cada año), lo cual indica que en el costo de las relaciones largas existe algo que es cognitivamente exigente.

Esto plantea dos temas importantes. Uno es por qué los vínculos de pareja (y las relaciones con vínculos parecidos) son tan costosos en términos cognitivos. El otro es la pregunta más difícil de justamente qué son en realidad los vínculos de pareja. Se entiende perfectamente que los vínculos de pareja (y especialmente los vínculos de por vida) implican riesgos importantes para sus miembros: una mala elección de pareja tendrá un enorme efecto en el buen estado del individuo debido a los costos de oportunidad de tener que empezar todo de nuevo en el asunto de encontrar una pareja. Tales circunstancias pueden presentarse ya sea porque una pareja es menos

fértil que otros miembros de la población o porque es menos confiable (ya sea en términos de brindar atención a los pequeños o en términos de infidelidad, o ambos). El hecho de que una mala decisión pueda tener consecuencias catastróficas en el bienestar del individuo —literalmente al reducir su rendimiento reproductivo de por vida a cero en el peor de los casos— impondrá inevitablemente una intensa selección en la capacidad para tomar buenas decisiones acerca de la idoneidad de futuras parejas. No obstante, la necesidad de mantener un alto nivel de coordinación y sincronía de la conducta deben también imponer importantes demandas a la cognición. Esto es particularmente obvio en aquellos casos en los que la pareja tiene que compartir responsabilidades de crianza, como en muchas especies de aves en que un miembro de la pareja debe quedarse en el nido mientras el otro se alimenta. Si la pareja no coordina su conducta efectivamente, el individuo que se queda cuidando los huevos puede eventualmente verse forzado a escoger entre morir de hambre o abandonar los huevos. Indudablemente, los vínculos de pareja son alianzas de colaboración para la reproducción conjunta y sus miembros tienen que estar dispuestos a reconocer y tomar en cuenta los intereses de la pareja. De hecho, deben ser capaces de prever las necesidades de su pareja y considerarlas en la planificación de sus propios requerimientos. Parece probable que esta sea la necesidad específica que debe haber provocado la evolución de aquellas habilidades cognitivas sociales asociadas a la teoría de la mente en los humanos.

Sin embargo, esto sí nos lleva a formular una pregunta seria: simplemente, ¿qué es un vínculo de pareja? Tendemos a reconocer los vínculos de pareja en virtud del hecho que la reproducción conjunta es un componente clave. Pero en realidad, esa ya no es más la experiencia de los animales sino la nuestra. Las especies de aves en pareja tienen vínculos para poder dedicarse al cuidado biparental, generando así beneficios reproductivos importantes. El cuidado biparental es el objetivo máximo, pero para alcanzarlo deben encontrar una solución efectiva para el objetivo próximo de crear y mantener un vínculo de pareja funcional. Su percepción de lo que está implicado (y por consiguiente, los mecanismos cognitivos de respaldo) es más bien lo desconocido intangible. Evidentemente, tenemos dificultad tratando de especificar exactamente lo que está implicado incluso en nuestro propio caso: reconocemos una relación cuando vemos o experimentamos una, pero no tenemos ningún sistema métrico adecuado por el cual podemos definirlo operacionalmente. Las relaciones de este tipo son algo que *sentimos*, no que identificamos directamente.

Quizás todo hasta aquí sea obvio a partir de la documentación de psicología social acerca de las amistades: aquí, varias décadas de investigación han identificado dos dimensiones clave para las relaciones —estar cerca y sentirse cerca (Sternberg, 1997; Berscheid, Snyder & Omoto, 1989)—. El primero de estos es fácil de especificar,

porque simplemente tiene que ver con el tiempo que la pareja pasa juntos (en efecto, proximidad espacial) o la frecuencia de la interacción. El segundo es más difícil, porque tiene que ver con nuestras experiencias emocionales internas y tenemos gran dificultad de expresarlo con palabras. Las intuimos como una forma de cognición «caliente», pero no son tan accesibles a la descripción verbal consciente como lo es la cognición «fría». Y por esa razón, nos enfrentamos a un doble dilema en el caso de los animales, dado que carecemos de los medios para describir las experiencias internas en otras especies. Es posible que ya haya llegado el momento de lidiar directamente con el polémico asunto de las experiencias mentales de los animales: es posible que ya no podamos empujarlo lentamente debajo del tapete conductista si queremos progresar en la comprensión de la naturaleza de las relaciones entre humanos u otros animales.

3. LOS VÍNCULOS QUE UNEN

La hipótesis del cerebro social, entonces, sostiene que conservar y cuidar los tipos de relaciones intensas encontradas en los vínculos de pareja (en la mayoría de aves y mamíferos) y amistades (entre los primates antropoides) implica un serio trabajo cognitivo que a su vez se refleja en el tamaño del cerebro de las especies. Pero es, al mismo tiempo, aparente, tanto desde la literatura de la psicología social humana acerca de las amistades como desde la literatura etológica acerca de cómo los primates cuidan sus relaciones, dicha cognición es solo parte de la historia. Existe un componente profundamente emocional en las relaciones que proviene, en ambos casos, de hacer cosas juntos. Esto nos lleva, quizás inevitablemente, al asunto de cómo los primates unen sus grupos sociales.

Para los primates antropoides, y en particular los monos y simios del Mundo Antiguo, el acicalamiento social es el principal mecanismo utilizado para el vínculo social. El acicalamiento es una actividad intensa en la cual un individuo hurga en el pelaje del otro, retirando trozos de plantas, piel seca y otros desechos. Es una actividad muy personal, en la cual el acicalador normalmente se encuentra profundamente concentrado en su tarea. El acicalamiento disminuye el ritmo cardíaco en el acicalado y reduce la frecuencia de los signos de tensión y estrés (bostezos, rasguños, etcétera) (Goosen, 1981; Castles, Whiten & Aureli, 1999), hasta el punto en que el receptor del acicalamiento puede relajarse tanto que se queda dormido. Muchas especies dedican considerables partes de su día al acicalamiento social —en el mayor de los casos, hasta casi una quinta parte del tiempo que pasan despiertos—. El acicalamiento tiene muchas características similares al masaje: es físicamente estimulante y ligeramente doloroso, y por tanto, produce la liberación de endorfinas (los analgésicos propios

del cerebro) (Keverne, Martensz & Tuite, 1989). Las endorfinas probablemente son las responsables de estos efectos soporíferos.

Empero, su verdadero significado parece ser que estos efectos psicofarmacológicos producidos por el acicalamiento juegan un papel crucial en los procesos de generar la confianza y la reciprocidad que forman la base de las relaciones sociales del primate. No sabemos a ciencia cierta cómo las endorfinas dan origen a las relaciones, pero una posibilidad es simplemente permitir que dos individuos se sientan lo suficientemente relajados en compañía del otro como para pasar tiempo juntos. El tiempo es un bien que debemos invertir para crear un vínculo, y parece haber más o menos una relación directa entre el tiempo vivido juntos y la fortaleza de un vínculo —al menos en términos de consecuencias funcionales, tales como la disposición para venir al auxilio de la pareja que acicala— (véase Dunbar, 1980, 1984 y 1989). Es posible que las endorfinas simplemente brinden un reforzador cercano que hace que valga la pena pasar tiempo con alguien más, o es posible que estas neuroendocrinas estén íntimamente involucradas en el proceso de la vinculación. De cualquier modo, parece que su papel es fundamental.

4. LOS HUMANOS Y EL CEREBRO SOCIAL

Los humanos modernos tienen cerebros mucho más grandes (y especialmente los neocórtex) que otros primates, y legítimamente podemos preguntar qué nos puede decir la relación que existe entre el tamaño del neocórtex y el tamaño de los grupos de primates acerca de los tamaños de los grupos humanos. Tal como se aprecia en el gráfico 1, hay grados bastante diferenciados en esta relación dentro de los primates: los simios se encuentran a la derecha de los monos, y los monos a la derecha de los prosimios, lo que indica que, a medida que se va de los prosimios hasta los simios, pasando por los monos, el cuidado de los grupos de un tamaño dado requiere proporcionalmente una mayor capacidad computacional (Dunbar, 1998). Es así que la línea de regresión apropiada desde la que se deben prever los tamaños de los grupos humanos es la de los simios. La interpolación de la proporción del neocórtex humano moderno a la ecuación de los simios resulta en un tamaño de grupo previsto de ~150 (gráfico 1).

Una investigación de la literatura etnográfica arrojó que ese es en realidad el tamaño típico de las comunidades cazadoras-recolectoras (Dunbar, 1993; Hamilton y otros, 2007). Quizá lo más sorprendente sea que esta cifra de ~150 aparece frecuentemente en muchos aspectos de la organización humana histórica y contemporánea (cuadro 1). Fue el tamaño promedio por aldea registrado para casi todos los condados ingleses en el *Libro de Domesday*, así como durante el siglo XVIII,

y es el tamaño típico de la compañía de la mayoría de ejércitos modernos, el número de destinatarios de una típica lista de reparto de tarjetas de Navidad en Gran Bretaña, y el tamaño de la red social en los experimentos inversos del «mundo pequeño», entre otros. Así, a pesar de las diferencias marcadas en escala y organización, una amplia gama de fenómenos sociales contemporáneos parecen coincidir mucho con los mismos tipos de patrones de agrupación. La única diferencia sustancial entre las redes sociales de las sociedades cazadoras-recolectoras y agrícolas tradicionales, y las sociedades posindustriales modernas parece ser que, en las sociedades tradicionales, todos los miembros de la comunidad tienen más o menos la misma red de 150 conocidos, mientras que en las sociedades urbanas modernas nuestras redes están bastante fragmentadas —mis 150 están conformados por un grupo de subredes que apenas coinciden entre sí—. Usted y yo podríamos tener en común un pequeño grupo de amigos, digamos del trabajo, pero no existe coincidencia alguna con los subgrupos restantes —no tenemos los mismos parientes, ni compartimos círculos de aficionados, grupos religiosos, amigos de esposos, amigos de la puerta del colegio (las frecuentes amistades temporales que se hacen a través de los amigos de colegio de nuestros hijos), o amigos de clubes deportivos—. Las redes en las sociedades modernas están fragmentadas y dispersas (a menudo a distancias geográficas considerables), mientras que en las sociedades tradicionales forman típicamente una comunidad unificada simple —a pesar de que dicha comunidad puede estar distribuida en una amplia área geográfica (igual que en muchas comunidades cazadoras-recolectoras contemporáneas)—.

La cifra de ~150 parece marcar un límite perceptible en cuanto a la calidad de la relación: parece haber una diferencia marcada en la calidad de las relaciones que tenemos con quienes se encuentran dentro del círculo escogido versus quienes están afuera. Mi definición informal para este límite a nuestro mundo social consiste en todos a quienes conocemos como personas, todos aquellos con quienes tenemos una relación personal definible. Quienes están dentro de este círculo son personas por quienes tenemos algún sentido de obligación, en quienes confiamos que nos ayudarían si así se los pidiésemos, con quienes sería recíproco nuestro sentido de compromiso personal. Nosotros sabemos dónde encajan estas personas en nuestra red de relaciones, ellas saben dónde encajamos nosotros dentro de las suyas, y en ambos casos lo sabemos porque nos conocemos personalmente. Algunas veces, ese conocimiento puede ser indirecto (amigos de amigos, o un abuelo en común), pero define a quienes les debemos compromisos personales; si los ofendemos o rechazamos de alguna manera, dicha ofensa nos va a atormentar por el efecto que tiene en las relaciones que nos une. Por el contrario, al superar este círculo de 150, las personas dejan de ser individuos, por lo menos en lo que respecta a nuestras relaciones.

Aunque los reconocemos como individuos (es decir, les ponemos nombres a sus rostros), nuestras relaciones con ellos son menos personales y más tipológicas. Necesitamos reglas prácticas para guiar nuestras interacciones con ellos y no depender de conocimientos personalizados. En esos casos, la regla normalmente la marca algún distintivo pertinente que indica el estatus de un individuo y cómo dirigirnos a él, y esto casi siempre requiere que lo identifiquemos formalmente con algún distintivo para reconocerlo —uniformes, insignias de grados, tipos de discurso, etcétera—.

Así como con todos los grupos sociales de primates, las redes sociales humanas son bastante estructuradas. No interactuamos por igual con todos los miembros de nuestro mundo social inmediato. Más bien, parece que nuestro mundo social está compuesto por una serie de círculos de conocidos incluidos jerárquicamente, que se ven reflejados tanto en la intimidad percibida de la relación como en la frecuencia de la interacción (Hill & Dunbar, 2003). Estos círculos de conocidos parecen tener una estructura muy consistente: cada anillo incluye casi el doble de gente de la que hay en el anillo inmediatamente dentro de aquel, de modo que los números acumulativos de personas incluidas en los círculos sucesivos muestran una proporción de escala constante de aproximadamente 3 (Zhou, Sornette, Hill & Dunbar, 2005; véase también Hamilton y otros, 2007). En términos generales, incluyen progresivamente a 5, 15, 50, 150, 500 y 1500 individuos (véase Dunbar, 1992; Zhou y otros, 2005) y, por lo que sabemos, podrían extenderse incluso más allá en otra serie de círculos con las mismas proporciones.

5. EL PAPEL DE LA COGNICIÓN

El hecho de que el tamaño del cerebro tenga correlación con el tamaño del grupo social implica que existe un límite cognitivo. Sin embargo, es sorprendente lo poco que sabemos sobre los tipos de cognición que podrían estar involucrados en la gestión de las relaciones sociales. Si bien probablemente todos están de acuerdo con que esto es en cierto modo una forma de «cognición social», gran parte de lo que eso supone sigue siendo confuso. El único aspecto sobre el que sabemos más es aquello que se conoce como teoría de la mente (Leslie, 1987 y Perner, 1991). La teoría de la mente es la capacidad de reflexionar sobre los estados mentales de otro individuo. Así, se trata de un nivel en una serie reflexiva potencialmente infinita de estados mentales y creencias sobre estados mentales conocidos como los niveles de la intencionalidad (Dennett, 1989). Se sabe muchísimo sobre la teoría de la mente (intencionalidad del segundo orden) pues los psicólogos del desarrollo han investigado sobre ello con notable meticulosidad. En términos sencillos, se trata

del Rubicón cognitivo que los niños cruzan aproximadamente a la edad de cuatro a cinco años, aunque algunos individuos (autistas) nunca lo logran incluso de adultos (Baron-Cohen, Leslie & Frith, 1985). Sin embargo, el problema con la teoría de la mente resulta ser que conocemos mucho sobre su historia natural, pero como Roth y Leslie (1998) lo han señalado, casi no tenemos idea de lo que esto realmente implica en términos cognitivos.

Sin embargo, a pesar de que los procesos exactos involucrados pueden ser un tanto opacos, quizás podemos usar la noción de intencionalidad para que nos ayude a saber algo sobre el problema de cómo los humanos se diferencian de otros primates, ya que los órdenes de intencionalidad forman una escala natural, y quizás así obtengamos un índice de competencia cognitiva social. La afirmación de que la intencionalidad misma no es más que algún aspecto de la función ejecutiva (Ozonoff, 1995; Mitchell, 1997; Barrett, Henzi & Dunbar, 2003; Stylianou, 2007) parece brindar una justificación para ello. Mi inquietud aquí no es tanto acerca de las discusiones sobre si la teoría de la mente (o más generalmente, la intencionalidad) es o no modular o el resultado de la función ejecutiva (aunque me inclino más por esto último), sino acerca de la afirmación más simple de que la escala de intencionalidad parece brindarnos un parámetro de competencia cognitiva social (tal como lo cataloga la capacidad de retener en la mente varios estados mentales de individuos al mismo tiempo).

De ser así, nuestro principal interés al respecto es saber cuáles podrían ser los límites naturales del razonamiento intencional en los humanos. Hemos evaluado a adultos normales en varios estudios individuales, y parece que los límites de la función en el caso de los adultos son consistentemente del quinto orden («Yo *creo* que *supones* que yo *imagino* que tú *quieres* que yo *crea* que...») (Kinderman, Dunbar & Bentall, 1998; Stiller & Dunbar, 2007). Aproximadamente dos tercios de los individuos tienen su límite en la intencionalidad del quinto orden o por debajo de este, mientras que tres cuartos tienen su límite en el sexto orden o por debajo de este. Existe una variación individual considerable al respecto (Stiller & Dunbar, 2007), y hemos demostrado que los niveles más altos se pierden durante la fase activa de por lo menos dos condiciones clínicas muy conocidas (desorden bipolar: Kerr, Dunbar & Bentall, 2003; esquizofrenia paranoide: Swarbrick, 2000). Estas competencias se desarrollan en el periodo de tiempo que transcurre entre los cinco años (cuando los niños recién adquieren la teoría de la mente o la intencionalidad del segundo orden) y el inicio de la adolescencia (cuando finalmente adquieren competencias adultas de quinto orden) (Henzi y otros, 2007).

6. LA INTENCIONALIDAD Y EL MUNDO VIRTUAL

Aquí resulta interesante saber qué se puede lograr con los diferentes niveles de la intencionalidad. Si las competencias intencionales nos permiten tener en cuenta diversos estados mentales individuales a la vez, es probable que se impongan restricciones a los fenómenos culturales que requieran que pensemos de manera intencional. Esto tal vez sea más evidente en el caso de los juegos imaginativos. Leslie (1987) señaló que la teoría de la mente puede ser crucial para que los niños puedan emprender juegos ficticios (por ejemplo, representar roles) en los que deben imaginar que el mundo es diferente a como realmente es (las muñecas pueden tomar el té, por ejemplo, o el timón detrás de una silla es un auto de verdad). La teoría de Leslie se puede extender al teatro. Consideremos el caso de la audiencia que ve *Otelo* de Shakespeare. La audiencia tiene que creer que Yago pretende que Otelo imagine que Desdémona está enamorada de Casio, una actividad que implica cuatro niveles de intencionalidad. Sin embargo, debe notarse que, a estas alturas, el tipo de historia con el que están lidiando no es precisamente exigente (o, es más, particularmente fascinante). ¿Por qué debería importarles a Otelo que Desdémona fantasee con Casio? En resumidas cuentas, en la vida diaria solo unos cuantos mostraríamos cierta perplejidad ante tal trivial fenómeno, y la historia terminaría ahí como un relato aburrido. En la obra de Shakespeare resulta atractivo que Yago sea capaz de persuadir a Otelo de que Casio corresponde los sentimientos de Desdémona, creando así un triángulo amoroso y redoblando las apuestas para que todos nos enganchemos con el drama (especialmente cuando, teniendo la ventaja del punto de vista del espectador, somos testigos del plan intrigante de Yago). A estas alturas, por supuesto, la audiencia trabaja en la intencionalidad del quinto orden, encontrándose así al límite natural de la gran mayoría de la población.

Pero al crear esta historia, el mismo Shakespeare debe estar en un nivel superior al de su audiencia, el sexto: su *intención* debe ser que la audiencia *crea*... Sugiero que esto puede explicar por qué la capacidad de disfrutar buena literatura es un universal humano generalizado, pero la capacidad de *componer* buena literatura no —la narración de historias requiere competencias sociales cognitivas que trascienden el rango normal de la gran mayoría de la población—. Así, al sentarnos a escribir esas novelas a las que tanto hemos aspirado escribir, nuestros límites naturales de la intencionalidad del quinto orden hacen que la mayoría de nosotros escriba historias aburridas.

Podemos usar el mismo marco para explorar las exigencias cognitivas de la religión, ya que esta también requiere que trabajemos con un mundo imaginado —un mundo que no podemos ver o sentir directamente puesto que solo existe en nuestras mentes—. Las experiencias trascendentales que vivimos durante los eventos religiosos

son sin duda alguna bastante reales, pero no son experiencias físicas reales creadas por el mundo que vulnera nuestros sentidos. Pertenecen, más bien, a un mundo virtual interno de la imaginación.

Un análisis de las exigencias al mentalizar las creencias religiosas sugiere que es perfectamente posible tener una religión con cualquier orden de intencionalidad; sin embargo, la *forma* de la religión depende de los niveles de intencionalidad en los cuales uno puede trabajar (cuadro 2). Parece haber una diferencia crítica entre las formas de religión posibles en las intenciones del cuarto y quinto orden. En el cuarto orden, puede existir lo que llamo una religión social: puedo imponerles a creer lo que yo creo como cierto sobre los intereses de alguna deidad respecto a nosotros, pero ustedes no tienen por qué estar de acuerdo con esa postura a pesar de aceptar que estoy convencido de la verdad que yo creo: no hay coacción alguna para que ustedes crean. Solo podría hacer que estén de acuerdo conmigo si usara la fuerza policial (o del tipo secular o celestial). Pero en ese caso, cabe la posibilidad de que su nivel de compromiso personal e intelectual con el proyecto sea muy limitado. Sin embargo, el quinto orden parece ser otro Rubicón crucial: en el quinto orden, parece que cuando acepto que ustedes crean que esta postura es verdadera, me comprometo *ipso facto* con la veracidad de su postura. Estamos ahora frente a lo que yo llamo «religión común» —un conjunto de creencias que nos unen en una sola comunidad cuyos miembros comparten la misma visión del mundo—. En este nivel, tenemos un mecanismo muy poderoso que compele a la voluntad comunal, que nos hace adherirnos a ese proyecto común (sea lo que eso sea). No necesitamos una fuerza policial para comportarnos de una manera religiosa: todos estamos profunda y personalmente comprometidos con ese proyecto y cumplimos con estas creencias según nuestro libre albedrío.

7. ¿QUÉ NOS HACE TAN DIFERENTES?

Aquí resurge mi pregunta inicial: ¿por qué los humanos no son solo grandes simios? He sugerido que la respuesta yace en nuestra capacidad de vivir en un mundo mental virtual. Podemos participar en actividades que van mucho más allá de las competencias de incluso los grandes simios. En efecto, están tan lejos de las competencias de otras especies que no cabe la posibilidad de que el proverbial chimpancé sentado frente a una máquina de escribir pueda producir las obras de Shakespeare en un tiempo razonable (y aquí me refiero a una escala de milenios). La pregunta pendiente es por qué los humanos hemos necesitado capacidades cognitivas así de extraordinarias.

Desde una perspectiva evolutiva, esto resulta particularmente enigmático porque el tejido cerebral es excepcionalmente costoso. Aiello y Wheeler (1995) señalaron hace algún tiempo que el tejido cerebral es una de las materias del cuerpo más caras de mantener, y que ofrece una pendiente muy empinada por la cual la selección natural debe orientar a la evolución cerebral en caso de que el tamaño del cerebro debiera aumentar (se trata de su «hipótesis del tejido caro»). Existe cierta evidencia comparativa tentativa (Dunbar 2003a y 2003b) —considerando el soporte adicional que permiten recientes estudios de neuroimágenes humanas que aún no han sido publicados— por la cual las competencias intencionales guardan correlación con algunos aspectos del volumen cerebral. En ese caso, sucede que la presión de selección de la capacidad que controla los niveles superiores de la intencionalidad incurre en un costo energético muy importante para los sujetos involucrados. Las ventajas que se obtienen al invertir en tales capacidades deben entonces ser considerables. ¿Cuáles podrían ser?

Las sociedades primates son contratos sociales implícitos. Como todos los contratos sociales, su estabilidad y funcionalidad dependen del intercambio de beneficios a corto y largo plazo entre sus miembros. Al igual que los vínculos de pareja, un grupo se mantendrá estable solo como una coalición siempre que sus miembros estén dispuestos a comprometer algunos de los beneficios de corto plazo a fin de sacar provecho a largo plazo mediante los efectos en el nivel grupal. Estos efectos aparecen en la forma de una mayor supervivencia, y por ende mayores índices reproductivos, principalmente como resultado de los menores índices de depredación (Dunbar, 1988; Shultz, Nöe, McGraw & Dunbar, 2004), aunque también podría haber beneficios más directos en términos de la supervivencia de la descendencia individual (Silk, Alberts & Altmann, 2003; Silk, 2007).

No obstante, todos los contratos sociales de este tipo se enfrentan al mismo problema: los oportunistas. Siempre hay un beneficio significativo que obtienen los individuos que sacan provecho de la sociabilidad pero que no pagan todos los costos (Enquist & Leimar, 1993; Nettle & Dunbar, 1997; Dunbar, 1999). Ya que la intromisión de los oportunistas es proporcional al tamaño de la población, las grandes comunidades que caracterizan a los humanos se enfrentan a un desafío mucho mayor en comparación con los desafíos de nuestros primos primates. En consecuencia, se necesitan mecanismos más sofisticados a fin de garantizar que los oportunistas no agobien a la comunidad. Si bien existen numerosos mecanismos cognitivos para controlar a los oportunistas (Dunbar, 1999), la falta de acicalamiento a gran escala significa que inevitablemente no contamos con mecanismos basados en las endorfinas que yacen en los orígenes de la sociabilidad primate. Este es un tema serio,

puesto que crea un «vacío emocional» de proporciones bastante importantes (Dunbar, en prensa), que podría desestabilizar seriamente la frágil base de la que depende de la colaboración a nivel grupal. Algo debía llenar ese vacío.

En realidad, el patrón de la evolución del tamaño cerebral sugiere que este tema se tornó serio en las etapas finales de la evolución humana, luego de la aparición de los humanos arcaicos (la especie *Homo heidelbergensis* y sus aliados) hace aproximadamente medio millón de años (Dunbar, en prensa). En ese momento, el tamaño del cerebro aumentó rápidamente, lo cual implicó el respectivo incremento del tamaño de las comunidades que debían vincularse entre sí. Sin embargo, nuestra capacidad de crear las grandes comunidades que supuestamente se necesitaban para garantizar la supervivencia (y las razones por las cuales necesitamos dichas comunidades siguen sin esclarecerse) debió depender de la solución satisfactoria del tema de la vinculación. Si bien la conformidad social y la «buena conducta» pueden siempre ser impuestas mediante una acción punitiva sobre los reincidentes (Orstrom, Gardner & Walker, 1994; Clutton-Brock & Parker, 1995; Fehr, Fishbacher & Gächter, 2002), existen límites en los que dicha acción puede verdaderamente hacer cumplir la cooperación social. Puesto que siempre hay beneficios que se obtienen del hecho de ser oportunistas y cabe la gran posibilidad de que cualquier oportunista pase desapercibido y/o se libre de un castigo, la tentación de serlo es tal que parte de los individuos también lo serán, independientemente de cuán efectivo sea el castigo para reducir su frecuencia absoluta (para referirse a un problema idéntico, relacionado a depredación y conservación, véase Cowlshaw & Dunbar, 2000). La cooperación social es mucho más efectiva cuando los sujetos actúan libremente, pues de manera voluntaria se adhieren al proyecto común (Orstrom, Gardner & Walker, 1994). Los mecanismos que crean un sentido de comunidad tienen, entonces, más probabilidades de producir los niveles necesarios de altruismo que el uso del control social.

Como se ha observado por mucho tiempo en las ciencias sociales, tanto la religión como la narración de historias cumplen una función importante en la vinculación social de todas las culturas humanas (Durkheim, 2001). La religión lo hace a través de la capacidad que tienen los rituales para desencadenar la liberación de endorfinas, pues en muchos casos se trata de actividades solo ligeramente estresantes que son especialmente buenas para liberar endorfinas. La religión, ciertamente, también tiene la ventaja de contar con una dimensión intelectual, y aquí las demandas cognitivas cobran importancia al crear el mecanismo de vinculación de proceso dual que hallamos en el acicalamiento social primate. En la medida en que el componente intelectual de la religión sea una forma de narración, la narración misma ingresa al marco como un importante mecanismo para la vinculación de la comunidad. Sin embargo, la narración de historias sobrepasa las explicaciones meramente religiosas

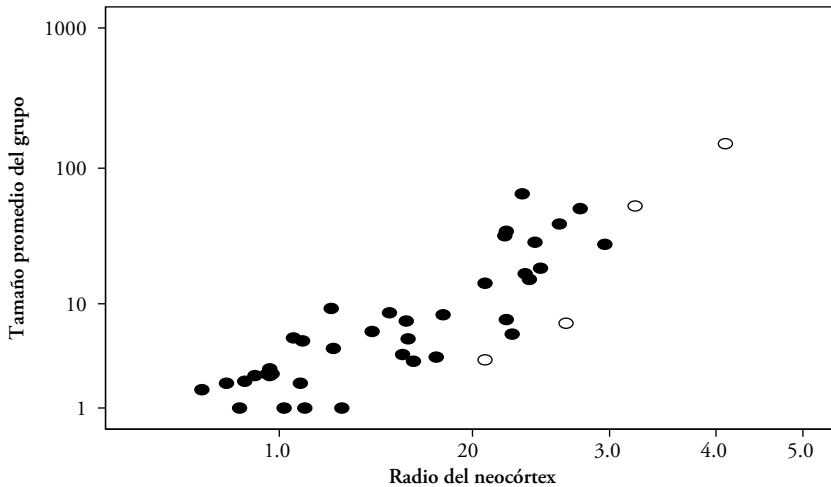
sobre cómo es y debería ser el mundo; nos brinda su propia oportunidad de desarrollar una forma de entretenimiento que puede ser un mecanismo de vinculación tan poderoso por méritos propios.

Uno podría argüir que ambos son meros epifenómenos —productos secundarios no adaptativos del hecho de que poseemos grandes cerebros—. Esta misma afirmación se ha hecho, obviamente, en el caso de la música («la música es un *cheesecake* evolutivo», manifestó Pinker [1998]). Sin embargo, tal afirmación resulta poco compatible con el tiempo, el dinero y el esfuerzo que, independientemente de la cultura, los humanos invierten en los tres fenómenos. En efecto, todas las sociedades humanas valoran la narración de historias en sí misma como una forma de entretenimiento, y muchas de estas historias están íntimamente implicadas en la creación de un sentido de comunidad: relatos sobre el origen, historias tribales, y fábulas son algunos de los «relatos de fogata», y todos contribuyen directamente a brindar un sentido de comunidad. Es más, su mera representación a menudo tiene una contribución directa, en particular a través de la risa (otro poderoso liberador de endorfinas: Dunbar, 2004). A modo de confirmación, Van Vugt, Hardy, Stow & Dunbar (s/f) han demostrado recientemente de manera experimental que la risa hace que los extraños (no los amigos existentes) sean significativamente más generosos con los demás en juegos de bienes públicos. Es importante resaltar que tanto la religión como la narración de historias adquieren sus propiedades de reforzamiento de la comunidad a través del lenguaje: sin el lenguaje no sería posible contar una «gran historia» a fin de crear efectos vinculantes o persuadir a las personas para ser parte de rituales religiosos de forma regular. Parece que es esencial la participación regular en estos ejercicios de vinculación con la comunidad, casi como si fuere una forma de inoculación que requiere refuerzos a intervalos fijos para mantener un nivel de desempeño consistente.

Pero, además de la respuesta trivialmente obvia del lenguaje, ¿qué es lo que limita estas «actividades de la mente» únicamente a los humanos? Una respuesta tiene que ser que solo los humanos pueden sobrellevar la intencionalidad del quinto orden que es necesario para permitir que estos fenómenos produzcan los efectos que tienen. Lo máximo que cualquier especie no humana puede hacer parece ser la intencionalidad del segundo orden —e incluso eso sea probablemente cierto solo en los grandes simios— (O’Connell & Dunbar, 2003), ya que el consenso indica que todas las otras especies animales pueden aspirar solo al primer orden. Si las habilidades de mentalización realmente son una propiedad emergente de las competencias de la función ejecutiva (Barrett, Henzi & Dunbar, 2003) y estas a cambio son una función de las facultades de cálculo (y por ende del tamaño) del cerebro (Dunbar, 2003b), la explicación simple para esta sorprendente diferencia entre nosotros y todas las otras especies yace en el tamaño de nuestro cerebro (y, quizás más específicamente,

los lóbulos frontales, pues generalmente se considera que son los *locus operandi* de aquellas capacidades a las que convencionalmente nos referimos como funciones ejecutivas). En efecto, las diferencias entre nosotros y todos los otros animales se resumen en el hecho de que el tamaño de nuestro cerebro nos permite hacer algo que simplemente no es posible hacer con un cerebro más pequeño. Si bien la organización cerebral y los aspectos de la eficiencia neural (véase a Burki & Kaessmann, 2004, por ejemplo) deben, ciertamente, jugar un papel, existen buenos argumentos para pensar que lo que importa es simplemente el tamaño de la computadora (véase a Duncan, 2001, por ejemplo). Sean cuales fueren las diferencias de estructura y organización, sí existe un asunto de tamaño que debemos explicar y nuestros cerebros de gran tamaño no pueden estar ahí accidentalmente nada más o como un producto colateral trivial de algo más. En resumen, la diferencia entre nosotros y nuestros primos más cercanos no pasa simplemente por tener una mayor inteligencia (como sea que queramos medir eso), sino lo que esa mayor inteligencia nos permite hacer (principalmente, vivir en un mundo virtual), y por qué necesitamos esa capacidad emergente en lo más mínimo (es decir, permitir relacionarnos en comunidades mucho más grandes de lo que le sería posible a un mono o un simio).

Gráfico 1



Cuadro 1. Ejemplos de agrupaciones sociales humanas que corresponden al tamaño calculado de ~150 sujetos²

Agrupación	Tamaño regular	Fuente
Pueblos neolíticos (Medio Oriente, 6500-5500 a. C.)	150-200	Oates (1977)
Manípulo («dos centurias») (Ejército romano: 350-100 a. C.)	120-130	Montross (1975)
Libro Domesday (1085): Tamaño promedio de un pueblo	150	Hill (1981) y Bintliff (1999)
Pueblos ingleses del siglo XVIII (promedio de los medios del condado)	160	Laslett (1971)
Sociedades tribales (promedio y rango de comunidades; $N=9$)	148 (90-222)	Dunbar (1993)
Sociedades cazadoras-recolectoras (tamaño promedio del clan; N)	165	Hamilton y otros (2007)
Comunidades agrícolas Hutterite (Canadá) (promedio, $N=51$)	107	Mange y Mange (1980)
Distritos Amish de «Nebraska» (promedio, $N=8$)	113	Hurd (1985)
Congregaciones eclesíásticas (tamaño recomendado ideal)	200	Urban Church Project (1974)
Comunidad rural montañosa de E. Tennessee	197	Bryant (1981)
Tamaño de red social (promedio, $N=2$ experimentos del «mundo pequeño»)	134	Killworth, Bernard y McCarty (1984)
Goretex Inc: tamaño de la unidad de una fábrica	150	Gladwell (2000)
Compañía (promedio y rango de los 10 ejércitos de la Segunda Guerra Mundial)	180 (124-223)	MacDonald (1955)
Listas de distribución de tarjetas de Navidad (promedio de todos los receptores: $N=43$)	154	Hill y Dunbar (2003)
Especialidades de investigación (ciencias y humanidades) (modo, $N=13$)	100-200	Becher (1989)

² Los intervalos de confianza alrededor del promedio calculado son 100-200 (Dunbar, 1993).

Cuadro 2. Formas de creencia religiosa posibles gracias a los diferentes niveles de la intencionalidad³

Intencionalidad Nivel	Posibles afirmaciones de creencia	Forma de religión
1 ^{er}	Yo <i>creo</i> que Dios [...existe]	ninguna
2 ^{do}	Yo <i>creo</i> que Dios está <i>dispuesto</i> [...a intervenir si no obedeces sus leyes]	hecho supernatural
3 ^{er}	Yo <i>deseo</i> que tú <i>creas</i> que Dios está <i>dispuesto</i> [...a intervenir...]	religión personal
4 ^{to}	Yo <i>deseo</i> que tú <i>creas</i> que yo <i>quiero</i> que Dios esté <i>dispuesto</i> [...a intervenir...]	religión social
5 ^{to}	Yo <i>deseo</i> que tú <i>creas</i> que Dios <i>entiende</i> que yo <i>quiero</i> que esté <i>dispuesto</i> [...a intervenir...]	religión común

REFERENCIAS BIBLIOGRÁFICAS

- Abbott D. H., E. B. Keverne, G. F. Moore & U. Yodyingvad (1986). Social Suppression of Reproduction in Subordinate Talapoin Monkeys, *Miopithecus talapoin*. En James G. Else & Phyllis C. Lee (eds.), *Primate Ontogeny* (pp. 329-341). Cambridge: Cambridge University Press.
- Aiello, Leslie & Peter Wheeler (1995). The expensive tissue hypothesis: the brain and the digestive system in human evolution. *Current Anthropology*, 36, 199-221.
- Baron-Cohen, Simon, Alan Leslie & Uta Frith (1985). Does the Autistic Child Have a Theory of Mind? *Cognition*, 21, 37-46.
- Barrett, Louise, Peter Henzi & Robin Dunbar (2003). Primate Cognition: From «What Now?» to «What If?». *Trends in Cognitive Sciences*, 7, 494-497.
- Becher, Tony (1989). *Academic Tribes and Territories*. Open University Press: Milton Keynes.
- Berscheid, Ellen, Michael Snyder & Allen M. Omoto (1989). The Relationship Closeness Inventory: Assessing the Closeness of Interpersonal Relationships. *Journal of Personality and Social Psychology*, 57, 792-807.
- Bintliff, John (1999) Settlement and Territory. En Graeme Barker (ed.), *Companion Encyclopedia of Archaeology* (pp. 505-545). Londres: Routledge.
- Bryant, Carlene (1981). *We're All Kin: A Cultural Study of a Mountain Neighborhood*. Knoxville (TN): University of Tennessee Press.
- Burki, Fabien & Henrik Kaessmann (2004). Birth and Adaptive Evolution of a Hominoid Gene that Supports High Neurotransmitter Flux. *Nature Genetics*, 36, 1061-1063.

³ Cuadro realizado por el autor.

- Byrne, Richard & Andrew Whiten (eds.) (1988). *Machiavellian Intelligence*. Oxford: Oxford University Press.
- Castles, Duncan L., Andrew Whiten & Filippo Aureli (1999). Social Anxiety, Relationships and Self-Directed Behaviour among Wild Female Olive Baboons. *Animal Behaviour*, 58, 1207-1215.
- Clutton-Brock, Tim H. & Geoff A. Parker (1995). Punishment in Animal Societies. *Nature*, 373, 209-216.
- Cowlishaw, Guy & Robin Dunbar (2000). *Primate Conservation Biology*. Chicago: Chicago University Press.
- Dennett, Daniel (1987). *The Intentional Stance*. Cambridge: The MIT Press.
- Dunbar, Robin (1980). Determinants and Evolutionary Consequences of Dominance among Female Gelada Baboons. *Behavioral Ecology and Sociobiology*, 7, 253-265.
- Dunbar, Robin (1984). *Reproductive Decisions: An Economic Analysis of Gelada Baboon Social Strategies*. Princeton: Princeton University Press.
- Dunbar, Robin (1988). *Primate Social Systems*. Londres: Chapman & Hall.
- Dunbar, Robin (1989). Reproductive Strategies of Female Gelada Baboons. En Anne Rasa, Christian Vogel y Eckart Voland (eds.), *Sociobiology of Sexual and Reproductive Strategies* (pp. 74-92). Londres: Chapman & Hall.
- Dunbar, Robin (1993). Coevolution of Neocortex Size, Group Size and Language in Humans. *Behavioral and Brain Sciences*, 16, 681-735.
- Dunbar, Robin (1998). The Social Brain Hypothesis. *Evolutionary Anthropology*, 6, 178-190.
- Dunbar, Robin (1999). Culture, Honesty and the Freerider Problem. En Robin Dunbar, Chris Knight & Camilla Power (eds.), *The Evolution of Culture* (pp. 194-213). Edinburgo: Edinburgh University Press.
- Dunbar, Robin (2003a). The Social Brain: Mind, Language and Society in Evolutionary Perspective. *Annual Review of Anthropology*, 32, 163-181.
- Dunbar, Robin (2003b). Why are Apes so Smart? En Peter Kappeler y Michael Pereira (eds.), *Primate Life Histories and Socioecology* (pp. 285-298). Chicago: Chicago University Press.
- Dunbar, Robin (2004). Language, Music and Laughter in Evolutionary Perspective. En D. Kimbrough Oller y Ulrike Griebel (eds.), *Evolution of Communication Systems: A Comparative Approach* (pp. 257-274). Cambridge: The MIT Press.
- Dunbar, Robin (2007a). Cognitive Constraints on the Structure and Dynamics of Social Networks. *Group Dynamics*, 12, 7-16.
- Dunbar, Robin (2007b). Mind the Gap or Why Humans Aren't just Great Apes. En VV. AA., *Anales de la Academia Británica*. Volumen 154 (lectura 10). Oxford: Academia Británica.

- Dunbar, Robin (2009). Why only Humans Have Language. En Rudolf Botha y Chris Knight (eds.), *The Prehistory of Language* (pp. 12-35). Oxford: Oxford University Press.
- Dunbar, Robin (2012). Mind the Bonding Gap: Constraints on the Evolution of Hominin Societies. En Stephen Shennan (ed.), *Pattern and Process in Cultural Evolution* (páginas). Berkeley: University of California Press.
- Dunbar, Robin & Susanne Shultz (2007a). Evolution in the Social Brain. *Science*, 317, 1344-1347.
- Dunbar, Robin & Susanne Shultz (2007b). Understanding Primate Brain Evolution. *Philosophical Transactions of the Royal Society of London*, 362B, 649-658.
- Duncan, John (2001). An Adaptive Coding Model of Neural Function in Prefrontal Cortex. *Nature Neuroscience*, 2, 820-829.
- Durkheim, Émile (2001[1900]). *Elementary Forms of Religious Life*. Oxford: Oxford University Press.
- Emery, Nathan, Nicola Clayton & Chris Frith (2007). *Social Intelligence: From Brain to Culture*. Oxford: Oxford University Press.
- Enquist, Magnus & Olof Leimar (1993). The Evolution of Cooperation in Mobile Organisms. *Animal Behaviour*, 45, 747-757.
- Fehr, Ernst, Urs Fischbacher & Simon Gächter (2002). Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms. *Human Nature*, 13, 1-25.
- Finlay, Barbara & Richard Darlington (1995). Linked Regularities in the Development and Evolution of Mammalian Brains. *Science*, 268, 1578-1584.
- Finlay, Barbara, Richard Darlington & Nicholas Nicastro (2001). Developmental Structure in Brain Evolution. *Behavioral and Brain Sciences*, 24, 263-308.
- Gladwell, Malcolm (2000). *The Tipping Point*. Londres: Little, Brown and Co.
- Goosen, Cornelis (1981). On the Function of Allogrooming in Old-World Monkeys. En Anton Bruneto Chiarelli y Robert S. Corruccini (eds.), *Primate Behaviour and Sociobiology* (pp. 110-120). Berlín: Springer.
- Hamilton, Marcus J. y otros (2007). The Complex Structure of Hunter-Gatherer Social Networks. *Proceedings of the Royal Society B*, 274(1622), 2195-2203.
- Henzi, P. y otros (2007). Look Who's Talking: Developmental Trends in the Size of Conversational Cliques. *Evolution and Human Behavior*, 28(1), 66-74.
- Hill, David (1981). *An Atlas of Anglo-Saxon England*. Oxford: Blackwell.
- Hill, Russell A. & Robin Dunbar (2003). Social Network Size in Humans. *Human Nature*, 14, 53-72.
- Hurd, James P. (1985). Sex Differences in Mate Choice among the «Nebraska» Amish of Central Pennsylvania. *Ethology and Sociobiology*, 6, 49-57.

- Jerison, Harry (1973). *Evolution of the Brain and Intelligence*. Londres: Academic Press.
- Kerr, Natalie, Robin Dunbar & Richard P. Bentall (2003). Theory of Mind Deficits in Bipolar Affective Disorder. *Journal of Affective Disorders*, 73(3), 253-259.
- Keverne, Eric Barrington, Nicholas D. Martensz & Bernadette Tuite (1989). Beta-Endorphin Concentrations in Cerebrospinal Fluid of Monkeys Are Influenced by Grooming Relationships. *Psychoneuroendocrinology*, 14(1-2), 155-161.
- Killworth, Peter D., H. Russell Bernard & Christopher McCarty (1984). Measuring Patterns of Acquaintanceship. *Current Anthropology*, 25, 385-397.
- Kinderman, Peter, Robin Dunbar & Richard Bentall (1998). Theory-of-Mind Deficits and Causal Attributions. *British Journal of Psychology*, 89, 191-204.
- Laslett, Peter (1971). *The World We Have Lost*. Londres: Methuen.
- Leslie, Alan (1987). Pretense and Representation - The Origins of «Theory of Mind». *Psychological Review*, 94, 412-426.
- MacDonald, Charles B. (1955). 'Company'. *Encyclopedia Britannica* (pp. 143-144). Décimo cuarta edición. Londres: Encyclopedia Britannica.
- MacLeod, Carol E. y otros (2003). Expansion of the Neocerebellum in Hominoidea. *Journal of Human Evolution*, 44(4), 401-429.
- Mange, Arthur & Elaine Johansen Mange (1980). *Genetics: Human Aspects*. Nueva York: Holt Rinehart & Winston.
- Mitchell, Peter (1997). *Introduction to Theory of Mind*. Londres: Arnold.
- Montross, Lynn (1975). 'Tactics'. *Encyclopedia Britannica*. Décimo quinta edición. Londres: Encyclopedia Britannica.
- Naroll, Raoul (1956). A Preliminary Index of Social Development. *American Anthropologist*, 58, 687-715.
- Nettle, Daniel & Robin Dunbar (1997). Social Markers and the Evolution of Reciprocal Exchange. *Current Anthropology*, 38, 93-99.
- Oates, Joan (1977). Mesopotamian Social Organisation: Archaeological and Philological Evidence. En John Friedman y Michael J. Rowlands (eds.), *The Evolution of Social Systems* (pp. 457-485). Londres: Duckworth.
- O'Connell, Sanjida & Robin Dunbar (2003). A Test for Comprehension of False Belief in Chimpanzees. *Evolution and Cognition*, 9, 131-139.
- Orstrom, Elinor, Roy Gardner & James Walker (1994). *Rules, Games and Common-Pool Resources*. Ann Arbor: University of Michigan Press.
- Ozonoff, Sally (1995). Executive Functions in Autism. En Eric Schopler y Gary B. Mesibov (eds.), *Learning and Cognition in Autism* (pp. 199-218). Nueva York: Plenum Press.

- Perner, Josef (1991). *Understanding the Representational Mind*. Cambridge: The MIT Press.
- Pinker, Steven (1998). *How the Mind Works*. Londres: Allen Lane.
- Roth, Daniel & Alan M. Leslie (1998). Solving Belief Problems: Toward a Task Analysis. *Cognition*, 66(1), 1-31.
- Shultz, Susanne & Robin Dunbar (2007). The Evolution of the Social Brain: Anthropoid Primates Contrast with Other Vertebrates. *Proceedings of the Royal Society B*, 274(1624), 2429-2436.
- Shultz, Susanne, Ronald Nöe, William Scott McGraw & Robin Dunbar (2004). A Community-Level Evaluation of the Impact of Prey Behavioral and Ecological Characteristics on Predator Diet Composition. *Proceedings of the Royal Society B*, 271(1540), 725-732.
- Silk, Joan (2002). The «F»-Word in Primatology. *Behaviour*, 139, 421-446.
- Silk, Joan (2007). Social Components of Fitness in Primate Groups. *Science*, 317, 1347-1350.
- Silk, Joan, Susan C. Alberts & Jeanne Altmann (2003). Social Bonds of Female Baboons Enhance Infant Survival. *Science*, 302(5648), 1232-1234.
- Smuts, Barbara (1985). *Sex and Friendship in Baboons*. Nueva York: Aldine.
- Sternberg, Robert J. (1997). Construct Validation of a Triangular Love Scale. *European Journal of Social Psychology*, 27, 313-335.
- Stiller, Joan & Robin Dunbar (2007). Perspective-Taking and Social Network Size in Humans. *Social Networks*, 29, 93-104.
- Stylianou, Maria (2007). *Does Executive Function Training Improve Mentalising Ability?* Tesis de doctorado, Universidad de Liverpool.
- Swarbrick, Rebecca (2000). *A Social Cognitive Model of Paranoid Delusions*. Tesis de doctorado, Universidad de Manchester.
- Urban Church Project (1974). *Let My People Grow!* Londres: Informe presentado ante el Sínodo General de la Iglesia de Inglaterra.
- Van Vugt, Mark, Charlie Hardy, Julie Stow & Robin Dunbar (s/f). Laughter as Social Lubricant: a Biosocial Hypothesis about the Pro-Social Functions of Laughter and Humor. <http://psycnet.apa.org/journals/gdn/12/1/17/>
- Whiten, Andrew & Richard W. Byrne (eds.) (1997). *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge: Cambridge University Press.
- Zhou, Wei-Xing, Didier Sornette, Russell A. Hill & Robin Dunbar (2005). Discrete Hierarchical Organization of Social Group Sizes. *Proceedings of the Royal Society B*, 272(1561), 439-444.

LA NATURALIZACIÓN DE LOS CONCEPTOS MORALES. SOBRE LA TESIS DE LA IMAGINACIÓN MORAL DE MARK JOHNSON

Claudia Muñoz Tobar

Universidad de Concepción

En este trabajo presento una síntesis de la epistemología de la moral de Mark Johnson, filósofo de la Universidad de Oregon. Su propuesta, que es un desarrollo de la llamada tesis de la corporeización del significado, integra, entre otros, aportes de la filosofía del lenguaje, la filosofía de la mente, la fenomenología, la lingüística cognitiva y la psicología cognitiva. Me centraré principalmente en dos libros suyos, *El cuerpo en la mente* (1991) y *Moral Imagination* (1993). En el primero de estos libros Johnson presenta su proyecto de una teoría de la imaginación que, nos dice, intenta completar la consideración kantiana de las funciones imaginativas mediante la ampliación de estas al ámbito de la razón práctica. Este libro es más conocido, sin embargo, porque en él Johnson introduce las nociones de esquema de imagen y de proyección figurativa, dos estructuras de la imaginación de las que depende la naturaleza corporeizada de los conceptos y, en este sentido, de la mente. En el segundo libro, *Moral Imagination*, Johnson desarrolla más exhaustivamente su proyecto de una teoría de la imaginación en la forma de una epistemología de la moral. Aspectos de este desarrollo se encuentran también en un libro del año 2007, *The Meaning of the Body*, que está dedicado en gran parte al conocimiento estético.

La teoría de la imaginación moral se podría caracterizar como una explicación de corte cognitivista del sistema de conceptos que da forma a la teoría moral tradicional. Por teoría moral tradicional Johnson se refiere al modelo mediante el cual entendemos la moral como un sistema de reglas, racionalmente derivadas, que nos dicen qué debemos hacer en una situación concreta y la acción moral como el seguimiento de esas reglas (Johnson, 1993, pp. 1-12). En su análisis, Johnson se refiere a este modelo en términos de la Teoría de la Ley Moral (TLM), un modelo cognitivo complejo que subyace tanto a la conceptualización y deliberación moral cotidianas como a toda la tradición occidental de la filosofía moral.

Johnson aclara que su libro tiene dos lecturas, una moderada y otra fuerte. De acuerdo con la lectura moderada, el libro propone la necesidad de complementar la teoría moral tradicional con una teoría de la racionalidad imaginativa, ampliando así el ámbito de alcance de la imaginación kantiana. La finalidad de esta teoría sería mostrar que los principios y leyes que se imponen sobre los casos morales particulares, y que se suponen derivados racionalmente, se basan más bien en conceptos metafóricos y que su aplicación requiere varias formas de la imaginación (1993, p. 77). De acuerdo con la otra lectura, la más fuerte, el libro propone que el examen de la naturaleza imaginativa de nuestro sistema conceptual y de la razón conduce a una revisión crítica de la teoría moral tradicional y a delinear una concepción diferente de la teoría moral que no dependería, como la tradicional, del descubrimiento y de la aplicación de las leyes morales (p. 77). Este programa de trabajo, en sus versiones moderada y fuerte, no sería posible sin los aportes de la ciencia cognitiva de segunda generación, que ha aportado evidencia acerca del carácter imaginativo de las capacidades cognitivas por las cuales creamos y comprendemos los conceptos morales básicos y razonamos acerca de ellos (p. 1). Estas capacidades cognitivas no son exclusivas del razonamiento moral, sino que pertenecen al razonamiento y a la razón en general. La tesis es que la razón es esencialmente imaginativa (p. 6; 1991, p. 2), corporeizada y figurativa (1991, p. 223), mientras que toda la teoría moral tradicional se fundaría en un concepto de razón no imaginativa, es decir, desencarnada y literal. Johnson se compromete así con una forma de naturalismo que afirma la naturaleza corpórea de los conceptos y que ofrece herramientas para analizar la estructura cognitiva del pensamiento occidental (Lakoff & Johnson, 1999).

Dividiré mi presentación en tres secciones. Primero, haré una breve reseña de la teoría de la corporeización de la mente en la versión de Johnson y de cómo, según él, esta teoría representa una superación del dualismo ontológico. En esta primera parte me referiré básicamente a dos nociones, ya mencionadas al comienzo, la de esquemas encarnados o esquemas de las imágenes y la de proyección figurativa metafórica, ambas, como señalé, son fundamentales en su teoría de la imaginación (1991, p. 222). En segundo lugar, mostraré cómo aplica Johnson estas nociones en su epistemología de la moral, la que puede caracterizarse, según ya anuncié, como una explicación cognitiva de la deliberación moral y de la teoría moral. Finalmente, en la tercera parte de mi exposición, muestro por qué la forma de naturalismo moral con la que Johnson está comprometido no ofrece respuestas a preguntas sobre justificación moral. En este caso, propongo una forma de responder a la objeción de Rosas (2000, p. 105), quien ha visto en esto una deficiencia del naturalismo de Johnson, minimizando su aporte al proyecto naturalista en la filosofía moral. Aunque coincido con Rosas en que el enfoque conceptual no ofrece este tipo de respuestas, mi punto aquí

será que esto no afecta al éxito de la propuesta naturalista de Johnson, no solo porque los objetivos de la teoría de la imaginación moral se enmarcan en un contexto distinto al de la reflexión sobre el origen de la moral, sino también porque el problema de la justificación se genera, precisamente, en el modelo tradicional que Johnson critica. Una vez que su análisis pone de manifiesto la estructuración imaginativa de la moral, es decir, los conceptos metafóricos que subyacen al modelo tradicional occidental y definen los supuestos en que se basa, Johnson propone una nueva forma de comprensión de la moral, es decir, un nuevo modelo conceptual que exige nuevas metas a la filosofía moral.

1. TESIS DE LA CORPOREIZACIÓN

Johnson caracteriza su propuesta para una teoría de la imaginación como un intento de basar la explicación del significado y del razonamiento humanos en una elaboración de las operaciones de la imaginación construidas en el sentido amplio que Kant les dio (1991, p. 242). Lo interesante de la consideración kantiana, para este autor, es que trata la imaginación en su aspecto creativo en el arte, la literatura y la ciencia e intenta relacionarla con las otras funciones de la imaginación que ha descrito en la *Crítica de la razón pura*: las funciones reproductiva, productiva y esquemática (1991, p. 243), asignándole así un relevante papel en el conocimiento objetivo. La perspectiva kantiana es importante porque, de acuerdo con el análisis de Johnson, habría dos formas en que la filosofía habría considerado la imaginación en su relación con el conocimiento (pp. 224-228), una que le niega rotundamente este papel (la tradición surgida de la interpretación de los escritos platónicos y el romanticismo), y otra, la línea aristotélica, que sí se lo concede. En esta última línea se ubicaría Kant, pero va más allá. Junto con superar el prejuicio de que la imaginación no tiene ningún papel en el conocimiento objetivo, Kant amplió las funciones de la imaginación para dar una forma de racionalidad al juicio estético. Sin embargo, no habría ocurrido algo similar en su teoría moral, en la que la imaginación aparece como un peligro para la deliberación y la toma de decisiones racionales (pp. 231-261). La teoría de la imaginación de Johnson pretende unificar conceptualmente la razón, o quizás sea mejor decir que su propósito es dar continuidad imaginativa a las tres formas de la racionalidad ilustrada: teórica, estética y práctica.

1.1. Estructuras imaginativas y dualismo ontológico

La hipótesis de la corporeización de la mente tiene varias versiones, pero en general todas tienen en común la idea básica de que mente y cuerpo no son dos entidades separadas sino partes de un *continuum* o bien dos aspectos de un mismo proceso

orgánico (Johnson, 2007, p. 1). Esta es la idea que Johnson desarrolla en *El cuerpo en la mente* (1991), obra en la que examina las estructuras imaginativas que dan cuenta del proceso por el cual el cuerpo se abre paso hacia la mente, a fin de dotar de significación la vida humana en todas sus dimensiones: significado, razonamiento y acción. La corporeización del significado y de la mente, en la versión de Johnson, consiste en la tesis de que todo proceso racional (pensamiento, comprensión y acción) es moldeado por los procesos corporales (entiéndase programas sensoriales, perceptivos y motores), más inmediatos y cognitivamente determinantes. La razón es corpórea en el sentido de que surge del cuerpo (Lakoff & Johnson, 1999) a través de un proceso estéticamente dependiente (Johnson, 2007), es decir, conducido y determinado por nuestras capacidades imaginativas.

En el contexto de la perspectiva cognitivo conceptual, llamada también experiencia lista, con la que se asocia a Lakoff y Johnson, se han identificado varias estructuras imaginativas, que son las que intervienen en los procesos de conceptualización. Se trata de los esquemas de las imágenes (Johnson, 1991) y las proyecciones figurativas (o elaboraciones figurativas de las estructuras esquemáticas), que incluyen la metáfora (Lakoff & Johnson, 2004) y la metonimia (2004 y 1999). El tratamiento de Johnson en *El cuerpo en la mente* apunta principalmente a dos de estas estructuras: los esquemas de imágenes y las proyecciones metafóricas. Aunque estas estructuras de la imaginación intervienen en la cognición moral, en esta primera parte de la presentación me interesa destacar dos aspectos del examen experiencialista, primero, la forma en que se relacionan estas dos estructuras, esquema de imagen y proyección figurativa, para dar origen a los conceptos. Es sobre la base de esta forma de relación que la tesis de la corporeización puede oponerse al dualismo metafísico; el segundo aspecto del examen conceptual que trato aquí es su intención de mostrar que la imaginación no sería una facultad sino, más bien, una propiedad masiva de la racionalidad humana, el carácter esencial de la comprensión y del significado.

1.2. Esquema de imagen

Tal como lo entiende Johnson, el esquema de imagen es una estructura imaginativa similar al esquema kantiano. Kant consideró el esquema de la imaginación como un paso intermedio entre la percepción y el concepto puro, una estructura imaginativa que no sería ni empírica ni conceptual, pero sí una condición para la formación de toda imagen empírica (por ejemplo, la de perro) y para la aplicación de los conceptos sobre ellas. El esquema de Johnson, lo mismo que el de Kant, es una estructura fundamental para la formación de cualquier imagen rica (1991, p. 77), pero a diferencia del esquema kantiano, no sería un mero paso intermedio entre la percepción y el concepto, pues el esquema de imagen, aunque preconceptual (pp. 63-64), representa

la base experiencial o corpórea de la formación y ampliación de los conceptos vía elaboraciones figurativas. Por una parte, el esquema de imagen actúa directamente en la experiencia, por ejemplo, cuando vamos caminando de nuestra casa al trabajo se haya activo el esquema origen-recorrido-meta, pero, por otra parte, este mismo esquema participa, indirectamente, en la comprensión de otros dominios de experiencia, que no son necesariamente motores, como, por ejemplo, en la comprensión del desarrollo evolutivo que nos conduce de niños a adultos. En el primer caso, un mismo esquema está involucrado en la realización de conductas motoras habituales o repetitivas, actividades o experiencias desde las que el esquema se origina también, mientras que en el segundo, el mismo esquema nos permite dar sentido a experiencias nuevas y más complejas. En ninguno de estos dos casos de su aplicación tenemos acceso consciente al esquema de la imagen ni a su estructura.

Los esquemas de imágenes no son solo perceptuales, porque también hay esquemas que surgen del funcionamiento específico de nuestros programas motores. Los esquemas emergen de la relación de nuestros sistemas perceptual y motor en nuestro trato y relación con el entorno. Son, por decirlo así, un producto en cuya estructura gestáltica confluyen tanto nuestra propia condición biológica como las propiedades externas, las del mundo en el que nos desenvolvemos. Este origen hace de los esquemas de las imágenes estructuras corporeizadas o experienciales.

Entonces, los esquemas llegan a formarse debido a que realizamos actividades de percepción y motoras de manera repetitiva, y los utilizamos para dar significado al tipo de experiencias de las que son esquemas o resúmenes abstractos, y también para tratar con sucesos y experiencias nuevos o novedosos y darles sentido. Es especialmente para realizar esto último, es decir, para cumplir la función creativa de aplicar los esquemas de imágenes a nuevas experiencias, que se requiere la operación de las proyecciones figurativas. El hecho de que comprendamos nuestro desarrollo de niños a adultos como un recorrido que va desde un punto inicial u origen (la niñez) hasta un punto de término o meta (la adultez) se debe a una elaboración figurativa del esquema origen-recorrido-meta que nos permite aplicarlo en la comprensión del dominio del desarrollo del individuo. A este tipo de proyección se le llama metáfora, una de las formas, la más importante y más explotada, en que un esquema de imagen, es decir, el cuerpo, se abre paso hacia la mente dando lugar a un complejo sistema que Lakoff y Johnson han denominado sistema conceptual ordinario (2004, p. 39).

1.3. Proyecciones figurativas

Una proyección figurativa puede definirse básicamente como una operación cognitiva por la cual un esquema de imagen se abre paso hacia la mente en la forma de una conceptualización. De esta manera, los esquemas intervienen en la comprensión

de diversos dominios de la experiencia y en el modo en que razonamos acerca de ellos. Las proyecciones figurativas ocurren también entre y dentro de dominios categorizados, de distintos niveles de complejidad; un ejemplo de la primera forma de proyección es la comprensión de la historia humana en términos del desarrollo ontogénico, de modo tal que no nos parece nada novedoso hablar de una infancia y de una adultez de la humanidad como puntos extremos de un recorrido lineal. Un ejemplo de la segunda clase proyección figurativa consiste en comprender un concepto a partir de las propiedades de alguno de sus aspectos más destacados o prototípicos, tal es el caso de la comprensión de una categoría cotidiana como la de madre, que se ha vuelto bastante más compleja en nuestra época, a partir del submodelo más simple de la madre biológica (Lakoff, 1987, pp. 74-76).

Como señalé, la clase de proyección figurativa que ha recibido más atención tanto en lingüística cognitiva como en psicología y en psicolingüística, es la metáfora. Si recordamos el ejemplo anterior, una proyección metafórica es una operación por la cual un dominio, normalmente más concreto y delineado corporalmente, o un esquema (por ejemplo, el esquema origen-recorrido-meta), llamado dominio fuente, es usado para mapear y así estructurar parcialmente un dominio distinto, más abstracto (por ejemplo, el desarrollo del individuo), llamado dominio meta. Las proyecciones figurativas en esta versión de la hipótesis de la corporeización tienen este carácter unidireccional, ocurren desde lo más concreto a lo más abstracto (Rohrer, 2007, p. 28). Actualmente, en el marco de la semántica cognitiva ha habido un creciente desarrollo del estudio de la metonimia, otra forma de proyección figurativa que, a diferencia de la metáfora, actúa dentro de un mismo dominio de experiencias (Lakoff, 1987, p. 288 y Barcelona, 2002, p. 215). Lamentablemente, Lakoff y Johnson usaron los rótulos de metáfora y metonimia para referirse a estas dos clases de proyección figurativa, lo que suele producir confusiones cuando se intenta analizar su papel en la cognición. Las metáforas y metonimias, como proyecciones figurativas de la imaginación, no están voluntariamente a nuestra disposición, como sí lo están muchas de las expresiones metafóricas y metonímicas que leemos y escuchamos habitualmente. Las proyecciones figurativas son, más bien, estructuras inconscientes que subyacen a estas últimas. Tanto el lenguaje metafórico como el metonímico son formas de expresión superficial de tales estructuras, indicios de su existencia y medios, quizás los más importantes (al menos así lo entienden los lingüistas cognitivos), para llegar a ellas. En consecuencia, se debe distinguir entre expresiones metafóricas, como «lo atrae su belleza» o «los sentimientos le impiden tomar buenas decisiones», y las metáforas conceptuales que las motivan: La belleza es fuerza y Los sentimientos son fuerzas. Estas dos metáforas tienen como dominio fuente el mismo esquema imaginativo, que surge de la experiencia de que las cosas ejercen fuerza sobre nosotros

y de que también nosotros ejercemos fuerza sobre ellas. La fuerza no es un tipo de experiencia simple para nosotros, por lo que Leonard Talmi (1988, pp. 52-56) propone la existencia de un sistema esquemático de dinámica de fuerzas (*force dynamics*) en lugar de la estructuración más simple de un esquema de fuerza, con lo que quiere señalar, en general, que la experiencia de la fuerza física tiene una lógica interaccional que queda plasmada sistemáticamente en el esquema dinámico, entre dos fuerzas que se enfrentan: la de una entidad con una tendencia intrínseca al descanso o al movimiento, o fuerza agonista (*agonist*), y la de otra entidad que se opone a ella, o fuerza antagonista (*antagonist*).

1.4. Categorías de prototipo y Modelo Cognitivo Idealizado (MCI)

Para mostrar qué condiciones conceptuales definen la aplicación de las normas a las situaciones concretas en el marco del modelo tradicional de la moral, Johnson se vale de los desarrollos en teoría de la categorización. La psicología cognitiva (Rosch, 1973a, 1973b y 1975) ha aportado evidencia de la estructura de prototipos de las categorías, en oposición al supuesto clásico de que estas se definen por condiciones necesarias y suficientes. Cuando las personas realizan actividades de categorización normalmente se verifican efectos de prototipo (Lakoff, 1987, pp. 41-42), el más simple de los cuales es el de gradiencia. Así, algunas entidades nos parecen más representativas de la categoría, mientras que otras nos lo parecen menos; de la misma manera hay propiedades de la categoría que suelen ser más relevantes que otras a la hora de definir las. Lo interesante es que este fenómeno no se da únicamente cuando tratamos con categorías de estructura prototípica (es decir, categorías con un centro cognitivo, el prototipo, y miembros más cercanos o lejanos a él), sino también cuando lo hacemos con categorías clásicas (Lakoff, 1987, pp. 68-76). Lakoff introduce el concepto de Modelo Cognitivo Idealizado (MCI) para explicar este fenómeno (1987, p. 68). Su tesis es que el conocimiento categorial se haya determinado, en su mayoría, por este tipo de estructuras conceptuales. Los efectos de prototipo se deberían, en gran parte de los casos, a la inconsistencia entre la gran simplicidad estructural que caracteriza nuestras idealizaciones (los MCI) y la gran complejidad de las situaciones concretas que modelizan. Los efectos de prototipicidad se producen por el desajuste entre un MCI dado, construido a partir de casos representativos, más simples y claros, y los ejemplos novedosos con los que nos topamos. Estos últimos son ubicados en los límites de la categoría, más alejados del centro prototípico. De la misma manera, el desajuste entre un MCI y la realidad puede conducir o a una mayor complejidad del modelo mismo, como sucede, por ejemplo, con el MCI a partir del cual definimos la categoría madre (Lakoff, 1987) o al surgimiento de nuevas categorías, como, las de mentira piadosa o mentira blanca, que definen

contextos que no se ajustan al modelo clásico de la mentira (Lakoff, 1987; Johnson, 1993 y Sweetser, 1989). Este análisis le permite a Johnson afirmar la primacía natural de la categorización prototípica por sobre la categorización clásica y explicar de qué manera la TLM, en tanto es ella misma un MCI, nos guía en nuestra comprensión y nuestro trato con las diversas situaciones de la vida moral.

2. MORAL CORPOREIZADA

La moral corporeizada es una forma de decir que los conceptos morales obedecen a los mismos principios de estructuración corpóreo-imaginativa que participan en otras dimensiones conceptuales de la vida humana. Johnson analiza la estructura imaginativa del Modelo Moral Occidental, que subyace, igualmente, a las elaboraciones de la filosofía moral. La TLM es el MCI a partir del cual se definen metafóricamente muchos de los conceptos morales básicos heredados de nuestra tradición (como los de libertad, voluntad, razón y deber, entre otros). La comprensión acerca de cómo se relacionan entre sí estos conceptos, en la mente moral, está estructurada sobre la base de un único concepto central: la metáfora de la *razón como fuerza*. Solo haré aquí una síntesis de esta explicación, deteniéndome en la base corpóreo-esquemática de algunos de estos conceptos definidos en el modelo, principalmente los de acción, voluntad, pasión y razón.

2.1. Teoría de la Ley Moral

La TLM puede definirse como un MCI complejo, en la medida en que incluye a lo menos cuatro submodelos interrelacionados: de la naturaleza humana, del funcionamiento de la mente, de lo que cuenta como un asunto moral y del origen de las leyes morales (Johnson, 1993, p. 13). De acuerdo con la TLM, la conducta moral consiste en el seguimiento de reglas que nos dicen qué debemos hacer en una situación concreta. Estas reglas morales no son dictadas por una voluntad individual y relativa, sino por Dios, o en su formulación racionalista, por la Razón Universal. Es esencial para el funcionamiento de la TLM la adecuación entre las reglas para la acción y las situaciones concretas que regulan, de hecho, el modelo se estructura a partir del tratamiento con situaciones muy simples y claras. El «respeto por la vida humana», por ejemplo, es un principio que, de acuerdo con la TLM, tiene una aplicación unívoca porque la situación concreta que regula es una que supone un concepto también unívoco de la vida humana, el que incluye la concepción del ser humano como persona. El modelo supone entonces la existencia de una definición clara y fija de lo que es ser una persona. Lo que sea la vida humana se haya definido, por tanto, en el marco del mismo modelo que se intenta aplicar. Si no estuviera claro qué se entiende por

persona no se podría o sería muy difícil aplicar a una situación concreta la regla que dice que debemos respetar la vida humana. Entonces, hay un supuesto que subyace a este modelo, el supuesto de que los conceptos tienen una definición absoluta basada en condiciones necesarias y suficientes. La TLM incluye también una noción de los agentes que siguen estas reglas; se trata de agentes morales porque, a diferencia de los animales no humanos, poseen una razón que les permite conocer y comprender las reglas mismas, es decir, captar su racionalidad inherente; y este hecho de ser las únicas criaturas capaces de comprenderlas hace que sean también las únicas criaturas que están sujetas a esas reglas, las únicas que pueden y deben guiar sus vidas de acuerdo con ellas (Johnson, 1993, p. 20). La TLM contiene además un modelo, estructurado metafóricamente, sobre el funcionamiento de nuestra facultad de deliberación y acción morales que Johnson denomina Modelo de la Facultad de la Psicología Moral (1993, p. 15). Este modelo motiva una comprensión de cómo interaccionan los diferentes componentes de nuestra mente moral, la que solo puede entenderse sobre la base de una concepción dual de la naturaleza humana, en parte animal, y en este sentido corpórea y pasional, y en parte mental, y por lo mismo espiritual y racional. Los componentes de la psicología moral son fundamentalmente: *razón, pasión, voluntad y libertad*. Cada uno de estos componentes es entendido, metafóricamente, como una entidad y las interrelaciones entre ellas se comprenden a partir de elaboraciones metafóricas de esquemas de imágenes, fundamentalmente el esquema de Dinámica de fuerzas (pp. 15-16). Así, la metáfora dominante en el modelo tradicional occidental de la moral es la metáfora de la *razón como fuerza* (p. 25). Una acción es moralmente correcta si la fuerza de la razón ha sido ejercida con éxito sobre nuestra voluntad. Esto quiere decir que la voluntad ha logrado vencer la fuerza opuesta que sobre ella imponía la *pasión*. En este MCI del funcionamiento de los componentes de nuestra moralidad, la *voluntad* se ve sometida a una dinámica con dos fuerzas opuestas que luchan por dominarla, ante lo cual tiene solo dos posibilidades, ceder ante la fuerza de la *razón* o ceder ante la fuerza de la *pasión*. En el marco del modelo las acciones se conceptualizan figurativamente, a partir del esquema origen-recorrido-meta, como movimientos a través de un camino y la libertad como la falta de obstáculos o restricciones a esos movimientos. Los dilemas morales son conceptualizados como bifurcaciones, es decir, como caminos alternativos que los agentes deben elegir.

Quiero detenerme ahora en el siguiente aspecto de la TLM. Esta define cuándo una situación es moralmente problemática, determinando así nuestra conciencia o reconocimiento de que estamos en esta clase de circunstancia. Conceptualizamos las situaciones moralmente problemáticas como casos en que nuestra voluntad se ve enfrentada a dos fuerzas opuestas, la *razón* y la *pasión*. Una solución moralmente adecuada a este conflicto consiste en dejar que la *razón* (la Ley Moral) ejerza su fuerza

sobre nuestra voluntad, bloqueando la fuerza opuesta de la *pasión*. Lo que sea una situación moralmente problemática se define, entonces, sobre la base de un caso prototípico, uno en el que la persona se ve enfrentada a la decisión de seguir lo que la razón le dice (por ejemplo, obedecer a sus padres) o de seguir lo que le dice su deseo o pasión (por ejemplo, no obedecer). Tal como señalé más atrás, los MCI están estructurados en función de estos casos representativos, que son, por lo mismo, los que se ajustan perfectamente al modelo idealizado. Sobre la base de esta reflexión, Johnson considera que un verdadero problema moral no corresponde precisamente al tipo de los casos prototípicos, sino que se manifiesta cuando nos toca tratar con aquellos que no se ajustan perfectamente, con situaciones que nos enfrentan a los límites de nuestro modelo moral, lo que exige de nosotros una reacción creativa de la imaginación. Esta acción creativa de la imaginación consiste en extender figurativamente nuestro conocimiento de los casos prototípicos a nuevas situaciones, ideando formas alternativas de solución y esto es algo que ocurre jugando de alguna manera con los límites categoriales, extendiendo nuestras categorías o bien redefiniéndolas, en un proceso que exige de nosotros una capacidad que Johnson llama ‘imaginación empática’ (*empathetic imagination*, 1993, p. 199).

Hasta aquí, hemos señalado que la TLM está anclada, según Johnson, en una consideración del significado de los conceptos, la teoría clásica de la categorización, que no se ajusta a las características de nuestra cognición categorial, básicamente prototípica. Además, el modelo de la TLM supone la naturaleza dual del ser humano y, con ello, una forma de racionalidad desencarnada que no se condice con su naturaleza esencialmente imaginativa. Y, por último, la TLM define los conceptos morales básicos a partir de proyecciones metafóricas que fortalecen esa concepción dual, y de las leyes morales como absolutas. Estas deficiencias nos enfrentan a conflictos ilusorios que el modelo no permite resolver adecuadamente y que dificultan la resolución creativa de los problemas morales.

2.2. Límites de la moral corporeizada

El sistema de principios de estructuración conceptual determina no solo el significado de los conceptos mediante los cuales comprendemos el mundo, sino también el modo en que razonamos con ellos para tomar decisiones y actuar en las diversas dimensiones de la vida. El enfoque experiencalista de la cognición moral se enfrenta a dos dificultades. La primera es que el conocimiento del origen y estructura metafórica de los conceptos morales no satisface plenamente las aspiraciones filosóficas en el ámbito práctico. Mientras la filosofía moral se ocupa de la búsqueda de un sustento universal para la justificación de las decisiones morales, ubicándose,

según sus respuestas a este problema, en los polos del relativismo o del absolutismo, el enfoque conceptual tiene el objetivo distinto de investigar la proveniencia y fuentes de estructuración de los conceptos morales por los que las personas toman esas decisiones.

Un segundo problema que debe enfrentar el análisis conceptual es la lectura relativista de su afirmación de que los esquemas de las imágenes pueden ser elaborados metafóricamente de maneras diversas. Esta dificultad no afecta solo al estatus de la teoría de la imaginación moral dentro de la filosofía moral, sino al enfoque mismo, ya que Johnson lo considera una superación de la dicotomía absolutismo-relativismo (1993, pp. 3-4, 80 y ss.). En lo que resta, sin embargo, no me referiré a este problema, sino que me dedicaré solo a la primera dificultad. Es posible enfrentarse a esta señalando que, para Johnson, un modelo sí puede ser mejor que otro en la medida en que armonice mejor con la naturaleza de nuestras capacidades cognitivas o imaginativas, puesto que un modelo de estas características no nos enfrentaría ya con conflictos ilusorios, que tienen su fuente en la base dualista de la TLM. Desde el punto de vista de la filosofía moral, modificar el modelo consiste en modificar las metáforas centrales en que se sostiene, y que definen los conceptos morales básicos. Dada la tesis de que la razón es esencialmente imaginativa, el nuevo modelo debería basarse en una metáfora a la luz de la cual la deliberación moral no se comprenda bajo el supuesto de la separación entre la razón y las motivaciones y condiciones asociadas al cuerpo, y a la dinámica de estas dos fuerzas opuestas. El concepto desencarnado de la razón se cambiaría, entonces, por uno que suponga una continuidad entre las experiencias básicas de nuestra cognición, esencialmente corporales, y las experiencias novedosas, complejas y de niveles mayores de abstracción.

3. IMAGINACIÓN O CREATIVIDAD MORAL

La concepción de un sistema fuertemente dominado por la dicotomía metafísica entre lo físico y lo mental, habría impedido a Kant reconocer el papel más amplio y determinante de la imaginación, es decir, no solo su función en el conocimiento objetivo y estético, sino también en el moral. El aporte de la epistemología moral corporeizada consiste en extender la imaginación a la racionalidad como tal, en todas sus aplicaciones (conocimiento, arte y acción moral), lo que llevaría a modificar el modelo de la TLM.

La imaginación moral, tal como la imaginación en el arte, hace de la deliberación una actividad que nos permite conceptualizar y razonar creativamente cuando tratamos con experiencias morales novedosas, o no prototípicas, proyectando el acervo experiencial resumido en los que solemos llamar nuestros principios morales.

Una experiencia moral no prototípica tiene la misma relación que guardan los casos no representativos de una categoría con el MCI que la define. Digamos que se escapan del MCI, estructurado de acuerdo con casos representativos, más simples y accesibles. Como vimos, los dilemas o conflictos morales son tipos de situaciones no representativas. El dilema moral no puede explicarse como un caso en que nuestra voluntad se haya simplemente en medio de la lucha de dos fuerzas opuestas. Sentimos que estamos en un dilema moral cuando la situación que debemos enfrentar no puede ser resuelta recurriendo a la lógica del modelo tradicional que nos dice cuál de estas dos fuerzas debe vencer para que nuestra decisión sea moralmente correcta. Puede ser el caso de que no esté claro cómo aplicar esa lógica porque no es posible identificar dos fuerzas antagónicas, sino que nos sentimos, más bien, en una situación en que uno de los conceptos definidos dentro del marco del modelo imperante no se ajusta a la realidad más compleja que nos toca enfrentar; sucede esto, por ejemplo, cuando tenemos que pronunciarnos y tomar decisiones acerca de asuntos como el aborto, o nuestro trato con los animales no humanos o el medioambiente, porque ello nos exige tomar posturas que nos enfrentan a los límites de nuestros conceptos, definidos por un modelo de categorías fijas o clásicas, y basado en una metáfora de entidades o fuerzas en conflicto.

3.1. Moral como arte

En escenarios moralmente conflictivos lo que sucede, gran parte de las veces, es que los enfrentamos creativamente. Tomamos (elaboramos) los materiales brindados por la experiencia (es decir, los casos prototípicos que el modelo describe) y afrontamos con ellos la situación nueva, tomando una decisión. Para este caso, Johnson propone que lo que hemos aprendido de la ciencia cognitiva acerca de la naturaleza imaginativa de la razón nos ofrece la posibilidad, tanto en nuestra vida moral como en la filosofía moral, de examinar y modificar todos o parte de nuestros modelos heredados. Este proceso consistiría en llevar adelante elaboraciones imaginativas que transformen el modelo a partir del cual comprendemos y definimos los conceptos morales básicos, lo que supone el reemplazo de la metáfora básica sobre la que se asienta. Johnson propone una proyección del proceso creativo en el arte a la actividad deliberativa. A partir de la nueva metáfora de la moral como arte (1993, pp. 210-212), deliberar moralmente ya no puede comprenderse en términos de la derivación racional de leyes ni de la aplicación de normas a casos particulares. Esta nueva metáfora no nos exige justificar las decisiones morales, con base en una reflexión sobre la racionalidad de las normas que las determinan. La racionalidad o «corrección» de nuestras decisiones depende, de acuerdo con el modelo tradicional,

de que hayan sido tomadas dejándonos dominar solo por la fuerza de una norma que debe su racionalidad a su independencia con respecto a cualquier determinación corporal o material, y a cualquier inclinación subjetiva. Entender la moral como arte, en cambio, nos lleva a concebir las decisiones para la acción como el resultado de un razonamiento que extiende nuestro saber acerca de experiencias anteriores (materiales con que contamos) a nuevas experiencias (nuevos materiales), dando origen a formas novedosas (de comportamientos y experiencias). Bajo esta nueva mirada, las leyes o principios que se definen en el marco de un modelo moral se entienden más bien como resúmenes de nuestra experiencia moral colectiva que abordan problemas importantes que deberían ser considerados en nuestras deliberaciones morales acerca de situaciones problemáticas nuevas, ensayando imaginativamente diversos escenarios, poniéndose en la perspectiva de los otros o bien asumiendo diversas perspectivas posibles para uno mismo (1993, p. 199). Esto exige pensar los principios morales de una manera distinta a la tradicional, es decir, no como un conjunto de reglas restrictivas que fijan lo que debemos hacer una situación dada (p. 80). Por esta razón, la filosofía moral se enfrenta a nuevas tareas. Nuestras capacidades imaginativas no hacen rectas o erróneas nuestras decisiones morales, sino que son el mecanismo natural que colabora en la aplicación de nuestro modelo moral a nuevas situaciones, y su cultivo puede contribuir, nos dice Johnson, a que la deliberación moral nos permita vivir mejores vidas, individual y colectivamente. El efecto liberador de esta epistemología moral no es comparable con las respuestas acerca de la fundamentación de nuestras normas, a las que aspira la filosofía moral, bajo la influencia del modelo de la TLM. La teoría de la imaginación moral nos dice lo que somos a través del conocimiento de la naturaleza corporal e imaginativa de los conceptos mediante los cuales conducimos nuestras vidas, y de la forma de razonar con ellos, ofreciendo una comprensión psicológica y filosóficamente más realista de la moral, que puede, por ejemplo, explicar y así liberar a las personas de las tensiones que suelen encontrar entre su punto de vista moral tradicional y la manera en que efectivamente experimentan los dilemas morales (p. 1). Y dado que la imaginación moral es una capacidad esencial de la cognición humana que es posible cultivar, la propuesta de Johnson puede tener un lugar en la reflexión sobre la «vida buena». La teoría de la imaginación moral supone una forma de racionalidad que no puede definirse como la búsqueda de razones universales que fundamenten las decisiones particulares. La ampliación imaginativa de la racionalidad, que en cierto sentido conecta la filosofía kantiana con aspectos de la vida moral concreta, conduce, más bien, a una reflexión acerca de lo que la propia tradición moral puede aportar, como fuente de las elaboraciones imaginativas, en el proceso de florecimiento humano.

REFERENCIAS BIBLIOGRÁFICAS

- Barcelona, Antonio (2002). Clarifying and Applying the Notions of Metaphor and Metonymy within Cognitive Linguistics: An Update. En René Dirven y Ralf Pörings (eds.), *Metaphor and Metonymy in Comparison and Contrast* (pp. 207-277). Berlín-Nueva York: Mouton de Gruyter.
- Johnson, Mark (1991). *El cuerpo en la mente. Fundamentos corporales del significado, la imaginación y la razón*. Madrid: Debate.
- Johnson, Mark (1993). *Moral Imagination. Implications of Cognitive Science for Ethics*. Chicago: The University of Chicago Press.
- Johnson, Mark (2007). *The Meaning of the Body. An Esthetics of Human Understanding*. Chicago: The University of Chicago Press.
- Lakoff, George (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Lakoff, George & Mark Johnson (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Nueva York: Basic Books.
- Lakoff, George & Mark Johnson (2004). *Metáforas de la vida cotidiana*. Madrid: Cátedra.
- Rohrer, Tim (2007). Embodiment and Experientialism. En Dirk Geeraerts y Herbert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 25-47). Oxford University Press.
- Rosas, Alejandro (2000). Explicación y justificación: Hacia el naturalismo en la filosofía moral. En *Mentes reales: La ciencia cognitiva y la naturalización de la mente* (pp. 101-118). Colombia: Ediciones de la Universidad Nacional de Colombia.
- Rosch, Eleanor (1973a). Natural Categories. *Cognitive Psychology*, 4(3), 328-350.
- Rosch, Eleanor (1973b). On the Internal Structure of Perceptual and Semantic Categories. En Timothy Moore (ed.), *Cognitive Development and the Acquisition of Language* (pp. 111-144). Nueva York: Academic Press.
- Rosch, Eleanor (1975). Principles of Categorization. En Eleanor Rosch y Barbara Bloom Lloyd (eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale: L. Erlbaum.
- Sweetser, Eve (1989). The Definition of Lie: an Examination of the Folk Models Underlying a Semantic Prototype. En Dorothy Holland y Naomi Quinn (eds.), *Cultural Models in Language and Thought* (pp. 43-66). Cambridge: Cambridge University Press.
- Talmy, Leonard (1988). Force Dynamics in Language and Cognition. *Cognitive Science*, 12(1), 49-100.

LA EVOLUCIÓN DE LA ATRIBUCIÓN PSICOLÓGICA: LECTURA DE MENTES Y METACOGNICIÓN

Pablo Quintanilla

Pontificia Universidad Católica del Perú

Esta contribución desea abordar la pregunta sobre la conexión lógica, causal y genealógica entre el conocimiento de los estados mentales ajenos, el de los estados mentales propios y el de la realidad objetiva compartida. En este texto sostendré que no puede darse ninguna de estas formas de conocimiento sin las otras dos. Me propongo defender que tanto en la dimensión conceptual como en la filogenética y en la ontogenética, estas tres formas de conocimiento constituyen un inseparable triángulo. En este artículo, sin embargo, me concentraré solo en las dos formas de conocimiento que constituyen la atribución psicológica, de un lado, lectura de mentes o conocimiento de la subjetividad ajena y, del otro, metacognición, autoconocimiento y autoconciencia.

Comenzaré rastreando la idea de que el conocimiento tiene forma triangular, en Peirce y Davidson, para luego detenerme en la atribución psicológica. Argumentaré en contra de las propuestas que sostienen que hay anterioridad lógica y causal de uno de estos tipos de conocimiento, como en el caso de Goldman (2006), que privilegia la autoconsciencia, o de Carruthers (2009) que sostiene la anterioridad lógica y causal, pero no temporal, de la lectura de mentes. También sostendré que la evidencia empírica prueba que estos tipos de conocimiento se dan juntos en la evolución de la especie y en el desarrollo del individuo, potenciándose mutuamente. De esta manera la presión del entorno habría seleccionado, al mismo tiempo y por las mismas causas, a los individuos con más habilidades de lectura de mentes y metacognitivas.

1. LA TRIANGULACIÓN

En «Algunas consecuencias de cuatro incapacidades»¹, publicado originalmente en 1868, Charles S. Peirce objeta lo que él denomina «el espíritu del cartesianismo», a saber, ciertos presupuestos epistemológicos que atraviesan la filosofía moderna, generando innecesarias paradojas y falsos problemas. En eso, Peirce se adelanta a toda una tradición, tanto en la filosofía anglosajona como continental, de cuestionamiento de los presupuestos fundamentales de la modernidad. Para nuestros fines actuales comenzaré concentrándome en la crítica que Peirce hace a la idea de que el autoconocimiento es una cierta forma de introspección inmediata e incorregible. El cuestionamiento a lo por él denominado «espíritu del cartesianismo» le permite sostener las siguientes cuatro tesis:

- (1) No tenemos ningún poder de introspección, sino que todo conocimiento del mundo interno se deriva de nuestro conocimiento de los hechos externos por razonamiento hipotético.
- (2) No tenemos ningún poder de intuición, sino que toda cognición está lógicamente determinada por cogniciones previas.
- (3) No tenemos ninguna capacidad de pensar sin signos.
- (4) No tenemos ninguna concepción de lo absolutamente incognoscible (Peirce, 1868, p. I).

De estas cuatro tesis nos interesa especialmente la primera, esto es, la idea de que el autoconocimiento no procede de la inmediata introspección sino más que es el producto del conocimiento de hechos externos. Esa forma pionera de externismo, en la cual ya aparece una concepción triangular del conocimiento, requiere previamente de aclarar algunos conceptos clave del padre del pragmatismo.

Peirce acuñó los conceptos de primeridad, segundidad y terceridad (*firstness*, *secondness* y *thirdness*) que, aunque complejos y difíciles de precisar, pueden ser usados para explicar el punto que nos concierne. Dice Peirce:

Primero es la concepción del ser y del existir independientemente de otra cosa. Segundo es la concepción del ser relativo a algo diferente. Tercero es la concepción de la mediación por la cual un primero y un segundo se ponen en relación (CP 6.32).

Sin pretender describir estos tres conceptos en lenguaje peirceano sino, más bien, tratando de convertir lo extraño en familiar, la primeridad aludiría a un fenómeno

¹ Todas las referencias de Peirce se hacen a los *Collected Papers* editados por Charles Hartshorne y Paul Weiss (1931).

en sí mismo, independientemente de su relación con otros, antes de ser categorizado lingüísticamente, procesado cognitivamente y antes de entrar en relación con otro objeto o con un sujeto. Es cuestionable que eso pueda ser pensado, pero en caso que lo fuese y si tuviéramos que usar otras categorías para aclarar lo que eso significa, podríamos hablar de cualidades primarias o, quizá, con más precisión, de propiedades monádicas. Sin embargo, esos términos no serían del todo exactos, pues la primeridad es la totalidad no diferenciada. Tan pronto un rasgo de esa totalidad es categorizado o descrito lingüísticamente, pasa a ser considerado segundidad. Así, por ejemplo, primeridad aludiría al color azul o a una sensación de dolor, antes de ser llamados así, categorizados cultural o lingüísticamente, y antes de ser parte de una cosmovisión o una teoría.

Cuando el color azul o la sensación de dolor son denominados de esa manera, o simplemente son diferenciados de otros colores o sensaciones posibles, el fenómeno en cuestión entra en relación con otra cosa, emergiendo la segundidad. No me parece incorrecto decir que la segundidad es una propiedad emergente de la primeridad o que puede ser vista como una propiedad relacional. Cuando eso ocurre, aparece ya una relación entre ese fenómeno y otro objeto o un sujeto, aparecen también las distinciones y categorías y, en consecuencia, el procesamiento cognitivo. Pero en la segundidad está incluida la primeridad.

La terceridad es el producto de la relación entre los otros dos y los contiene. Podría decirse, por tanto, que la terceridad es una propiedad emergente de la segundidad y que se constituye cuando dos objetos o fenómenos, relacionados entre sí, entran a su vez en relación con un tercer objeto. «Es algo que pone a un primero en relación con un segundo» (CP 8.332). Peirce asocia la terceridad al surgimiento de la normatividad (CP 1.536), también a la aparición del lenguaje, lo simbólico y la cultura, y al mundo de las convenciones y los hábitos (CP 8.328).

La concepción de lo mental de Peirce está en la misma línea de su concepción de terceridad. Para él, los rasgos más importantes de la mente son la actividad de interpretar signos, la intencionalidad y la capacidad de causar acciones. Pensar es interpretar signos y el pensamiento es relacional en su propia naturaleza, es decir, se forma en la comunicación con otros en relación a un tercero con el que los dos anteriores interactúan simultáneamente.

Donald Davidson y Marcia Cavell, probablemente sin mayor influencia directa de Peirce, han desarrollado el concepto de triangulación en otras dimensiones (Davidson, 2001; Cavell, 1993 y 2006). El objetivo central de Davidson, al desarrollar el concepto de triangulación, es mostrar que la interpretación y el pensamiento son siempre procesos sociales e intersubjetivos que son posibles solo cuando intérprete e interpretado interactúan en relación a un mundo que ambos asumen

compartir y que es, de hecho, real, objetivo y común. También considera que las perspectivas de primera, segunda y tercera persona son inseparables y que se constituyen mutuamente. Adicionalmente, sostiene que el conocimiento de nuestra propia mente, el de las mentes ajenas y el del mundo objetivo son «tres variedades de conocimiento» que conforman una suerte de trípode, tal que si faltara uno de ellos los otros dos no serían posibles.

Marcia Cavell, por su lado, influida por el pensamiento de Davidson y por el psicoanálisis, ha llevado la propuesta triangular al terreno ontogenético. Así, sostiene que la subjetividad del bebé emerge solo cuando este se pone en contacto con otro en relación a un objeto común. Cuando eso ocurre, el bebé no solo comienza a tener propiamente estados mentales sino también a atribuirlos a otra persona, pero siempre acerca de una realidad por ambos compartida. En este punto Cavell explícitamente sigue la línea de George Herbert Mead (Cavell, 2006, p. 62).

Para los autores hasta ahora mencionados, el error fundamental de Descartes, y sus herederos, fue utilizar la introspección, la subjetividad o la perspectiva de la primera persona como punto de partida para derivar las otras dos formas de conocimiento, cuando la subjetividad solo es posible dándose simultáneamente a la intersubjetividad y la objetividad. Habría errado más profundamente Descartes al asumir que es posible alguna forma de autoconocimiento que no pase por el conocimiento de los estados mentales ajenos y de la realidad objetiva.

Integrando las diversas posiciones mencionadas, e intentando ir un poco más lejos que ellas, podríamos sostener un concepto robusto de triangulación, que relaciona los tres siguientes conceptos: yo (subjetividad, primera persona, introspección, autoconocimiento, autointerpretación), otros (intersubjetividad, segunda persona, lectura de mentes, conocimiento de los estados mentales ajenos, aleointerpretación) y realidad compartida (objetividad, tercera persona, conocimiento del mundo exterior).

En la dimensión lógica, los tres conceptos (yo, otros y mundo compartido) se interdefinen mutuamente, en el sentido de que no puede explicarse uno de ellos, o siquiera comprenderse, si no es en relación con los otros dos. En la dimensión ontogenética, las tres experiencias se constituyen simultáneamente, es decir, el niño desarrolla conciencia autobiográfica, esto es, autoconciencia y procesos metacognitivos complejos, cuando también está en condiciones de reconocer a otras entidades del mundo como personas dotadas de subjetividad, poseedoras de estados mentales y de un punto de vista propio. Esto, a su vez, es posible cuando el niño puede reconocer que existe una realidad objetiva que es compartida con los otros y que es diferente de la perspectiva individual que cada uno tiene de ella. Este hito del desarrollo se da entre los dos y los tres años de edad, y su logro suele medirse con las pruebas de la falsa creencia. Cuando esto ocurre, el niño está en condiciones de comprender

las distinciones realidad/apariencia, verdadero/falso y subjetivo/objetivo. Asimismo, desde una dimensión filogenética estas tres formas de conocimiento se habrían dado juntas en la especie, es decir, la autoconciencia o conciencia autobiográfica, la lectura de mentes y la comprensión de la objetividad del mundo exterior, que hace posible su conocimiento, debieron haberse dado juntos en la evolución, aproximadamente cuando los homínidos desarrollaron el lóbulo frontal lo suficiente, en interacción con comunidades sociales complejas, como para atribuir y atribuirse estados mentales en, por lo menos, tercer grado de intencionalidad, esto es, cuando pudieron entender y formular pensamientos del tipo «María cree que Jorge piensa que p», «creo que Jorge piensa que p» o «temo que mi deseo de que p es inconveniente». Es probable que los australopitecinos, quienes vivieron entre cuatro y dos millones de años antes del presente, ya tuvieran rudimentos de estas capacidades. Si creemos que un desarrollo más complejo de estas habilidades requiere de lenguaje, tendríamos que sostener que eso ocurrió hace aproximadamente 500 000 años, en que aparentemente se desarrolló alguna forma de comunicación y pensamiento simbólico, que no involucraba sintaxis, la que debió haber evolucionado hace unos cien mil años.

Puede sostenerse que estas tres formas de conocimiento evolucionaron simultáneamente, con la finalidad de permitir a nuestros ancestros primates adaptarse a situaciones sociales complejas, en la línea de la hipótesis de la inteligencia social (Dunbar, 1998; Byrne & Whiten, 1988). Es razonable pensar que la capacidad de lectura de mentes solo puede florecer cuando el intérprete puede comparar y relacionar los estados mentales atribuidos al agente con sus propios estados mentales, y con los eventos del mundo exterior que él reconoce como reales y que asume que el otro también puede reconocerlos. No parece razonable que uno pueda atribuir un estado mental a otra persona sin ser capaz de comprender qué se siente estar en ese estado mental. En otras palabras: ¿Cómo podría yo saber qué estados mentales atribuir a otra persona, en ciertas circunstancias particulares, si no me pregunto también qué clase de estados mentales tendría yo si estuviera en la situación en que yo creo él está? Si esto es correcto, tanto para tener estados mentales sobre nuestros propios estados mentales, es decir, metacognición, como para poder atribuir estados mentales a los demás, es necesario asumir que el interpretado y uno mismo comparten una misma realidad objetiva, con lo cual los tres vértices del triángulo se constituyen simultáneamente.

En el terreno ontogenético, esto habría coincidido con la aparición de las capacidades de metarrepresentación y simulación, que se desarrollan entre los tres y los cinco años de vida en el niño. La primera es la capacidad de representarse las representaciones ajenas y la segunda es la capacidad de imaginar que la realidad podría ser o podría haber sido diferente de como es. Esto permite al individuo imaginar ser

otro en condiciones contrafácticas, de manera que él pueda atribuirle al agente los estados mentales que él cree que tendría si estuviera pasando por las circunstancias que él cree el agente está pasando, frente a las mismas circunstancias del entorno. Otra manera de formular esto es como la habilidad para imaginar mundos posibles, algo que hasta donde se sabe no está presente en ninguna otra especie excepto el ser humano, a partir de la edad ya mencionada.

Hasta este punto, la tesis que sostengo es que el conocimiento de cualquiera de los vértices del triángulo pasa necesariamente por los otros dos. Esto es, el conocimiento de los estados mentales de otros individuos solo es posible si uno los compara con sus propios estados mentales frente a la realidad objetiva compartida. A su vez, el conocimiento de la realidad objetiva solo es posible si uno compara los estados mentales ajenos con los propios. Análogamente, el autoconocimiento es solo posible si uno compara los estados mentales ajenos con lo que uno asume que es la realidad objetiva compartida, pues son los estados mentales atribuidos a los otros en relación a lo que consideramos que es el mundo que compartimos, lo que nos sirve de criterio para determinar la corrección de nuestros propios estados mentales dotados de contenido proposicional y, así, para constituir, en sentido estricto, conocimiento, en cualquiera de sus tres variedades.

No hay un acceso directo ni inmediato a la objetividad ni a los estados mentales ajenos, y tampoco a los contenidos de los propios estados mentales, esto es, no hay introspección inmediata, sino siempre mediada por la intersubjetividad y la objetividad. Esto no implica, sin embargo, el abandono de la autoridad de la primera persona. Sin duda hay una asimetría entre el autoconocimiento y el conocimiento de los estados mentales ajenos, pero ambas formas de conocimiento requieren de interpretación y de atribución de estados mentales en procesos metacognitivos.

Excepto en el caso de experiencias fenoménicas básicas, como dolores o sensaciones primarias, el autoconocimiento requiere de procesos metacognitivos de, por lo menos, segundo o tercer grado. Así, tenemos creencias acerca de nuestras creencias, deseos o afectos; deseos acerca de nuestras creencias, deseos y afectos; o afectos acerca de nuestras creencias, deseos y afectos. Los ejemplos dados describen procesos metacognitivos de segundo grado, pero con frecuencia el autoconocimiento implica tres o cuatro grados. Es posible, verbigracia, temer que las creencias que uno tiene respecto de sus propios deseos sean falsas. O desear que los afectos que uno tiene respecto de ciertas creencias propias no fueran los que son. O creer que los deseos que tenemos acerca de ciertas emociones (por ejemplo, tenerlas o no) son inevitables. En todos estos casos el autoconocimiento requiere de procesos metacognitivos de, por lo menos, segundo grado y está asociado a la posibilidad de atribuir estados mentales a los demás frente al mundo objetivo común que compartimos.

Para conocer el contenido de nuestros estados mentales no debemos ver hacia dentro, como en la introspección, sino hacia los diversos vínculos que establecemos en relación al mundo que compartimos. Así como el psicoanalista Donald Winnicott solía decir que no es posible observar al bebé si no es en relación con su madre, análogamente no observamos estados mentales (nuestros o ajenos) si no es en relación con otros y con la realidad objetiva acerca de la que tales estados mentales están dirigidos. Tampoco observamos a las personas si no es en sus vínculos con otros y en el contexto del mundo que compartimos. De esta manera, la comunicación no es la transmisión de pensamientos previamente formados en nuestras mentes sino el contexto en que los pensamientos se forman. La comunicación es un proceso simultáneo a la constitución de pensamientos en mentes individuales.

Asumamos, como es comúnmente admitido, que el autoconocimiento es, o por lo menos incluye, la capacidad para conocer los estados mentales de uno mismo. Ya se vio que esto requiere de metacognición en, por lo menos, segundo grado. Lo que deseo sostener ahora, es que uno solo puede atribuirse a sí mismo estados mentales y, sobre todo estados mentales *correctos*, en relación a los estados mentales ajenos y al mundo objetivo compartido. Debe recordarse que el autoconocimiento es una modalidad de conocimiento y, por tanto, implica algún criterio de corrección, pues no es solo auto atribución de estados mentales. Lo que necesito probar, por tanto, es dos cosas.

Para poder atribuirse a uno mismo estados mentales es necesario poder reconocer los estados mentales ajenos y poder compararlos con los de uno mismo y con la realidad compartida.

Para poder atribuirse a uno mismo estados mentales *correctos* es necesario poder reconocer los estados mentales ajenos y poder compararlos con los de uno mismo y con la realidad compartida.

Para ello haré uso de evidencia empírica ontogenética. Más o menos a los tres años, cuando los niños pueden triangular entre su conciencia de sí mismos, sus atribuciones de estados mentales a otras personas y el mundo objetivo que comparten, varias funciones diferentes pero conectadas entre sí aparecen: metacognición, metarrepresentación, la distinción entre realidad y fantasía, y la plena capacidad lingüística. La metarrepresentación requiere de la habilidad para aceptar y comprender la coexistencia de distintos e inconsistentes entre sí sistemas de creencias: las creencias de uno mismo, las atribuidas al yo pasado de uno mismo, las que uno cree que tendría en ciertas circunstancias específicas del futuro, y aquellas atribuidas a otras personas. Los niños menores de tres años no suelen ser capaces de hacer eso, incluso si poseen habilidades de simulación muy básicas.

2. LAS CONEXIONES DEL TRIÁNGULO

Ahora deseo abordar la pregunta sobre la conexión lógica, causal y genealógica entre estos tres tipos de conocimiento: el de los estados mentales ajenos (el uso de la atribución psicológica para lo que suele llamarse *mindreading* o lectura de mentes), el de los estados mentales propios (que es posible gracias a mecanismos metacognitivos y a la conciencia biográfica), y el de los objetos y eventos del mundo exterior, tanto físico como social (percepción sensorial y conocimiento tácito). Sin embargo, me concentraré solo en los dos primeros.

Las propuestas en torno a la triangulación, de los autores mencionados, son originales y pioneras. Lo que ellos no han hecho, sin embargo, es entrecruzar sus tesis con las investigaciones empíricas recientes que proceden de la teoría evolutiva, la psicología experimental y las ciencias cognitivas. Eso es lo que me propongo hacer ahora, con una doble finalidad. Por una parte, deseo expandir el territorio interdisciplinario entre la filosofía de la mente, la epistemología y las ciencias empíricas. De otro lado, me propongo ampliar el triángulo para incorporar una dimensión evolutiva, tanto ontogenética como filogenética, y una causal.

Como es sabido, hay una muy reciente y abundante producción en lo que se llama «filosofía experimental», que empezó a desarrollarse aproximadamente a partir del año 2007, y que sería el uso de evidencia empírica para fundamentar o refutar tesis filosóficas. Tradicionalmente se ha creído que eso no es posible, por asumirse que la filosofía se encuentra en una dimensión diferente de las ciencias empíricas. No quisiera detenerme en esa discusión metafilosófica, solo deseo plantear que si uno tiene una concepción naturalizada de la filosofía y si además uno tiene una concepción post kuhniiana de la ciencia, se puede aceptar fácilmente que la filosofía y la ciencia son territorios superpuestos, de manera que las tesis filosóficas podrían corroborarse empíricamente o, por lo menos, no deberían ser incompatibles con las teorías científicas bien fundadas. La filosofía experimental tiene la ventaja de que permite anclar nuestras reflexiones sobre datos empíricos objetivos, evadiendo así el siempre presente riesgo de que la filosofía sea un tipo de discurso especulativo puramente *a priori*, que corre el riesgo de producir argumentos que recirculen infinitamente sin que haya criterios claros que nos permitan decidir entre ellos.

Ciertamente no creo que la filosofía experimental sea la única forma de hacer filosofía, pero sí creo que es digna de ser proseguida y examinada. Así, en lo que sigue presentaré un panorama de lo que podría ser un programa filosófico en esa dirección. Pienso que la concepción triangular del conocimiento es apropiada para hacer trabajo interdisciplinario en diversas áreas que se pueden corroborar empíricamente como, por ejemplo, el fenómeno de la adquisición del lenguaje y el proceso

por el que el infante construye su *sí* mismo, es decir su *self*; construye su habilidad para atribuir estados psicológicos a los demás, y construye su concepción de realidad objetiva, así como de los conceptos de verdadero y falso.

Como en esta ocasión voy a intentar la ampliación y corroboración empírica del triángulo por el lado del desarrollo y de la evolución, argumentaré en contra de las propuestas que sostienen que hay anterioridad lógica y causal de uno de estos tres tipos de conocimiento, como en el caso de Descartes y Goldman, que privilegian la autoconsciencia, o de Carruthers que sostiene la anterioridad lógica y causal, pero no temporal, de la lectura de mentes, o, finalmente, de alguna forma de empirismo que afirme que la mente es solo una *tabula rasa* que se moldea a partir de la experiencia sensorial, con lo que todo conocimiento procede de la pura sensación.

Comencemos con el autoconocimiento. Durante la mayor parte de la historia de la filosofía occidental, la pregunta por el autoconocimiento ha sido abordada desde un punto de vista conceptual y especulativo, es decir, basada en el análisis de los conceptos involucrados. La concepción cartesiana y moderna que hemos heredado suele suponer que el autoconocimiento es un tipo de conocimiento de la realidad pero, en este caso, de una realidad a la que tenemos un acceso en cierto modo privilegiado y que está compuesta fundamentalmente por estados mentales subjetivos a los que ingresamos por introspección. Así, el autoconocimiento pasaría por el reconocimiento de nuestras propias creencias, deseos y afectos, y de la manera como estos causan nuestras acciones. Según esta concepción tradicional, conocerse sería estar en condiciones de determinar los estados mentales que causan nuestras propias acciones, mediante un acto de reflexión sobre uno mismo.

Aunque esta manera de operar ha sido cuestionada desde comienzos del siglo XX, sobre todo como consecuencia del descubrimiento de que hay estados mentales que causan nuestras acciones pero que, por su propia naturaleza, no son conocidos por nosotros, es una manera válida y necesaria de entender el autoconocimiento, pero no es la única. Así como el autoconocimiento puede ser entendido como un tipo de conocimiento de nuestra propia subjetividad a la que accedemos mediante el reconocimiento introspectivo de nuestros estados mentales, también es posible verlo como un tipo de conocimiento externo. En este caso, averiguar no de manera introspectiva sino empírica, los procesos por los cuales llegamos a determinar nuestros propios estados mentales es también una forma de autoconocimiento pero en un grado de abstracción mayor. Asimismo, saber cómo se constituyeron estos procesos, elaborar una genealogía del reconocimiento de nuestros propios estados mentales y poder establecer las conexiones que eso tiene con otras formas de conocimiento, es una forma de autoconocimiento. El tipo de conocimiento que tenemos a partir de la descripción de los procesos psíquicos que están presentes cuando tenemos procesos

cognitivos acerca de nuestros propios procesos cognitivos, es una forma de autoconocimiento de un nivel adicional, aquel según el cual describimos con cierta pretensión de objetividad lo que ocurre cuando tenemos momentos de introspección subjetiva. Este proceso, a su vez, está conectado y no es separable del proceso por el cual conocemos los estados mentales de otras personas, así como conocemos los hechos del entorno físico y social.

Todo conocimiento es el producto de una relación triangular en la que tenemos, en un vértice, al sujeto que se autoconoce mediante la introspección, es decir, gracias al reconocimiento por experiencia fenoménica de sus propios estados mentales y generando creencias y otros estados mentales de segundo orden acerca de sus propios estados mentales. En otro vértice, el sujeto conoce los estados mentales de otros individuos. En este caso, se trata de creencias acerca de agentes a los que atribuye estados mentales que, supone, son causas de sus acciones, las cuales son eventos físicos que también pueden ser descritos de manera intencional. Finalmente, en el tercer vértice el sujeto conoce el mundo exterior mediante la percepción sensorial, formándose creencias acerca de las relaciones causales entre hechos que, asume, son ontológicamente diferentes de él. Pero, ¿cuál es el orden lógico y causal de los elementos que conforman este triángulo? Es decir, ¿es alguno de ellos temporalmente anterior o condición de posibilidad de los otros?

Sin duda hay criaturas no humanas e incluso humanas, como bebés pequeños, que tienen procesos cognitivos complejos pero no procesos metacognitivos. De ellos podemos decir que tienen formas de comportamiento que involucran interacción con el mundo externo, percepción y procesamiento complejo de la percepción, pero creo que la palabra *conocimiento* sería inaplicable para esos casos. La razón es que para poder decir que alguien sabe o conoce que p es necesario que él o ella también sepa que p podría ser falso, es decir, es necesario que el agente en cuestión posea los conceptos de verdadero y falso, así como de apariencia y realidad. Para que esto sea posible, el agente debe tener procesos metacognitivos de por lo menos segundo grado de intencionalidad, es decir, suficientemente complejos como para tener creencias acerca de sus creencias, por ejemplo, para creer que sus creencias son verdaderas aunque podrían ser falsas.

Adicionalmente, para poder atribuir el concepto de conocimiento a alguien, el agente debe saber que si sus creencias son verdaderas, son acerca del mismo objeto sobre el cual otros individuos tendrían las mismas creencias, si es que tuvieran la evidencia relevante y no hubiese ningún error de procedimiento o percepción. Es decir, el agente debe poseer los conceptos de objetivo y subjetivo. En otras palabras, si uno cree que sus creencias son verdaderas, o sea que constituyen conocimiento, uno debe creer también que los demás tendrían que creer lo mismo en las mismas

circunstancias y frente a la misma evidencia relevante. Esto requiere que el agente tenga la habilidad para atribuir estados mentales a los demás en función a los objetos del mundo real compartido.

La habilidad de atribuir estados mentales en segundo grado aparece en los infantes entre los dos años y los tres años, cuando ellos están en condiciones de aprobar el test de la falsa creencia que mide la posibilidad de reconocer que los demás podrían tener estados mentales diferentes de los que uno tiene. Es a esa edad cuando el niño adquiere el concepto de perspectiva y desarrolla lo que, desde Premack y Woodruff (1978) se llama una *teoría de la mente*, es decir, una concepción acerca de la subjetividad ajena.

Desde los nueve meses de nacidos los bebés tienen una forma rudimentaria de triangulación, en lo que se llama *joint attention* o atención compartida. Ya a esa edad tienen la habilidad de reaccionar frente a los objetos del mundo a partir de las reacciones de otra persona. Pero la triangulación madura solo está presente a una edad cercana a los tres años, cuando el niño adquiere la capacidad de observar a otros individuos relacionándose entre ellos, asumiendo como marco de referencia común los objetos del mundo compartido y reconociendo que existe una realidad objetiva independiente de nuestras percepciones y apreciaciones de ella. Como se ha visto, a esa edad surgen también varias habilidades conectadas con la triangulación madura, como la habilidad metarrepresentacional de representarse no solo la realidad exterior sino también las representaciones que otros individuos tienen de esa misma realidad.

Eso requiere de la capacidad de simulación, que hace posible que un individuo pueda imaginar que la realidad podría ser diferentes de cómo es o que el individuo pueda imaginar ser otro en condiciones contrafácticas, de manera que él pueda atribuirle al agente los estados mentales que él cree que tendría si estuviera pasando por las circunstancias que él cree el agente está pasando. En sentido estricto, la capacidad de simulación es la habilidad de poder imaginar mundos posibles alternativos. La capacidad de simulación no solo nos permite monitorizar el comportamiento de los demás reaccionando ante él, haciendo predicciones y modificando el comportamiento propio en consecuencia, sino, de una manera más general, nos permite comprenderlos.

Antes de los tres años ya encontramos en los niños habilidades simulativas muy elementales, más asociadas con lo afectivo que con lo cognitivo. Estas habilidades parecen ser centrales para la adquisición del significado de las primeras palabras, lo que ocurre alrededor del año y ocho meses, cuando los niños aprenden a observar a los adultos relacionándose verbalmente entre ellos y con los objetos de la realidad (Bloom, 2003; Bloom & Markson 1998).

La capacidad de simulación requiere de la habilidad para asumir y entender la existencia de creencias inconsistentes en *sistemas diferentes*: las de uno en el presente, las atribuidas a uno mismo en el pasado, las que uno cree que tendría en circunstancias específicas del futuro, y las atribuidas a las demás personas. Los niños menores a tres años son incapaces de hacer eso, aunque puedan poseer habilidades de simulación muy básicas.

Los individuos que carecen de capacidad metarrepresentacional, como por ejemplo los autistas severos, carecen también de lenguaje, pero incluso formas de autismo menos severas tienen implicancias lingüísticas importantes. Por ejemplo, los autistas no comprenden oraciones metafóricas ni contrafácticas, es decir, oraciones que exigen la capacidad de imaginar escenarios diferentes a los reales. Tampoco distinguen entre realidad y apariencia, no tienen juego simbólico ni pueden imaginar que las cosas pudieran ser diferentes de como son. De otro lado, la aparición de la capacidad metarrepresentacional casi plenamente desarrollada coincide en los niños con la aparición del lenguaje verbal casi plenamente desarrollado, lo que ocurre entre los tres y cinco años.

Alrededor de los dos años surgen las primeras habilidades sintácticas, lo que permite a los niños construir frases de al menos dos palabras, en las que ya hay una relación predicativa. La capacidad de representar hechos posibles depende de la capacidad de asociar componentes léxicos de manera predicativa, aunque estos sean todavía muy pocos, por las limitaciones de procesamiento y almacenamiento de información de los niños de esa edad. Para algunos lingüistas, como por ejemplo Bickerton (1994) en este momento se constituye un protolenguaje que actúa de manera representacional. Para otros lingüistas lo que hay es una fase sintáctica básica. Hacia los tres años, con el surgimiento de las habilidades metarrepresentacionales, surge también la capacidad sintáctica más compleja y, lo que es de particular importancia, la recursividad, es decir, la capacidad de construir, a partir de un conjunto finito de piezas léxicas, en principio un conjunto infinito de oraciones, así como oraciones en principio infinitamente largas.

Entre los cinco y los seis años madura la capacidad metarrepresentacional en tercer grado de intencionalidad o más, coincidiendo con el fin del periodo crítico para la adquisición de una lengua. Puede decirse que es en este momento que el conocimiento lingüístico central (*core grammar*) está ya instalado con sus diversas competencias: fonológica, morfológica, sintáctica y semántica. Como las capacidades de metarrepresentación y simulación están asociadas al tamaño y desarrollo del lóbulo frontal, se cree que estas habilidades aparecieron en la especie, de manera bastante básica, con los Australopithecus, esto es, hace aproximadamente dos millones de años.

3. ESTRUCTURA Y EVOLUCIÓN DEL TRIÁNGULO

Ahora deseo tomar el camino filogenético para justificar la inseparabilidad de los vértices del triángulo. Discutiré cuatro modelos posibles acerca de las relaciones entre lectura de mentes y metacognición, según los ha planteado Peter Carruthers en «How we Know our Minds: The Relationship between Mindreading and Metacognition» (2009). Carruthers afirma que podemos ver las relaciones entre lectura de mentes y metacognición de cuatro posibles maneras:

- Según el primer modelo, atribuido a Nichols y Stich (2003), lectura de mentes y metacognición serían dos capacidades independientes, con diferentes mecanismos cognitivos, que a su vez pueden tener funcionamientos y posibles daños de manera separada. La capacidad de lectura de mentes incluiría tanto aspectos de simulación como un módulo nomológico. La metacognición, de otro lado, tendría diferentes mecanismos, pero incluiría introspección en el sentido de un acceso inmediato y no interpretativo a nuestros propios estados mentales.
- Según un segundo modelo, que Carruthers considera implícito en un texto suyo de mediados de los noventa (1996a), solo hay un mecanismo metarrepresentacional pero con dos modos diferentes de acceso: un modo basado en la percepción, para la interpretación de las otras personas y un modo introspectivo, para el conocimiento de nuestros propios estados mentales. El daño en el modo basado en la percepción afectaría las habilidades de lectura de mentes, el daño en el modo introspectivo dañaría el autoconocimiento, y el daño en el mecanismo metarrepresentacional afectaría a ambos.
- Para un tercer modelo, denominado también cartesiano y uno de cuyos representantes contemporáneos sería Goldman (1993 y 2006), la metacognición tiene prioridad respecto de la lectura de mentes. Primero tenemos introspección no interpretativa a nuestros propios estados mentales y luego usamos ese conocimiento para atribuir estados mentales a los demás.
- Finalmente, Carruthers propone su actual modelo (2009), de acuerdo con el cual la lectura de mentes tiene prioridad respecto de la metacognición y luego gira hacia nosotros mismos, produciendo metacognición. Hay solo una facultad con dos aplicaciones de ella. Siguiendo ese modelo, Carruthers concluye que:
 - La metarrepresentación evolucionó con propósitos sociales de tercera persona, más que de consciencia y control de la primera persona.

- El autoconocimiento resulta de girar el sistema de lectura de mentes (*mindreading*) hacia el yo.
- El autoconocimiento no tiene autoridad (no hay autoridad de la primera persona), ni es diferente de nuestro conocimiento de los demás.

Deseo proponer un quinto modelo según el cual la metacognición y la lectura de mentes, así como el conocimiento del mundo exterior, se constituyen simultáneamente, desde un punto de vista lógico y causal, y tanto en el orden filogenético como ontogenético. Así, en lo que sigue discutiré los cuatro modelos y luego volveré a mi tesis inicial, según la cual objetaré tanto el modelo cartesiano de Goldman como el de Carruthers, los cuales son víctimas del mismo error: suponer que hay un solo punto de partida para la creación del triángulo, sin percibir que un triángulo solo puede existir cuando desde el comienzo tiene tres bases.

No parece haber mucha diferencia entre los modelos segundo y cuarto, excepto que el último sostiene que el autoconocimiento es siempre interpretativo y, por tanto, no hay un acceso inmediato introspectivo a nuestros propios estados mentales. En su modelo actual, Carruthers sostiene que la capacidad de lectura de mentes evolucionó primero y luego giró sobre nosotros mismos, pero no en un sentido temporal sino lógico y causal. Es decir, él piensa que la presión del entorno seleccionó la lectura de mentes pero no la metacognición, de suerte que esta última es una especie de producto colateral o *by product* de la primera. Él no desea decir que primero hubo lectura de mentes y que posteriormente se dio la metacognición. Carruthers no cree que haya habido un momento en que hubiera lectura de mentes y no metacognición, pero eso suena contradictorio, pues también cree que la lectura de mentes causó la metacognición como un producto colateral, y eso solo es posible si él sostiene que la causa puede no ser temporalmente anterior al efecto, lo que es por lo menos contra intuitivo.

Sin embargo, parece más razonable suponer, como sostengo, que lectura de mentes y metacognición evolucionaron simultáneamente con la finalidad de permitir a nuestros ancestros primates adaptarse a situaciones sociales complejas, en la línea de la hipótesis de la inteligencia social (Dunbar, 1998; Byrne & Whiten, 1988). Es razonable pensar que la capacidad de lectura de mentes solo puede florecer cuando el intérprete puede comparar y relacionar los estados mentales atribuidos al agente con sus propios estados mentales y con los eventos del mundo exterior que él reconoce como reales y que asume que el otro también los reconoce como tales. No parece razonable que uno pueda atribuir un estado mental a otra persona sin ser capaz de comprender qué se siente estar en ese estado mental. En otras palabras: ¿cómo podría yo saber qué estados mentales atribuir a otra persona, en ciertas circunstancias particulares, si no me pregunto también qué clase de estados mentales tendría yo si estuviera

en la situación en que yo creo que él está? Si esto es correcto, tanto para tener estados mentales sobre nuestros propios estados mentales, es decir, metacognición, como para poder atribuir estados mentales a los demás, es decir lectura de mentes, es necesario asumir que el interpretado y uno mismo comparten una misma realidad objetiva, con lo cual los tres vértices del triángulo se constituyen simultáneamente.

El cuarto modelo que propone Carruthers sostiene que la lectura de mentes tiene prioridad sobre la metacognición, aunque no resulta claro qué puede significar *prioridad* en ese contexto. ¿Se trata de una prioridad ontogenética o filogenética? Él sostiene que la lectura de mentes fue seleccionada y la metacognición no, con lo cual tiene que ser una prioridad causal. Carruthers tiende a asumir que esta prioridad y anterioridad se da también en el desarrollo del niño. Sin embargo, incluso en el supuesto negado de que la lectura de mentes haya sido seleccionada primero en la especie, bien podría ser que se haya instalado genéticamente, de suerte que en el desarrollo del niño lectura de mentes y metacognición se den simultáneamente.

El cuarto modelo, propuesto por Carruthers, acepta alguna autoridad de la primera persona basada en una asimetría entre las autoatribuciones y las atribuciones a otras personas. La pregunta es cómo podría explicarse tal asimetría sin algún grado de introspección. Parece posible que nos interpretemos en una dimensión diacrónica, es decir, que nos atribuyamos a nosotros mismos estados mentales de una manera prospectiva o retrospectiva, esto es, considerando nuestros yoes futuros o pasados. Pero, ¿habrá autointerpretación en una dimensión sincrónica, es decir, en tiempo real? La conclusión de Carruthers es que la metarrepresentación evolucionó con propósitos sociales de tercera persona, pero no de autoconciencia ni de control de la primera persona.

No resulta claro, sin embargo, por qué ambas cosas tendrían que ser incompatibles. ¿Por qué no podría ser que la metarrepresentación fuese posible solo si puede ayudar a predecir el comportamiento de los otros pero que, a su vez, eso solo ocurre cuando el comportamiento ajeno puede ser comparado con las posibles reacciones que uno tendría frente a ellos? Para poder atribuir un sistema de estados mentales a otro individuo uno tiene que ser capaz de compararlo con sus propios estados mentales. Para atribuir exitosamente estados mentales a otras personas, es necesario tener la habilidad de simular *ser* ellos en condiciones contrafácticas, es decir, es necesario tener la capacidad de atribuirles el tipo de estados mentales que nosotros creemos que tendríamos si *nosotros* fuéramos quienes creemos que son los otros. De esa manera no solo imaginamos lo que pasa en sus mentes, sino también lo que pasaría en sus mentes si nos comportásemos de una manera o de otra. Al hacer eso, podemos predecir sus acciones y planificar los tipos de acciones que nosotros realizaríamos si ellos se comportaran de la manera como nosotros predecimos lo harán.

Esto significa que lectura de mentes y metacognición son inseparables y tendrían que haber evolucionado simultáneamente, sin prioridad de una por sobre la otra.

El descubrimiento de las neuronas espejo también podría ser un argumento a favor de que no puede haber lectura de mentes sin metacognición. Las neuronas espejo son activadas cuando un individuo realiza una acción (como por ejemplo tomar un objeto) o cuando el individuo observa a otros individuos realizando la misma acción (Perrett y otros, 1990; Rizzolatti & Craighero, 2004). También se cree que la corteza prefrontal se activa tanto cuando uno evalúa sus propios estados mentales como cuando uno evalúa los ajenos (Frith & Frith, 2003). Más aún, la corteza prefrontal de los primates parece haber coevolucionado con sus habilidades cognitivas, especialmente la capacidad de abstracción, en respuesta a las demandas de una sociedad socialmente compleja (Cheney & Seyfarth, 2007, p. 133). Las neuronas espejo están involucradas en el uso y en la adquisición del lenguaje, lo que nos conduce al siguiente punto. Es probable que la metarrepresentación y la metacognición plenamente desarrolladas requieran de lenguaje, tanto en las dimensiones filogenética como ontogenética. Esto resulta claro si uno toma en consideración la hipótesis del cerebro social.

Esta hipótesis sostiene que la causa principal del crecimiento del cerebro de nuestros ancestros homínidos (de unos 450 centímetros cúbicos hace aproximadamente tres millones de años, a 1500 centímetros cúbicos en promedio en la actualidad), fue la presión para competir al interior de grupos de individuos con relaciones sociales complejas. Esto habría conducido a sofisticadas funciones neurológicas y psicológicas, tales como la habilidad de predecir las acciones de otros individuos, atribuyéndoles los estados mentales que causan y explican su comportamiento, así como el reconocimiento de sus intenciones, sean estas amigables o agresivas. La planificación social requiere de formas de razonamiento que nos permitan simular escenarios futuros en los cuales podríamos vernos envueltos. Imaginar tales situaciones posibles y elegir entre ellas, nos permitiría probar el éxito o fracaso de nuestras acciones por adelantado. En esos casos parecería extraño que haya lectura de mentes sin metacognición.

Como he intentado sostener, el autoconocimiento requiere de la constitución de una relación triangular, en tanto, la consciencia que el sujeto tiene de sí mismo solo es posible cuando el sujeto es también consciente de la existencia de otros y de un mundo objetivo que asume compartir con ellos. Pero, además, el autoconocimiento requiere, como una condición esencial, conciencia de la identidad personal o sí mismo, que es lo que Damasio ha llamado conciencia extendida o autobiográfica. Ella presupone conciencia del tiempo y esto, a su vez, hace necesario otro tipo de triangulación entre el pasado, el presente y el futuro. Para que uno tenga conciencia de su propia identidad es imprescindible que tenga la capacidad de reconstruir su

pasado desde su propio presente, con frecuencia resignificándolo, es decir, reinterpretándolo a la luz de los estados mentales y acontecimientos que el individuo vive en el presente. Esto requiere de memoria episódica pero, además, el sujeto debe tener la capacidad de plantearse posibles escenarios futuros, esto es, debe tener la capacidad de simular mundos posibles alternativos que pueden emerger de posibles decisiones que el sujeto adopte en su situación presente. Así, el sujeto tomará en consideración los que él considera son los hechos actuales y los acontecimientos pasados de los cuales esos hechos se han derivado, y los comparará con las diversas simulaciones que realiza sobre el futuro. El producto de esa triangulación entre pasado, presente y futuro es múltiple: por una parte, permite al agente tomar las decisiones que él considera apropiadas para que se produzcan los hechos objetivos que él desea que tengan lugar, así como los estados mentales subjetivos que él espera que acontezcan en su «mundo interior». Pero, por otra parte, se constituye también la conciencia subjetiva del tiempo (Boroditsky, 2000)².

Finalmente se constituye la identidad personal, es decir, nuestro espacio subjetivo propio, el *sí mismo*, el *self*, entendido no solo como los estados mentales que el sujeto tiene acerca de sus estados mentales, sino también como los estados mentales no conscientes e inconscientes que participan causalmente en la conducta de la persona. El conocimiento e interpretación que uno tiene del mundo exterior se da también a partir de la comparación entre los estados mentales que uno tiene, los que atribuye a los demás, y los hechos que considera compartir con los otros. Tenemos, de esta manera, varios triángulos que se van integrando y constituyendo entre sí para dar lugar al conocimiento de la realidad en que vivimos, tanto interna como externa.

Ahora giremos ahora nuestra mirada al lenguaje y a la manera triangular como este se instala en el niño. Según la última versión del programa minimalista de Chomsky (2002), la característica del lenguaje humano, entendido en el sentido específico, es la recursividad, esto es, la capacidad de iterar, en principio de manera infinita, diferentes componentes sintácticos, ya sea con oraciones subordinadas o con conectores lógicos. La recursividad lingüística, bajo esta concepción, tiene un correlato neurológico que aparentemente solo se daría en humanos. Ahora bien, la recursividad también está presente en la metarrepresentación, que es condición necesaria para atribuir estados mentales a los demás, pues para poder atribuir un sistema de estados mentales a otro individuo, uno debe ser capaz de compararlo con sus propios estados mentales. En el caso de la metarrepresentación en tercer grado, como el contra engaño, la existencia de recursividad es aún más evidente. Este sugiere alguna conexión entre lenguaje y metarrepresentación. Como ya se ha visto, los individuos

² Boroditsky considera que desarrollamos nuestro sentido del tiempo por analogía con nuestro sentido del espacio.

que carecen de metarrepresentación, como por ejemplo los casos severos de autismo, carecen también de lenguaje; pero incluso formas menos severas de autismo también tienen importantes consecuencias lingüísticas. De otro lado, la aparición de metarrepresentación plenamente desarrollada (a los cinco años con metarrepresentación en tercer grado), es casi simultánea con la aparición de lenguaje casi plenamente desarrollado, entre los tres y los cinco años.

La recursividad es una condición necesaria, tanto para la existencia de lenguaje humano como para la capacidad metarrepresentacional. No ha sido probado que los primates no humanos sean capaces de algún tipo de lenguaje recursivo y, aunque no hay pruebas definitivas y no controversiales de que posean metarrepresentación, hay fuertes evidencias de que tienen formas básicas de simulación (Kuroshima, Fujita, Adachi, Iwata & Fuyuki, 2003). Los chimpancés muestran cierta capacidad de este tipo cuando interactúan con otros individuos uno a uno, especialmente en operaciones de engaño y contra engaño. Sin embargo, cometen muchos errores cuando tienen que atribuir diversos estados mentales a diferentes individuos, es decir, cuando interactúan con muchos individuos al mismo tiempo. La evidencia sugiere que, aunque es posible que haya algunas formas básicas de simulación en primates no humanos y en niños menores al periodo de adquisición de lenguaje, la plena metarrepresentación está presente solo cuando hay lenguaje. Esto implicaría que la capacidad metarrepresentacional plenamente desarrollada no habría estado presente antes de unos 100 000 años atrás, aunque podría haber habido formas básicas de simulación, conectadas con lo que Bickerton (1994) denomina un protolenguaje, que sería básicamente un lenguaje moderno sin sintaxis y sin recursividad (Jackendoff, 1999). De acuerdo con Bickerton, el protolenguaje estaría todavía presente en los humanos, lo cual sería especialmente evidente durante el proceso de adquisición del lenguaje y también en casos en que la adquisición es abruptamente interrumpida.

El punto es que las formas básicas de triangulación pueden encontrarse en los primates superiores como los chimpancés quienes, según algunos experimentos, podrían pasar la prueba de la falsa creencia y, por tanto, tener niveles rudimentarios de metarrepresentación. Estas son habilidades que ninguna otra especie parece tener y que, precisamente, están instaladas genéticamente en los bebés humanos, con lo cual se activan como módulos innatos cuando el desarrollo del bebé alcanza ciertos hitos de madurez.

Como hemos visto, la triangulación se empieza a constituir a nivel ontogenético desde una edad muy temprana. Un bebé de pocos meses de nacido ya establece relaciones entre su cuidador más cercano, él mismo y los objetos más significativos en su corta vida. De hecho, hacia los nueve meses el bebé reacciona con mucha claridad a las reacciones que el cuidador tiene acerca del mundo exterior que comparten y,

para los doce meses, el bebé ya señala con el dedo, con lo cual claramente establece un triángulo entre él, el cuidador a quién señala y el objeto señalado. El retraso o la ausencia de la capacidad de señalar suele ser un indicador temprano de algún tipo de trastorno en el desarrollo, con mucha frecuencia, autismo.

En casos normales, sin embargo, parece razonable afirmar que la lectura de mentes y la metacognición maduran más o menos simultáneamente, aproximadamente a los tres años, lo que coincidiría con la adquisición de un lenguaje casi plenamente desarrollado y con la aparición de los conceptos de verdadero y falso. Esto también coincidiría con la aparición de las bases del principio de realidad, esto es, con la distinción entre realidad y apariencia, u objetividad y subjetividad. En otras palabras, coincidiría con la distinción entre la manera como las cosas son objetivamente, como me parecen a mí y como te parecen a ti.

Por todo lo sostenido anteriormente, las habilidades de metacognición y lectura de mentes habrían sido seleccionadas por la selección natural, simultáneamente y por las mismas razones, por lo cual tendrían que ser lógica y causalmente inseparables.

REFERENCIAS BIBLIOGRÁFICAS

- Barton, Robert & Robin Dunbar (1997). Evolution of the Social Brain. En Andrew Whiten y Richard Byrne (eds.), *Machiavellian Intelligence II: Extensions and Evaluations* (pp. 240-263). Cambridge: Cambridge University Press.
- Bickerton, Derek (1994). *Lenguaje y especies*. Madrid: Alianza.
- Bloom, Paul (2003). *How Children Learn the Meaning of Words*. Cambridge: The MIT Press.
- Bloom, Paul & Lori Markson (1998). Capacities Underlying Word Learning. *Trends in Cognitive Science*, 2, 67-73.
- Boroditsky, Lera (2000). Metaphoring Structuring: Understanding Time through Special Metaphors. *Cognition*, 75, 1-28.
- Byrne, Richard & Andrew Whiten (1988). *Machiavellian Intelligence*. Oxford: Oxford University Press.
- Carruthers, Peter (1996a). Autism as Mindblindness. En *Theories of Theories of Mind* (pp. 257-276). Cambridge: Cambridge University Press.
- Carruthers, Peter (2009). How We Know Our Own Mind. The Relationship between Mindreading and Metacognition. *Behavioral and Brain Sciences*, 32, 121-182.
- Carruthers, Peter & George Botterill (1999). *The Philosophy of Psychology*. Cambridge University Press.
- Cavell, Marcia (1993). *The Psychoanalytic Mind. From Freud to Philosophy*. Cambridge: Harvard University Press.

- Cavell, Marcia (2006). *Becoming a Subject*. Oxford: Oxford University Press.
- Cheney, Dorothy & Robert Seyfarth (2007). *Baboon Metaphysics. The Evolution of a Social Mind*. Chicago: The University of Chicago Press.
- Chomsky, Noam, W. Tecumseh Fitch & Marc D. Hauser (2005). The Evolution of the Language Faculty: Clarifications and Implications. *Cognition*, 97, 179-210.
- Damasio, Antonio (1999). *The Feeling of What Happens. Body and Emotions in the Making of Consciousness*. Nueva York: Harvest Book.
- Davidson, Donald (2001). *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- Dunbar, Robin (1998). The Social Brain Hypothesis. *Evolutionary Anthropology*, 6, 178-190.
- Frith, Uta & Christopher Frith D. (2003). Development and Neurophysiology of Mentalizing. *The Royal Society of London B*, 358(1431), 459-473.
- Goldman, Alvin (1993). The Psychology of Folk Psychology. *Behavioral and Brain Sciences*, 16(1), 15-28.
- Goldman, Alvin (2006). *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. Oxford: Oxford University Press.
- Horowitz, Alexandra (2009). *Inside of a Dog: What Dogs See, Smell, and Knows*. Nueva York: Harvest Books.
- Jackendoff, Ray (1999). Possible Stages in the Evolution of the Language Capacity. *Trends in Cognitive Science*, 3(7), 272-279.
- Kuroshima, Hika, Kazuo Fujita, Ikuma Adachi, Kana Iwata & Akira Fuyuki (2003). A Capuchin Monkey (*Cebus apella*) Recognizes When People Do and Do Not Know the Location of Food. *Animal Cognition*, 6(4), 283-291.
- Nichols, Shaun & Stephen Pastiche (2003). *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding other Minds*. Oxford: Oxford University Press.
- Peirce, Charles Sanders (1931). *Collected Papers*. Edición de Charles Hartshorne y Paul Weiss. Cambridge: Harvard University Press.
- Perrett, David I. y otros (1990). Social Signals Analysed at the Single Cell Level: Someone is Looking at Me, Something Touched Me, Something Moved. *International Journal of Comparative Psychology*, 4(1), 25-55.
- Premack, David & Guy Woodruff (1978). Does the Chimpanzee Have a Theory of Mind? *Behavioral and Brain Sciences* [número especial], 1(4), 515-526.
- Rizzolatti, Giacomo & Laila Craighero (2004). The Mirror Neuron System. *Annual Review of Neuroscience*, 27, 169-192.

SECCIÓN II
ENSAYOS SOBRE DESARROLLO Y ESTRUCTURA
DE LA COGNICIÓN SOCIAL

EL EMERGENTISMO Y LA APARICIÓN DE LA CONCIENCIA¹

Ricardo Braun

Universidad de Lima

En este ensayo quiero explorar el origen de la conciencia a través del concepto de emergentismo, el cual ha cobrado especial importancia en el desarrollo de teorías de la mente que tienen la pretensión de ser materialistas y, al mismo tiempo, no-reduccionistas. Precisamente, la autonomía explicativa de la psicología depende del anclaje metafísico y epistemológico que tengan las entidades, metodología y vocabulario que la disciplina postula y emplea. Y claramente, la psicología contemporánea mantiene estos tres componentes en forma no reducida. El emergentismo, como teoría filosófica de largo recorrido histórico, puede proporcionar los elementos teóricos para conceptualizar mejor el fenómeno de la conciencia humana.

Divido el ensayo en tres acápites. El primero está dedicado a la aclaración conceptual de emergentismo y otros conceptos asociados. En el segundo hago una revisión histórica del origen de las teorías emergentistas. Finalmente, el tercero está dedicado a analizar dos de estas teorías contemporáneas, que escojo por su notoriedad e influencia: la de J. S. Searle y la de D. Chalmers.

1. ACLARACIÓN CONCEPTUAL

1.1. Los conceptos de emergentismo, conciencia y otros asociados

El término *emergencia* es utilizado en diferentes sentidos, pero no nos concentraremos en su uso popular, sino más bien en el académico. Este concepto es usado generalmente en las doctrinas de orientación no-reduccionistas tanto de la biología como de la psicología y otras disciplinas sociales.

¹ Agradezco al Instituto de Investigaciones Científicas de la Universidad de Lima por el apoyo en esta investigación. Asimismo, al Fondo Editorial de la Universidad de Lima por permitir reproducir una sección publicada anteriormente en la revista *PERSONA*.

Desde una perspectiva metafísica, una postura emergentista concibe que la realidad existe y se presenta en forma plural, de ahí que, en primer lugar, se pueda afirmar que el emergentismo es una forma de pluralismo (Bunge, 1977). Sin embargo, el pluralismo que le da sentido debe admitir que existen diferentes niveles de organización que son ontológicamente de igual valor.

En segundo lugar, el emergentismo sostiene que en cada nivel superior de organización aparecen (o emergen) nuevas e irreducibles propiedades que no están presentes en el nivel inferior. Estas propiedades serían propiedades *emergentes*, predicadas de un todo organizado y no de sus partes o componentes. Otra forma de presentar esta tesis es sosteniendo que las propiedades de mayor nivel son de segundo orden, pues son el resultado (o emergen) de las propiedades del primero.

Desde una perspectiva epistemológica, cualquier doctrina emergentista supone que el conocimiento de las propiedades emergentes no puede ser derivado del conocimiento de las propiedades de las partes que componen el ente, y, por tanto, que las leyes que se puedan formular en ese nivel no son deducibles de las que se obtienen de los niveles inferiores.

Recientemente el proyecto multinacional Human Strategies in Complexity. Philosophical Foundations for a Theory of Evolutionary Systems (Arshinov & Fuchs, 2003) ha tratado de identificar las características que tienen en común las teorías clásicas y contemporáneas de la emergencia:

- *Sinergismo*. La emergencia ocurre como producto de la interacción entre las entidades. La sinergia sería una propiedad de la materia que se manifiesta en la cooperación (sea de partes, elementos o individuos) produciendo efectos. La sinergia entre los individuos que interactúan sería la causa de la evolución de la persistencia de los sistemas emergentes.
- *Novedad*. En el nivel superior al de las entidades, interactuando sinérgicamente, aparecen nuevas cualidades. Estas cualidades no habrían sido previamente observadas y no habrían existido en otro sistema. Esta característica se podría resumir en el conocido principio que el todo es mayor que la suma de sus partes.
- *Irreducibilidad*. Las nuevas cualidades producidas no serían reducibles al o derivables del nivel de las entidades productoras.
- *Impredictibilidad*. La forma de la resultante emergente y el tiempo de la emergencia serían impredecibles.
- *Coherencial correlación*. Los sistemas complejos que tienen cualidades emergentes tendrían un comportamiento coherente por un cierto periodo de tiempo.

- *Historicidad.* Las cualidades emergentes no están predeterminadas sino aparecen como resultado del desarrollo dinámico de los sistemas complejos.

Lo valioso de esta caracterización de cualquier teoría emergente radica en que el trabajo interdisciplinario se nutre del aporte de científicos, filósofos y académicos de diferentes disciplinas, desde las ciencias formales hasta las lingüísticas y semióticas, para quienes las teorías con orientación emergentista podrían convertirse eventualmente en nuevos paradigmas para la ciencia actual y del futuro.

Aun cuando solemos asociar el concepto de emergencia con las ciencias biológicas y humanas, como se mencionó anteriormente, es importante aclarar que las propiedades emergentes pueden ser biológicas como no. Lo que se requiere para que ocurra un fenómeno de tal naturaleza en cualquier aspecto de la realidad es la súbita aparición de una nueva cualidad, la cual puede durar desde un rango de microsegundos hasta miles de años.

Según Reid, la emergencia biológica e intencional se distingue de las no-biológica por al menos cuatro fenómenos (2007). El primero es la reproducción biológica, que proporciona a patrones emergentes de complejidad viviente alguna garantía de persistencia. Un segundo fenómeno es el de integridad dinámica o también podría denominarse *automantenimiento*. Algunos consideran que esta característica es esencial en los fenómenos emergentes y lo denominan *autoorganización* (Arshinov & Fuchs, 2003, p. 6). Los dos primeros fenómenos darían lugar al tercero, la habilidad para conducir experimentos naturales en la emergencia evolutiva. Finalmente, el cuarto, que define la emergencia biológica y además intencional, sería también de tipo experimental: los organismos pueden ajustar su integridad dinámica en diferentes entornos ambientales, y así inducir cambios fisicoquímicos y biológicos en sí mismos. El incremento de la complejidad en los organismos o la adquisición de nuevas cualidades emergentes pueden ser igualados con el incremento de libertad para dirigirse a actividades específicas. Este cuarto fenómeno, entre otros, podría explicar el paso de la evolución bioquímica que, en el caso humano, nos llevó desde el comportamiento intencional hasta la aparición del lenguaje en tanto confrontación del cerebro desarrollado con el entorno de supervivencia encontrado (Bickerton, 2009)².

² La tesis del origen del lenguaje de Bickerton (2009) es consistente con esta clasificación de los fenómenos emergentes de Reid. Para el primero no solo fue necesaria la capacidad volumétrica cerebral sino la necesidad de buscar aliados para la actividad carroñera que nuestros antepasados tuvieron que emprender para su supervivencia. Según Bickerton, para poder sacar la carne de los huesos que nos dejaban los predadores más hábiles estructuralmente, tuvimos que requerir ayuda de otros humanos y para ello empezamos a utilizar lenguaje —y no solo sistemas de comunicación— que incorporó tiempos verbales, desplazamiento lingüístico y exaptación.

Una forma de entender el emergentismo es también describirlo en relación con una posición antagónica, a saber, el materialismo reduccionista y el dualismo. El reduccionismo aboga por las explicaciones y comprensión ontológica en términos y entidades, respectivamente, de las partes constituyentes de algún sistema físico. Así, una teoría fundamental fiscalista sería más profunda e incluyente, tendría un mayor poder explicativo y predictivo y, en consecuencia, una comprensión mayor del mundo. En este caso, la presuposición ontológica implícita es que la realidad física, sea lo que fuere, es la fuente *real* de todo lo existente, y cualquier cosa que quiera mantener el estatus de real debería ser susceptible de ser descrita o construida a partir de aquellos elementos de esta ontología fundamental.

El dualismo, por su lado, afirma el rol causal eficiente para ciertos fenómenos, tales como las actividades de un alma o mente, los cuales no podrían ser derivados, de ninguna forma, de las propiedades físicas, pues, además de no compartirlas, se oponen a ellas sustancialmente. Prácticamente en la filosofía contemporánea, así como la psicología académica, no hay lugar para el dualismo sustancial cartesiano. Sin embargo, muchas de nuestras discusiones, y en particular, este mismo trabajo, nos muestran que el reduccionismo no nos llega a convencer universalmente y la mayoría de filósofos de orientación analítica adoptan alguna versión del materialismo no-reduccionista, sin estar muy satisfechos con la cantidad de contraejemplos que se presentan constantemente. Y, este descontento, en parte, se debe a que no hemos desarrollado una teoría suficientemente consistente del fenómeno de la conciencia. En las palabras de Nagel, defensor por décadas de la irreductibilidad de lo mental:

Lo que necesitamos es algo que no tenemos: una teoría de los organismos conscientes como sistemas físicos compuestos de elementos químicos y que ocupan un lugar, que tienen también una perspectiva individual en el mundo, y [...], en algunos casos, una capacidad para la autoconciencia, también [...] La extraña verdad parece ser que ciertos sistemas físicos complejos biológicamente generados, de los cuales cada uno de nosotros es un ejemplo, tiene ricas propiedades no-físicas (1985, p. 51)³.

Por ello, cualquier teoría emergentista que pudiera concebirse tiene que tener condiciones como las señaladas por Crane, que restringen adecuadamente el concepto de emergencia dentro del consenso materialista mencionado (2001)⁴. Para este autor los requisitos serían la distinción y la dependencia. La *distinción* significa que

³ Todas las traducciones de los textos en inglés o en idioma distinto al castellano han sido realizadas por el autor [nota de editores].

⁴ Aunque Crane señala estas dos características como descriptivas del fiscalismo no reduccionista, a las que llama «las dos tesis», considero que estas pueden ser incorporadas a un emergentismo fiscalista.

las propiedades mentales son diferentes de las propiedades físicas y la *dependencia*, que ellas son propiedades de los objetos físicos (2001, p. 208). Las teorías reduccionistas de lo mental, asociadas a la teoría de la identidad tipo-tipo postulada fundamentalmente por J. J. C. Smart (1962), sostienen que las propiedades mentales son idénticas a las físicas. Por tanto, de una parte podríamos afirmar que la teoría de la identidad sería contraria a la noción de distinción y, que, de otra, las formas del dualismo cartesiano estarían negando el requisito de dependencia.

En consecuencia, la noción de emergencia, con los requisitos de distinción y dependencia, implicaría que las propiedades emergentes —mentales, para el caso de la filosofía de la mente— serían distintas a las físicas y, además, que el dualismo de tipo cartesiano sería falso. Más aún, la idea de que algunas propiedades emergen del sustrato físico conllevaría a que ellas no puedan ser explicadas ni predichas desde las ciencias de los sustratos físicos, es decir, desde las ciencias físicas. No obstante, a pesar de esta *distinción*, la *dependencia* de dichas propiedades con lo físico es innegable: si destruyéramos las moléculas que componen las neuronas, estas dejarían de existir y la actividad cerebral cesaría, o si usáramos drogas alucinógenas, estas alterarían significativamente nuestra actividad mental cognitiva, entre otros ejemplos.

1.2. Conciencia y eventos mentales

Una de las lecciones que hemos aprendido de las críticas al positivismo lógico es que algunos problemas son genuinos problemas y que las preguntas a esos problemas no son pseudopreguntas. Antes, como puede deducirse, el positivismo había agrupado a las preguntas sobre la conciencia dentro de las pseudopreguntas. Hoy, naturalmente, pensamos diferente. La conciencia está de nuevo con nosotros y desde hace ya varias décadas. Por tanto, su análisis no puede estar ausente de las discusiones sobre el estatus de las entidades y procesos de la psicología académica y sus cognaticios.

Ahora bien, las teorías emergentes tienen el mérito de enfrentar el problema de la conciencia desde su origen y, también, desde su existencia como entidad. Sin embargo, aún no existe una teoría unificada. Dado que los sentidos de este concepto son múltiples, su evaluación como constructo válido se hace más difícil, no solo para su comprensión sino también para su confrontación en el plano empírico. Por poner un ejemplo, en el pasado, la psicología científica estudiaba procesos que no eran conscientes, como los estudios que llevaron a Chomsky a la concepción de una gramática innata que se basaba en el conocimiento inconsciente de las reglas del lenguaje. Otras serán las conclusiones que veremos cuando se aborden más estudios desde el punto de vista de la conciencia, que, dado el prejuicio conductista que imperó hasta hace algunas décadas cercanas, fueron descuidados.

Por ello, desarrollaré un breve análisis del concepto de conciencia con miras a proporcionar el análisis conceptual necesario para cualquier discusión posterior.

Según N. Block, el concepto de conciencia (tanto académico como coloquial) es, en el fondo, la combinación de varios conceptos:

La conciencia es un concepto híbrido: existen varias «conciencias» muy distintas entre sí. La conciencia fenomenal es la experiencia; el aspecto fenoménico consciente de un estado es aquello que se siente al encontrarse en ese estado. La marca de la conciencia de acceso, por contraste, es la disponibilidad de usar, en el razonamiento, y guiar, racionalmente, el habla y la acción (1995, p. 1).

Claro, que para Block el resultado de este híbrido es cerca de desastroso para los efectos de la investigación, puesto que lo que se estudie sobre uno puede resultar contradictorio o confuso en otro.

Ciertamente la posición de este autor es incuestionable, pero las funciones denominadas mentales que suelen ser inconscientes, como la mencionada sobre la capacidad lingüística, suelen ser menos problemáticas para asimilarlas a funciones asociadas al modelo computacional, bien sea conexionista u otras alternativas presentes. Con todo, de los mencionados, dos los son los conceptos que deben ser especialmente tomados en cuenta al desarrollar una teoría emergente: el de la conciencia fenoménica y el de la autoconciencia.

La conciencia llamada fenoménica es definida con la celeberrima frase de T. Nagel «¿Qué es ser como un murciélago?» (1974). Esta frase resume la experiencia de ser un sujeto que experimenta estados cualitativos como ver el color rojo de una manzana, sentir el sabor del chocolate, escuchar la obertura de Parsifal, etcétera, esto que los filósofos llaman los *qualia*. En otras palabras, ¿cómo es que tenemos esas sensaciones de la manera tan particular como las experimentamos? Estas sensaciones, además, tienen la propiedad de ser percibidas por nosotros mismos junto a otras simultáneamente, que presumiblemente deben ser procesadas por diferentes partes de nuestro cerebro.

La otra concepción de conciencia hace alusión a nuestra capacidad para dirigir nuestras acciones, controlarlas racionalmente, decimos, a partir de una reflexión que hacemos de nosotros mismos. Nos damos cuenta de que somos agentes, o si se quiere, de que somos un *yo*, una persona. La autoconciencia supone que uno sabe de sí mismo, pero no que uno esté en lo correcto. Así, por ejemplo, yo me doy cuenta de que estoy con un dolor de estómago, aunque más tarde un médico me revele que tengo un cálculo en el riñón. No obstante, lo que el médico no puede desconformar es mi sensación, ese darme cuenta de que estoy teniendo un dolor (ahora corregido a riñón).

Una teoría emergente debe, idealmente, dar cuenta de cómo es posible que de elementos, como las partículas elementales, se generen estos dos tipos de conciencia (y acaso otros más), partiendo de la suposición que esas partículas no son conscientes en los sentidos explicados. Descartando la existencia de mentes inmateriales, sustancias ideales conscientes y otras suposiciones que tomamos como patentemente falsas, existe evidentemente una brecha explicativa entre los elementos fundamentales o bloques primarios de la realidad, que dan evidencia de no ser conscientes, y los sujetos, como nosotros, que tienen la experiencia de ser altamente conscientes. Lo que se requerirá —y hay que admitir que no tenemos tales teorías en el momento— es una teoría o teorías que expliquen cómo lo que no es consciente se convierte en tal.

1.3. Causación emergente

Toda teoría emergentista tiene que enfrentar el problema de la causación efectiva. Si aparecen nuevas propiedades a partir de otras de un nivel inferior, se esperaría que estas propiedades tengan un real rol causal, en virtud de las nuevas. Si nos limitamos a los actos conscientes o de la conciencia tendríamos que preguntarnos, ¿cómo es posible que ella sea causalmente eficaz? O de forma más precisa, ¿cómo es concebible que los estados conscientes *qua* conscientes puedan interactuar con estados neuroquímicos *qua* neuroquímicos? O si se quiere simplificar en un lenguaje ordinario, ¿cómo es posible la causación mental?

Para algunos, la causación mental es una realidad ontológicamente anclada en el mismo nivel superior mental y, en consecuencia, compatible con el discurso emergentista. Esta es la posición que defiende en este trabajo. Para otros, la conciencia tiene un papel meramente funcional y, en consecuencia, el mecanismo neural es el causante de todo lo que denominamos causación mental (Block, 1995). Otros están convencidos de que, dado nuestro conocimiento científico del mundo, la eficacia causal mental está excluida (Levine, 2001). Finalmente, hay quienes simplemente abandonan cualquier intento de solución y adoptan una posición escéptica ante la posibilidad de explicar la eficacia causal de lo mental (Baker, 1993). Sin embargo, a pesar de los cuestionamientos que encontramos en la literatura filosófica, desde el punto de vista de nuestra vida cotidiana, nos parece que nada podría ser más certero que la eficacia de nuestra conciencia⁵.

Al estudiar a diferentes autores que propugnan una teoría emergentista debemos examinar en qué medida se cumple la exigencia de dar una explicación causal a la conciencia y los actos mentales, puesto que no tendría sentido afirmar la existencia de niveles ontológicos autónomos y, al mismo tiempo, negar su eficacia causal o dejarla inerte.

⁵ Es esta la intuición cartesiana de lo cual no podría dudar, dude lo que quisiera.

Pero postular la eficacia causal de la conciencia, siendo un gran desafío, dentro de una postura fisicalista, permitiría naturalizarla. Las tendencias eliminativistas, por atractivas que parezcan, nos crearían la ilusión de la conciencia causal, y, sin embargo, nuestra vida y nuestro autoconcepto parten de la experiencia de la conciencia misma. Hay algo perturbador en que por un lado nuestros actos conscientes sean ubicuos en nuestra vida y, si seguimos las conclusiones eliminativistas, sean inexistentes. La naturalización de la conciencia pasa por darle un lugar dentro de la evolución adaptativa. Por azaroso que haya sido el proceso evolutivo, debe haber habido un fin adaptativo en nuestros ancestros que dio lugar a la aparición de la complejidad estructural y biológica que lleva al fenómeno de la conciencia y de toda nuestra vida mental. Y también tendríamos que tener en cuenta en esto último la ingente cantidad de recursos energéticos corpóreos que supone el mantenimiento de los fenómenos mentales.

Indudablemente el problema de la causación surge en el contexto de cualquier teoría emergentista. En efecto, uno de los problemas que ronda al emergentismo permanentemente es la posibilidad de que la realidad emergente de la que se predica resulte ser meramente un epifenómeno sin un rol causal eficiente o peor aún, ficticio.

1.4. Epifenomenalismo

El epifenomenalismo es la postura que sostiene que los fenómenos mentales son causados por eventos físicos en el cerebro, pero que no tienen efectos sobre ningún evento físico, incluido el mismo cerebro. Entonces, la conducta es causada por músculos que se contraen al recibir impulsos nerviosos, y estos impulsos son generados por la entrada de otras neuronas o de órganos sensoriales. La conclusión a la que se llega con el epifenomenalismo es que los eventos mentales son ineficaces, como lo son las experiencias cualitativamente mentales. En todo caso, el epifenomenalismo es tomado hoy como una visión crítica contra el no-reduccionismo, puesto que cualquier fenómeno que no sea físicamente describible y anclado ontológicamente en las entidades físicas no tiene poder causal dentro del concepto de cierre causal del mundo. Si esto fuera correcto, entonces, los procesos emergentes no serían causalmente relevantes, sino habría que buscar su base física como explicación de su poder causal. En otras palabras, no es lo emergente lo que causa algo sino puramente el arreglo físico correspondiente.

Lo que motiva el epifenomenalismo es la suposición que todos los eventos físicos tienen las causas necesarias y suficientes, y estos pertenecen a un único nivel de eventos físicos. Seguidamente, si los eventos biológicos o mentales o sociales fueran algo diferente a un evento físico, entonces cualquier contribución propia sería una

violación a la leyes físicas y a al principio de cierre causal universal. Podemos ver, claramente, que se trata de una tesis en contra del pluralismo ontológico.

Es interesante mencionar que el problema del epifenomenalismo surge fundamentalmente en la teoría de la mente y no, por ejemplo, en las teorías de la biología, o aun en la química orgánica, puesto que, finalmente, si el mundo fuera explicado y dependiera ontológicamente de las micropartículas, prácticamente todo sería un producto epifenomenalista. La importancia se puede aclarar aún más. Si se demostrara que las propiedades mentales son causalmente impotentes en comparación con las físicas, ¿cuál sería la razón para mantener una ontología que es inerte o inactiva? Se podría aducir que estas propiedades son como los objetos abstractos que forman parte de nuestro discurso coloquial y científico. Pero realmente los objetos abstractos no son entidades (*relata*) que tienen un rol causal, no esperamos que lo tengan. Si mantenemos nuestras categorías mentales es porque también confiamos que ellas tienen que ver con los cambios causales efectivos en el mundo: creemos que nuestros deseos, como nuestra voluntad, sí tienen una causación eficiente en nuestros asuntos. Como dice P. Humphreys «si el argumento de la exclusión [“que el verdadero trabajo causal se debe a las propiedades físicas”] es sólido, entonces, el raciocinio, los *qualia*, las esperanzas y temores de la humanidad serían simplemente humo en el fuego de los procesos cerebrales» (1997, p. 2).

T. H. Huxley fue uno de los primeros en utilizar esta perspectiva en el contexto de lo mental, a mediados del siglo XIX, cuando comparó los eventos mentales con el silbato de la máquina de vapor, que no contribuye en nada al funcionamiento mecánico de la locomotora. W. James a fines de dicho siglo rechazó esta postura utilizando un argumento basado en la selección natural. Decía que si los placeres y los dolores no tuvieran efectos, no habría razón alguna para nuestra aversión a esas sensaciones, que son causados por actividades que consideramos importantes en nuestra vida. Entonces, no se explicaría por qué tenemos una valoración afectiva a las sensaciones y a la utilidad de la actividad que las producen (Robinson, 2007).

Hasta aquí podemos ver que el problema del epifenomenalismo no es otro que el del materialismo eliminativista. Así, los eliminativistas, como críticos de la posición emergentista de una realidad ontológica nueva, comparten las siguientes convicciones acerca de lo mental:

- (1) Los conceptos mentales son equivalentes a conceptos vacíos puesto que sus términos se refieren a algo que no existe.
- (2) Existen estados mentales, pero ellos son simplemente cerebrales.

El eliminativismo de acuerdo con (1) es un eco de la supresión de muchos conceptos, con sus respectivos términos, que resultaron hacer referencia a cosas inexistentes,

como por ejemplo, los demonios, las esferas de cristal y conceptos parecidos. De acuerdo con (2), hay un reconocimiento de los estados mentales —en oposición al conductismo lógico y metodológico original— pero requiere una reconceptualización teniendo como marco a las neurociencias. Es decir, no es que los estados mentales no existan, sino que están inadecuadamente descritos en el lenguaje de origen popular mentalista y deben ser traducidos o reemplazados por conceptos neurológicos en una futura ciencia de la conducta.

Al respecto, Smart, que fue mencionado anteriormente, sostiene lo siguiente:

Creo que la ciencia nos está proporcionando, cada vez más, un punto de vista por el cual los organismos pueden ser vistos como mecanismos físico-químicos: parece que incluso el comportamiento del hombre mismo será explicable, algún día, en términos mecanicistas. Parece ser, en lo que concierne a la ciencia, que no hay nada más en el mundo que arreglos cada vez más complejos de los constituyentes físicos (1991, p. 169).

Sin embargo, el temperamento de Smart no lo llevaba a postular la necesaria eliminación en el momento del vocabulario con contenido ontológico mental, sino como una propuesta a futuro. P. M. Churchland, sin embargo, sí es de la convicción que tanto más pronto nos deshagamos de toda la ontología mental, tanto mejor. Su eliminativismo materialista se expresa así:

[E]s la tesis que [afirma que] nuestra concepción común de los fenómenos psicológicos constituye una teoría radicalmente falsa, una teoría tan fundamentalmente defectuosa que tanto los principios como la ontología de tal teoría eventualmente serán desplazadas, en vez de suavemente reducidas, por una neurociencia completa. Nuestra comprensión común y aun nuestra introspección serán entonces reconstituidas dentro del marco conceptual de una neurociencia completa, una teoría que podemos esperar sea mucho más poderosa que la psicología de sentido común que desplaza, y sustancialmente más integrada dentro de la ciencia física en general (1991, p. 601).

Por tanto, el epifenomenalismo se presenta como una objeción a la emergencia también porque negaría que los eventos emergentes sean una realidad ontológica, como serían otras realidades que sí participan causalmente en el mundo. Si no se supone que las propiedades emergentes tienen efectos, entonces, bien podrían no ser reales sino inertes, o epifenoménicas. Y mucho de lo que se dice en el emergentismo podría tener la objeción de ser epifenoménico. Como dice Flanagan: «La sospecha epifenomenalista es extraordinariamente difícil de disipar» (1992, p. 133). El materialista, reduccionista o no, es, pues, contrario al emergentismo como lo señala adecuadamente Horgan:

Una posición materialista debería aseverar, contrariamente al emergentismo (i) que la física es causalmente completa (es decir, que las fuerzas causales fundamentales son fuerzas físicas y que las leyes de la física no pueden ser violadas); y (ii) que cualquier hecho metafísicamente básico o ley —cualquier explicador inexplicado, por decirlo así— es un hecho o ley dentro de la física misma (1993, p. 560).

Se puede ver, entonces, que la tensión es clara entre el emergentismo y el materialismo. El primero sostiene que los eventos mentales tienen poderes causales que no son explicables en los términos ni ontología de los poderes causales de sus sustratos físicos. Pero el materialismo postula que los poderes causales de lo mental son explicables en términos de las propiedades y leyes físicas subyacentes. Por eso para los materialistas es importante negar que las propiedades mentales sean propiedades emergentes y, posiblemente si lo fueran, serían microrreducibles.

1.5. Superveniencia

El concepto de superveniencia también es central en la discusión la causación mental y, en consecuencia, de cualquier teoría emergentista. Como muchos conceptos que son de uso casi exclusivo de la comunidad filosófica, el de superveniencia es uno que ha resurgido en la discusión académica reciente en el problema de lo mental, particularmente impulsado por J. Kim (1993)⁶.

La superveniencia no es propiamente una teoría sino un concepto aplicado a las relaciones entre dos conjuntos de propiedades, las propiedades supervenientes y las de su base:

Las propiedades mentales supervienen sobre las propiedades físicas, en que necesariamente, para cualquier propiedad mental *M*, si algo tiene *M* en un tiempo *t*, existe una propiedad física base (o subveniente) *P*, tal que tiene *P* en *t*, y necesariamente cualquier cosa que tiene *P* en un tiempo tiene *M* en ese tiempo (Kim, 2000, p. 9).

⁶ El término *superveniencia* viene del latín *supervenire* y Kim sostiene que lo encontró por primera vez en Leibniz. Posteriormente, su uso más frecuente fue encontrado en el libro de Hare, cuando introdujo dicho concepto en la filosofía moral. Según Hare, la característica de *bien* (en sentido moral) es superveniente si decimos, por ejemplo, que «San Francisco es un buen hombre», y que, a la vez, sea imposible que exista otro individuo que se encuentre en sus mismas circunstancias, que se comporte igual a él y que no sea un buen hombre. Lo que señala Moore —y es a partir de ahí que se crea el concepto contemporáneo que usa Kim— es que lo que se predica de las propiedades de un objeto está constreñido por las otras propiedades que tiene este. De este modo, si otro objeto tiene las segundas propiedades del primero, necesariamente tendría que tener las primeras. Para el desarrollo de este concepto ver Kim (1993, especialmente, pp. 131-160).

Supongamos que una persona desea tomar un helado de chocolate. Ese deseo debe ser una instancia de una propiedad física (posiblemente una propiedad neurológica) tal que cuando cualquiera esté en ella y tenga aquella propiedad física, tenga que experimentar el deseo de tomar un helado de chocolate. Entonces, cada propiedad mental tiene una base física que garantiza esa instanciación. Además, sin esa base no puede darse la propiedad mental.

Ahora bien, los filósofos han empleado este concepto para imaginar mundos posibles, como el de los gemelos de Putnam⁷, porque pone la siguiente restricción de realización física: «Cualesquiera dos cosas que son duplicados físicos exactos, necesariamente son también duplicados psicológicos, esto es, los duplicados físicos son duplicados *tout court*» (2000, p. 10).

Aparentemente uno podría considerar que la noción de superveniencia haría más viable la identidad mente-cerebro, o descartar la emergencia, ya que la base física es lo que determinaría los fenómenos mentales. Además, podría sospecharse una acusación epifenomenalista a los fenómenos mentales. Sin embargo, esta noción más bien es consistente con una tesis monista de la realidad, pero no es necesariamente reduccionista, puesto que la realización física de cualquier objeto complejo puede ser variada o múltiple. Lo único que restringe es la posibilidad que alguna cosa emergente o no tenga el mismo arreglo físico y no tenga las mismas propiedades emergentes o no. Pero la superveniencia es perfectamente compatible con el emergentismo, y aún más, creo que lo fortalece.

En primer lugar, la superveniencia significa que un nivel de los fenómenos depende de otro nivel (por ejemplo, el mental del biológico y este del químico), pero al mismo tiempo no es reducible al último; de ahí que se hable de dos conjuntos de propiedades y no de uno. Utilizando el ejemplo que dio origen al concepto de superveniencia en la moral, una cosa es ser «San Francisco» y otra es ser «bueno». Es importante diferenciar aquí lo que constituye una sutileza importante en el concepto de superveniencia. Esta puede ser fuerte o débil. Será fuerte si cualquier diferencia en el nivel superveniente, por ejemplo, tener un deseo diferente al de comer helado de chocolate, resulta en alguna diferencia en los sistemas subvenientes, por ejemplo, el arreglo específico del estado del cerebro o del sistema nervioso central. Pero comprometerse a una superveniencia fuerte sería asumir tácitamente que lo que proporciona la explicación real del fenómeno en cuestión sería el nivel subveniente y no el superveniente como la teoría de la emergencia reclamaría.

⁷ Ver Braun (2009, pp. 167-179) para una explicación y aplicación contemporánea del concepto y ejemplos de los gemelos de H. Putnam.

En segundo lugar, el concepto de superveniencia puede ir acorde con la teoría de la evolución, en tanto, se considere que la base física es necesaria para la aparición de fenómenos emergentes, pero la base física no es lo mismo que los fenómenos emergentes. En otras palabras, se puede utilizar algunos principios y leyes de la física para describir las condiciones para la aparición de fenómenos, que dado su arreglo particular, da lugar a propiedades novedosas. Sin embargo, no son las leyes de la física las que describirán completamente esos fenómenos. La sensación de ver el color rojo no es la reflexión en un lector óptico de una longitud de onda de 7700 *angstroms*. Eso sería lo que algunos reduccionistas considerarían una muy buena física, pero una ¡muy mala biología!

1.6. Causación descendente

Como hemos visto, la causación eficaz es un problema que acosa al emergentismo, por ello surge la necesidad de postular el concepto de causación descendente, que posiblemente es el más importante que define a la ontología y a los procesos emergentes. Según algunos sería la característica más importante del emergentismo de tipo fuerte (Clayton, 2004, p. 49).

El concepto fue introducido por D. Campbell en 1972 y publicado en español en 1974. Aun cuando la formación de este autor fue en psicología, su ensayo estaba más centrado en la biología. Y es que, como muchos, encuentra en la biología un paso intermedio en los procesos emergentes que alcanzan una mayor complejidad en la aparición de la conciencia y la conducta humanas.

El objetivo de Campbell era dar cuenta del comportamiento de sistemas biológicos organizados pero dispuestos jerárquicamente. Así, tendríamos en orden ascendente: la molécula, la célula, el tejido, el órgano, el organismo, la población en reproducción, la especie, el sistema social hasta el ecosistema. Pero sostiene que esas nos son realidades explicables desde el punto de sus componentes y menos concebidas como meros «agregados», sino más bien «como realidades verdaderas que como conveniencias arbitrarias de clasificación, organizando, cada uno de los órdenes superiores, las unidades reales del nivel inferior» (1974, p. 237). Es decir son realidades en todo el sentido de independencia ontológica. Para los reduccionistas —y curiosamente, Campbell se incluye entre ellos— las realidades jerárquicamente superiores se restringen por las leyes que gobiernan el comportamiento de las inferiores, incluidos los niveles de la física subatómica (p. 237). Este sería considerado un principio reduccionista de dos⁸.

⁸ El otro principio que no es relevante en la causación descendente es que los logros teleonómicos en niveles superiores requieren para su cumplimiento mecanismos y procesos específicos de nivel inferior. La explicación no es completa hasta que no se hayan especificado estos micromecanismos.

Pero para este autor no es suficiente, porque no se explicaría adecuadamente el fenómeno de la selección natural. Por ello, introduce dos principios adicionales:

- El principio emergentista:

La evolución biológica en su exploración sinuosa de segmentos del universo tropieza con leyes, que operan como sistemas selectivos, que no están descritas por las leyes de la física y de la química inorgánica, y que no serán descritas por los sustitutos futuros de las aproximaciones actuales de la física y de la química inorgánica; y,

- El principio de la causación descendente:

Cuando la selección natural actúa mediante la vida y la muerte en un nivel superior de organización, las leyes del sistema selectivo de nivel superior, determinan, en parte, la distribución de los hechos y sustancias de nivel inferior (p. 237).

La causación descendente también puede ser entendida (y es la formulación más conocida de Campbell), como una reformulación del primer principio: «Todos los procesos en los niveles inferiores de una jerarquía son restringidos por las leyes de los niveles superiores y actúan en conformidad con ellas» (pp. 237-238).

Entonces, para Campbell la causación ascendente es condición necesaria en la descripción y explicación del fenómeno biológico, pero no es suficiente, es solo parcial. En este sentido, además de una explicación en términos bioquímicos, se requeriría una explicación causal descendente para complementarla. Se pregunta, ¿cómo se podría explicar el hecho de que las mandíbulas de una termita o de una hormiga obrera puedan estar tan bien diseñadas como para hacer el trabajo de roer madera y de cargar semillas? Ciertamente, las leyes de la palanca son una parte de la explicación del organismo, pero se requiere del principio de «propósito organísmico» para explicar la distribución particular de las proteínas en la mandíbula, es decir, las plantillas de ADN que guían su producción (pp. 238-239).

La causación de otros elementos del organismo también debe ser incorporada, como la práctica social y los elementos del entorno que influyen en el diseño. Dice Campbell: «Las mandíbulas de la hormiga soldado y la distribución de las proteínas en ellas [...] requieren para su explicación ciertas leyes de la sociología centradas sobre la organización social en la que existe una división del trabajo» (p. 239).

La causación descendente, por tanto, no estaría interfiriendo con principios y regularidades biológicas ni tampoco con las leyes de la física. En otras palabras, dentro de una concepción fisicalista, la realidad de la vida trae nuevas formas de interacciones causales en el mundo, pero ellas son a su vez, consistentes con la física,

aunque no explicables ni predecibles por dicha ciencia. Por poner un ejemplo, las leyes de la física no pueden explicar y predecir la aparición de ratas, pero las ratas y su comportamiento no violan sus leyes.

Ahora bien, por más intuitivamente correcto que aparezca el concepto de causación descendente no ha sido inmune a una serie de críticas, particularmente cuando se utiliza en el análisis del fenómeno de la conciencia. Para los objetores, las actividades de la conciencia, y en general de lo que llamamos eventos mentales se explican necesaria y suficientemente por la causación ascendente: las influencias causales ocurren exclusivamente desde las partes constituyentes al todo constituido. Uno podría creer que el contenido de un pensamiento fue lo que causó una determinada acción; sin embargo, las fuerzas causales operativas son los eventos microfísicos, que en nuestros cerebros tienen la forma de eventos electroquímicos.

Con todo, se podría afirmar la causación descendente desde otra perspectiva, la dualista. Los dualistas afirmarían que una entidad o entidades que son ontológicamente distintas a las causas físicas son las responsables de la influencia causal en nuestros cuerpos. Pero esta posición, poco convincente en nuestros días por razones empíricas, también tiene problemas en el nivel conceptual. Por ejemplo, cuando se afirma que existen cosas ontológicamente distintas, existe una tremenda ambigüedad por el sentido de *cualitativamente distintas* o, en el caso histórico cartesiano, de naturalezas distintas. Pero esas distinciones podrían aplicarse al mundo biológico (sin incluir a los humanos conscientes) sin ninguna sorpresa. Por ejemplo, un conejo y un electrón parecen compartir poco en sentido ontológico; sin embargo, no tendríamos que inclinarnos al dualismo para explicar que los movimientos del conejo pueden afectar a los electrones que lo componen.

Para comportamientos de los seres vivos elementales como los paramecios es posible que una explicación desde el ámbito epistemológico y ontológico, basada en la bioquímica, parezca suficiente —aun así algunos no compartirían como suficiente una explicación bioquímica— pero cuando extrapolamos la bioquímica a la acción humana la brecha se hace inalcanzable, y las explicaciones y ontología que suponen una causación ascendente parecen insuficientes. Por ejemplo, en las explicaciones de los procesos psicológicos o históricos la causación ascendente nos conduciría a versiones que por momentos parecen cómicas. Supongamos que quisiéramos explicar las causas de la Segunda Guerra Mundial a partir de los eventos del fin de la Primera Guerra Mundial. ¿Qué vocabulario y qué entidades deberían reemplazar a las que usualmente utilizamos en la historiografía para dar cuenta de lo que algunos llamaron la «conspiración» de los veteranos de la Primera Guerra Mundial? ¿Es que la física fundamental —partículas subatómicas— conspiró (o algún fenómeno que asociamos con la conspiración) para causar la Segunda Guerra? ¿No parece más razonable

acudir a nuestras tradicionales concepciones sociológicas/históricas como entes responsables de la causación de estos conocidos eventos?

Por supuesto, el hecho de que parezca cómico o antiintuitivo no lo hace verdadero como sistema explicativo. Sin embargo, las explicaciones poderosamente antiintuitivas, particularmente las que tratan acerca de nosotros mismos, nuestras acciones, voliciones, intenciones, pueden ser sospechosas de no ser adecuadas, y, por ello, la concepción de causación descendente puede ser poderosamente explicativa en el análisis causal de, al menos, un grupo de fenómenos de la compleja realidad.

1.7. Tipos de emergencia

Para finalizar esta sección de análisis conceptual sobre el emergentismo veamos algunos tipos de teorías emergentistas, puesto que como toda teoría, la variedad de presupuestos y postulados determinará sus compromisos —ontológicos y epistemológicos— en el esfuerzo por construir una tesis aceptable de la conciencia. Además, como en las siguientes secciones veremos posturas emergentistas es importante visualizarlas dentro de la tipología que a continuación describiremos.

Siguiendo, en parte, la división de Chalmers (2006) y el modelo propuesto por Stephan (1999), se dividirá las teorías emergentistas en dos, fuerte y débil. Y de la débil, se subdividirá en emergentismo sincrónico y diacrónico.

1.7.1. Emergentismo fuerte

El emergentismo fuerte sostiene que un fenómeno de nivel jerárquicamente alto es fuertemente emergente respecto a un dominio de nivel inferior, si el fenómeno de mayor nivel surge del dominio inferior, pero las verdades que conciernen al fenómeno no pueden ser deducibles, ni en principio, de las verdades del nivel inferior. Se podría decir que la evolución del cosmos produce niveles nuevos, ontológicamente distintos, que se caracterizan por sus propias leyes o regularidades y sus fuerzas causales. Esta concepción es la más usual en la discusión filosófica y caracterizó a los llamados «emergentistas británicos» de principios del siglo XX.

El emergentismo fuerte tiene un compromiso ontológico más radical que el débil, puesto que considera que las características emergentes ontológicamente son características de los sistemas o de los todos que poseen capacidades causales que no son reducibles a ningunas de las mismas capacidades de las partes.

1.7.2. Emergentismo débil

El emergentismo débil sostiene que un fenómeno de nivel jerárquicamente alto es débilmente emergente en relación con un dominio de nivel inferior, si el fenómeno de mayor nivel surge del dominio inferior, pero las verdades concernientes al

fenómeno son inesperadas dados los principios que gobiernan el dominio del nivel inferior. Los emergentistas débiles afirman que, conforme aparecen nuevos patrones y características, los procesos causales pertenecen al ámbito de la física.

El emergentista débil está de acuerdo con el fuerte en que la evolución produce nuevas estructuras, que como un todo emergente, determina a sus partes. Pero, a diferencia del fuerte, admite que la imposibilidad de reconocer la estructura de los componentes que dan origen a la emergencia del nuevo todo, se debe a nuestra ignorancia. En otras palabras, más que una posición ontológica es una posición epistemológica. Las características novedosas son solo novedosas en el nivel epistémico.

Como se puede inferir, la emergencia fuerte tiene consecuencias mucho más radicales que la débil, puesto que, si se quiere entender los fenómenos de los niveles más altos, se requerirá el conocimiento de nuevas leyes y principios fundamentales que no se conocen en la física.

Puesta así la diferenciación, pareciera que la posición débil es insostenible y sorprendería porqué debiera llamarse emergentismo después de todo, ya que es una posición perfectamente compatible con el fisicalismo reduccionista. La cuestión es más de tiempo que de ontología. Conforme se vaya descubriendo más el mundo, estaremos en capacidad para asimilar los eventos de nivel superior en los del inferior, tarea que ya se había empezado en tantos ejemplos conocidos desde el siglo XIX y ha continuado hasta nuestros días, en el caso de la biología molecular, por mencionar el modelo reduccionista más impresionante.

Como se mencionó los siguientes tipos son un subgrupo del emergentismo débil.

1.7.2.1. Emergentismo sincrónico

El emergentismo sincrónico es una versión del emergentismo débil al que se le añade la irreducibilidad. Según Stephan:

La relación atemporal entre la propiedad del sistema y su microestructura, es decir, el arreglo y las propiedades de las partes del sistema son el centro de interés. Para tal teoría se considera que es emergente si es irreducible, esto es, si no es reducible al arreglo y las propiedades de las partes del sistema (1999, p. 49).

1.7.2.2. Emergentismo diacrónico

En contraste con el sincrónico, el diacrónico tiene las características adicionales de la novedad y la impredecibilidad. El énfasis de esta postura es la predictibilidad de los fenómenos. Pero dado que no es posible augurar los fenómenos, ya sea por la particular estructura emergente o porque la propiedad que superviene es impredecible, la natural consecuencia del diacrónico es la irreducibilidad.

2. EL ORIGEN DE LAS TEORÍAS EMERGENTISTAS

2.1. Antecedentes en la Antigüedad

Podemos encontrar las raíces del emergentismo en la filosofía griega antigua a partir del concepto de orden como resultado del caos y en el debate acerca de si lo que existe es producto del azar o la necesidad (Burnet, 2005, p. 7). La pregunta acerca del ser requiere la pregunta acerca de su origen. Pero este origen, aun cuando se relacione inicialmente con un caos primigenio, sin forma y con mutaciones aparentes, es reconceptualizado por algo permanente. Para los griegos antiguos, lo inestable, disímil y cambiante de la realidad es lo que nuestros sentidos nos informan. Sin embargo, como sostiene Guthrie (1997, pp. 23-28), la actitud filosófica generalizada entre los griegos antiguos era la creencia que lo que subyacía al aparente caos que revelaba nuestra sensorialidad, era una permanencia y unidad escondidas, discernible, si no ya por los sentidos, por la mente. Esta actitud parece conservarse siglos más tarde, como refiere Stebbing:

Parece existir una tendencia muy enraizada en la mente humana que busca [...] algo que persiste a través del cambio. Consecuentemente, el deseo por la explicación solo puede ser satisfecho con el descubrimiento que lo que aparece como nuevo y diferente estuvo ahí todo el tiempo. Por ello, la búsqueda por una identidad subyacente, una cosa que persiste, una sustancia que se conserva a pesar de los cambios cualitativos y en términos del cual estos cambios se explican (1950, p. 404).

Aristóteles, según algunos comentaristas, parece haber sostenido el principio que el todo tiene propiedades diferentes al de las partes constituyentes, sosteniendo inclusive que la naturaleza del todo era mayor que el de las partes⁹. Aunque no se podría afirmar con seguridad que el Estagirita suscribiría una posición asociada a lo que hoy llamamos emergentismo, es posible sostener que no podría ser un materialista reduccionista o eliminativista, puesto que en su análisis de la psicología sostiene que sí existe un alma y fenómenos psicológicos y una distinción irreducible entre el alma y el cuerpo. Aristóteles cree que el alma tiene un rol causal por sí misma y no por (o reducible a) las partes del cuerpo (2007, I, p. 4). Asimismo, considera que son los actos mentales, tales como el deseo y la fantasía, los que constituyen la causa eficiente de la acción y el habla; los estados corpóreos no tienen sino un rol meramente instrumental (III, p. 10). De manera que se podría sostener que Aristóteles concibe una forma de la llamada causación «descendente» en el sentido que no todos los eventos de un ser con alma son resultado de las constricciones corpóreas (que inclinaría más hacia una causación «ascendente»), sino que al menos algunos son resultado de uno

⁹ Aristóteles en el contexto de la respuesta a las paradojas de Xenón de Elea. Confróntese Goldstein (1999).

o más eventos psicológicos¹⁰. Y una suposición de causación descendente es compatible con formas contemporáneas de la fundamentación del emergentismo.

Galeno (1968), en el siglo II d. C., en varios de sus escritos indica claramente que existe una diferencia entre estudiar las partes del cuerpo y estudiar el todo, pues aparecen cualidades distinguibles cuando se aborda el estudio del segundo. Del texto *Sobre la utilidad de las partes del cuerpo* se puede inferir cómo concebía que las partes, como los dedos y los espacios que podían separarlos, cooperaban cada uno con el todo.

Ciertamente ni los presocráticos ni Aristóteles ni Galeno pudieron dar respuesta a la difícil pregunta acerca de la producción de todos más complejos a partir del azar o la necesidad; de hecho, esta pregunta sigue siendo explorada aún entre los filósofos de orientación evolucionista materialista. El cambio emergente podría ser un accidente inusual o azaroso que incrementa la complejidad de la estructura y de la función. O bien podría ser la expresión de la operación de las tendencias físicas y biológicas hacia la complejidad.

2.2. La tradición alemana: Hegel, Kant y el organicismo

Usualmente Hegel no es considerado como un autor antecedente de las ideas emergentistas contemporáneas; sin embargo, podemos encontrar que en su visión metafísica existe una progresión de la vida a partir de lo no-viviente, y de ahí a lo consciente y, luego, a lo espiritual. Esta concepción la encontramos en el concepto hegeliano de dialéctica que fue posteriormente desarrollado por Engels, al que me referiré a continuación. De acuerdo con este autor la pregunta básica de la filosofía es la relación del pensamiento y el ser. Y, conforme a las leyes de la dialéctica, el ser (*Sein*) determina la conciencia (*Bewußtsein*). La dialéctica explica cómo las cualidades del ser dan paso a nuevas formas y el sistema pasa a un nuevo estado emergiendo nuevas cualidades. Se puede afirmar, entonces, que el nuevo nivel no tiene las mismas propiedades que el anterior. En este sentido, la emergencia es un salto cualitativo a un sistema irreducible al anterior. Si empleamos el lenguaje hegeliano, según la dialéctica, la síntesis de las nuevas propiedades emerge de la tesis y la antítesis. Engels proporciona un esclarecedor ejemplo:

Tomemos, en cambio, una plástica planta ornamental, por ejemplo, una dalia o una orquídea; si tratamos, según el arte de la jardinería, la semilla y la planta que nace de ella, conseguimos como resultado de esta negación de la negación no ya solo más semillas, sino semillas cualitativamente mejoradas que producen flores más hermosas, y cada repetición de este proceso, cada nueva negación de la negación, aumenta dicho perfeccionamiento (2003, p. 126).

¹⁰ Para una explicación de esta tesis ver Caston (1997)

En el ejemplo, la planta, por el proceso de negación de la negación, dará lugar al cambio cualitativo.

En el periodo de la Ilustración encontramos un rechazo a las ideas mecanicistas en los escritos de Kant. Él concebía que en el desarrollo de un organismo simple, las partes interactuaban para producir progresivamente unas series complejas de emergencias de formas funcionales, en contraste con una máquina que era interpretada como un ensamblaje de partes premanufacturadas. Así, en la *Crítica del juicio* escribe:

Un ser organizado no es, pues, una simple máquina, no teniendo más que la fuerza *motriz*; posee en sí una virtud creadora y la comunica a las materias que no la tienen (organizándolas), y esta virtud creadora que se propaga no puede ser explicada por la sola fuerza motriz (por el mecanismo) (1876, p. 125).

En otro pasaje del mismo libro, en relación a la mecánica newtoniana afirma:

Es, en efecto, absolutamente cierto que no podemos aprender a conocer de una manera suficiente y, con mayor motivo, a explicar los seres organizados y su posibilidad interior por principios puramente mecánicos de la naturaleza; y se puede sostener sin temor, con igual certeza, que es absurdo para los hombres intentar semejante cosa y esperar que algún nuevo Newton vendrá un día a explicar la producción de un tallo de yerba por leyes naturales a las que no presida designio alguno, porque este es un procedimiento que se debe rehusar a los hombres en absoluto (p. 139).

Esta visión kantiana puede ser entendida en términos contemporáneos como que la complejidad de la vida no puede ser ni explanatoria ni ontológicamente reducible a los mecanismos causales naturales. Sin embargo, la visión de Kant no tuvo aplicación en el desarrollo de la evolución biológica y menos en el de la conciencia o mente humanas, pues sostenía que los fenómenos vivientes eran irreducibles, puesto que tenían propósito, en lugar de suponer que evolucionaban naturalmente a través de etapas discontinuas de complejidad emergente.

2.3. J. S. Mill y el inicio propiamente del emergentismo

La diferencia establecida por Kant entre el conjunto de componentes mecánicos y los sistemas complejos es abordada también desde el campo empirista por J. S. Mill. Para algunos este autor es, sin duda, el antecesor más evidente del denominado «emergentismo británico» (Nagel, 1961, p. 372). Según Mill, existen dos tipos de fenómenos, el primero es producto del modo de causación mecánica, o también llamado *composición de causas*. En suma, el efecto es producto de la sumatoria de las acciones individuales. Se puede citar como ejemplo la clásica composición de fuerzas newtoniana. El autor proporciona la siguiente ilustración:

Si un cuerpo es impulsado en dos direcciones por dos fuerzas, una que tiende a moverlo hacia el norte y la otra al este, es causado a moverse en un tiempo determinado con exactitud tan lejos en ambas direcciones como las dos fuerzas por separado lo hubieran conducido, y se queda precisamente donde hubiera llegado si hubiera actuado primero por una de las dos fuerzas y luego por la otra. Esta ley de la naturaleza se llama, en la dinámica, el principio de la composición de las fuerzas e, imitando esa expresión bien escogida, le voy a dar el nombre de la composición de causas al principio que se ejemplifica en todos los casos en que el efecto conjunto de varias causas es idéntico a la suma de sus efectos por separado (1843, p. 267).

En física clásica este efecto es conocido con el nombre de *sumatoria de vectores*: el efecto de dos o más fuerzas actuando en conjunto es la suma vectorial del efecto que cada fuerza hubiera tenido si hubiera actuado sola. Mill llama a las leyes que gobiernan este tipo de fenómenos *leyes homopáticas*.

En contraste, el segundo modo, el modo químico de acción conjunta de fuerzas, no se explica por la sumatoria de los efectos de cada una de las causas, fenómeno típico de las transacciones causales químicas. Las leyes que describen estos casos son llamadas *heteropáticas* porque, según Mill, las nuevas regularidades que aparecen en estos casos en que la combinación de diferentes sustancias producen otra que tiene nuevas propiedades, que no son poseídas por sus elementos constituyentes. La ilustración es la siguiente:

La combinación química de dos sustancias produce, como se sabe, una tercera sustancia, con propiedades diferentes de aquellas de cualquiera de las dos sustancias tomadas en forma separada, o ambas tomadas en forma conjunta. No se observa rastro alguno de las propiedades del hidrógeno o del oxígeno en aquellas de su compuesto, el agua. El sabor del azúcar de plomo no es la suma de los gustos de sus elementos componentes, el ácido acético [no es] su óxido; ni es el color del vitriolo azul una mezcla de los colores del ácido sulfúrico y el cobre. Esto explica por qué la mecánica es una ciencia deductiva o demostrativa, y no la química (p. 267).

Aunque se utilice la específica expresión *emergentismo*¹¹, Mill es el punto de referencia para el desarrollo moderno y contemporáneo de esta corriente. En efecto, en el ejemplo del agua mencionado, las propiedades del compuesto son cualitativamente distintas de los componentes y, por tanto, no se pueden deducir de las propiedades de los constituyentes, lo que nos lleva a suponer que Mill propone que los compuestos

¹¹ El término es atribuido a G. E. Lewes (1875) en su libro *Problems of Life and Mind*. Londres: Kegan, Paul, Trench, Trubner & Co. (ver Sarkar & Pfeifer, 2006).

manifiestan una novedad cualitativa en sus propiedades, irreducibles a la mera suma de las propiedades de sus bases. Él no lo afirma, pero podría interpretarse del ejemplo que en la emergencia surgen características no solo epistemológicas sino también ontológicas. Esto se infiere de la incapacidad de la derivabilidad a través de la adición¹².

Las aplicaciones de la diferencia entre leyes homopáticas y heteropáticas alcanzan no solo a las ciencias físicas sino también a las biológicas y las humanas. Así, Mill hablando de las sustancias vegetales y animales, argumenta que sus cuerpos continúan siguiendo las leyes mecánicas y químicas *en tanto* no se apliquen las nuevas leyes que los rigen como seres organizados (p. 431). En otras palabras, aparece por vez primera en la historia del emergentismo la noción que los seres vivos tienen propiedades ontológicas describibles en términos mecánicos y otras en términos de organización (o composición siguiendo su vocabulario).

Más aún, en la terminología especializada de la filosofía de la mente contemporánea se utiliza constantemente (a favor y en contra) la noción de *qualia*, término que alude a las propiedades que experimentamos los seres mentales como son sensaciones, sentimientos, percepciones, la manera cualitativamente particular de sentir lo que vemos, oímos, olemos, etcétera (Guttenplan, 1995, p. 514). Uno de los clásicos argumentos en contra de la reducibilidad de lo mental vincula la realidad de los *qualia* a una ontología distinta a la física y, por supuesto, desde el punto de vista epistemológico, indescriptible a partir del lenguaje de la física (Jackson 1986). En consecuencia, Mill es un evidente antecesor de la irreducibilidad de los *qualia* a partir de la división de leyes antes descrita. En la siguiente cita se puede leer su argumentación en contra de la reducción:

Existe un fenómeno *sui generis* llamado color, que nuestra conciencia testifica no como un grado particular de algún otro fenómeno, como el calor, el olor o el movimiento, sino como algo *intrínsecamente* diferente a cualquier otro; se sigue que existen leyes últimas del color, que aun cuando los hechos del color admitan una explicación, aquellos no pueden ser explicados a partir solamente de las leyes del calor o del olor, o solo del movimiento, sino que, por más lejos que se lleve la explicación, siempre quedará en él la ley del color. Con esto no quiero decir que no pueda ser demostrado que algún otro fenómeno, alguna acción química

¹² Se podría argumentar que el efecto heteropático es resultado de una limitación epistemológica cuando Mill señala lo siguiente: «[En la química] la mayoría de las uniformidades a las que pertenecen las causas cuando están separadas cesan totalmente cuando están juntas, y no somos capaces, al menos en el estado presente del conocimiento, de predecir el resultado que sigue a cualquiera de las nuevas combinaciones hasta que hayamos hecho el experimento específico» (1843, p. 267). Sin embargo, la incapacidad de predicción no alteraría la suposición de la ontología distinta del compuesto con respecto a sus componentes.

o mecánica, por ejemplo, lo precede invariablemente, y es la causa de todo fenómeno del color. Aun si se probara esto, sería una extensión considerable de nuestro conocimiento de la naturaleza, pero no explicaría cómo o por qué un movimiento, o una acción química, puede producir la sensación del color (1843, p. 346; las cursivas son mías).

La propuesta de Mill para abordar el fenómeno de la sensación de ver el color es análoga a la distinción entre leyes homopáticas y heteropáticas, a saber, leyes derivadas y últimas. Las derivativas son aquellas que pueden ser deducibles de otras más generales y las últimas no. Las leyes que describen los fenómenos como el de la visión serían consideradas últimas y, en tal sentido, no podrían derivarse de las leyes de la mecánica —leyes derivativas— que describirían los diferentes fenómenos corporales asociados a la visión. En la sección referente a los límites de la explicación de las leyes de la naturaleza Mill establece que fenómenos como el de la visión deben quedar restringidos a leyes últimas, y en consecuencia, estos no podrían ser simplificados en leyes derivativas, como aconseja para otros casos (pp. 346 y ss.). En el fondo, hay una propuesta para distinguir propiedades derivables y de las últimas, que no son dependientes de otros factores, en consecuencia, irreducibles, en nuestro vocabulario contemporáneo.

El proyecto de Mill es diferente al que los posteriores emergentistas británicos, que veremos más adelante, le atribuyen. En efecto, la normativa defendida en *Sistema de Lógica* era la metodología basada en la experiencia y la inducción, con un componente claramente determinista, acorde con la actitud prevalente en el siglo XIX. El autor estaba defendiendo lo que él consideraba eran las fuentes del conocimiento científico: el método experimental y el método inductivo. Por eso vemos constantemente la búsqueda de regularidades como parte del sistema explicativo, bien sea en leyes o en propiedades que pudieran ser generalizadas y, por supuesto, manipuladas para la predicción. Años más tarde los emergentistas británicos utilizaron las ideas de Mill en una forma que no estuvo pensada por el autor original: para buscar una tercera vía entre el mecanicismo y el vitalismo y para explicar la aparición de nuevas entidades y propiedades como producto de la evolución (Stephan, 1992).

2.4. Los emergentistas británicos

Los conceptos de efectos y leyes heteropáticos tuvieron impacto sobre algunos teóricos británicos —filósofos y biólogos— que transformaron las ideas de Mill en la doctrina de la *evolución emergentista*. Los más destacados fueron Scott Alexander, Lloyd Morgan, y C. D. Broad.

2.4.1. S. Alexander

En *Espacio, tiempo y deidad* (1920), Alexander introduce el concepto de emergencia tomando ideas de Mill, entre ellas las características de novedad e impredecibilidad de las propiedades emergentes. Es importante mencionar que tanto este autor como los otros emergentistas británicos no descartaban la idea que la existencia de nuevas propiedades emergentes era resultado causal de las propiedades del nivel inferior, es decir, si se dieran los mismos componentes en un arreglo idéntico, el todo resultante en el nuevo nivel superior debería ser consecuentemente idéntico (1920, p. 330). Por ejemplo, Broad afirma esta idea así:

Esta posibilidad [emergente] es perfectamente compatible con la visión que el comportamiento característico de un cuerpo viviente es completamente determinado por la naturaleza y arreglo de los componentes químicos que lo componen, en el sentido que cualquier todo que está compuesto por tales componentes en tal arreglo mostrará un comportamiento vital y que nada más lo podrá hacer de esa manera (1925, pp. 67-68)¹³.

Esta idea resulta familiar en nuestros días pero bajo el nombre técnico de superveniencia, en el que, como ya se ha mencionado, las propiedades emergentes supervienen sobre las propiedades de sus constituyentes (Kim, 1984).

Dentro de un proyecto del desarrollo de un sistema metafísico bastante ambicioso, Alexander desarrolla específicamente el concepto de emergencia en el que diferentes realidades aparecen como resultado de niveles inferiores del espacio-tiempo (1920). Presta especial atención al fenómeno de la conciencia y, de alguna manera anticipándose a los críticos de la posterior teoría de la identidad, afirma que «la experiencia nos lleva a conectar los procesos mentales con nuestro cuerpo, y en particular con nuestro sistema nervioso central, y a localizar nuestros procesos mentales en los mismos lugares y tiempos como ciertos procesos neurales» (II, p. 5). Sin embargo, pronto advierte que gracias a la experiencia, suponemos aquí, experimentación, y a la reflexión nos damos cuenta de que esos procesos mentales constituyen un proceso altamente diferenciado y complejo de nuestro cuerpo humano. Por eso sugiere no identificar los procesos mentales con los neurales, puesto que no solamente tienen mayor complejidad sino que manifiestan conciencia. La conciencia, según Alexander no puede ser identificada totalmente con lo neural: «Puesto que mientras que el proceso mental es *algo* neural, no es *meramente* neural. Pues, para que la mente emerja se requiere una constelación de condiciones neurales o vitales que no se encuentran en las acciones vitales que no son mentales» (II, p. 6).

¹³ Es, en efecto, una idea compatible con la teoría de la identidad más rígida que no admite la corrección posterior a la que la teoría fue sometida, es decir, el funcionalismo.

La clara separación ontológica entre las propiedades de niveles inferiores y superiores queda establecida por este autor así:

La cualidad superior emerge del nivel inferior de existencia y tiene sus raíces allí, pero emerge desde allí, y no pertenece a ese nivel inferior, sino que constituye a su poseedor un nuevo orden de existencia con sus especiales leyes del comportamiento. La existencia de las cualidades emergentes así descrita es algo que debe ser notado, como algunos dirían, bajo la compulsión del hecho empírico bruto o, como yo prefiero llamarlo, en términos menos duros, que debe ser aceptada con una «piedad natural» del investigador. No admite una explicación (II, p. 46).

Añade más adelante:

La cualidad es por lo tanto, un hecho empírico [...] que un cierto tipo de arreglo de los existentes de un nivel inferior debe ser calificado con una nueva cualidad [...] La cualidad pertenece a las cosas como la mente o la conciencia pertenece a los procesos vitales de una cierta configuración (p. 47).

El problema que enfrenta Alexander es la reconciliación entre su determinismo de tipo laplaciano¹⁴ y la emergencia de la nueva cualidad que «es expresable *sin residuo* en términos de los procesos propios al nivel del que emerge» (II, p. 45; las cursivas son mías). En este sentido, si las cualidades emergentes pudieran ser expresadas sin residuo, no hay nada incompatible con que sean expresadas en términos de la física y el demonio laplaciano podría hacer perfectas predicciones de ellas. Y esta es una forma de reduccionismo, lo haya querido Alexander o no.

2.4.2. L. Morgan

Morgan resulta un interesante punto de conexión entre Alexander y Broad, y por ello conviene presentarlo. Para este autor, como lo fue para Mill, la química presenta innumerables casos de emergencia de propiedades emergentes:

Cuando el carbono que tiene ciertas propiedades se combina con el azufre que tiene otras propiedades se forma no una mera mezcla pero un nuevo compuesto, algunas de cuyas propiedades son bastante diferentes de cualquiera de los componentes [...] Pero las otras distintas propiedades son constitutivamente emergentes, que no podrían ser predichas con anterioridad a tal combinación. Por supuesto una vez que uno aprende lo que emerge en esta instancia particular, uno puede predecir lo que emergerá en una instancia semejante bajo circunstancias similares (1923, p. 3).

¹⁴ «Un calculador a quien se le da el estado del universo en un cierto número de instantes, o en un instante., con la ley del cambio podría, dados los poderes suficientes, calcular cuál sería la condición del mundo en cualquier instante posterior» (1920, II, p. 73).

Morgan suscribe igualmente la concepción de Mill acerca de la novedad e impredecibilidad:

Lo que se afirma que no puede ser predicho, entonces, es la expresión emergente de una nueva clase de relación entre eventos preexistentes. Uno no podría predecir el carácter emergente de eventos vitales a partir del solo posible conocimiento completo de los eventos físico-químicos, si la vida es un acorde emergente y no meramente una sumatoria, por compleja que fuera (pp. 5-6).

El concepto de evolución emergente de este autor es resultado de una combinación de la emergencia milliana junto con una evolución darwiniana en el nivel cósmico. Por ello rechaza una cosmología mecanicista y aboga por una emergencia «en todos sus grados ascendentes» (p. 8).

Finalmente, podríamos anotar que la propuesta de Morgan se enmarca en la denominada causación descendiente y, por tanto, los niveles ascendentes de complejidad mencionados deberían ser estudiados por las ciencias especiales y no por la física.

2.4.3. C. D. Broad

Culmina el grupo de emergentistas británicos C. D. Broad, considerado por muchos filósofos como la cumbre de este grupo. Su libro *La mente y su lugar en la naturaleza* es un esfuerzo por «demostrar la relación (o no) entre las diferentes ciencias» (2001, p. 3). La pregunta fundamental en la controversia entre mecanicistas y vitalistas acerca de los organismos vivientes es la siguiente: «¿Son los aparentes tipos de objetos materiales irreductiblemente diferentes?» (p. 43). De acuerdo con la propuesta de Mill y tratando de dar respuesta a la pregunta, Broad sostiene que en una teoría puramente mecánica todas las diferentes cosas estarían hechas básicamente de lo mismo y diferirían entre sí por su número, arreglo y movimientos de sus partículas. Sus diferentes comportamientos serían más bien aparentes, pero, en último término, iguales. Sin embargo, esta visión unificadora de la realidad es ingenua y superficial (p. 47). La visión mecanicista para este autor desconoce la diferencia entre propiedades macroscópicas y microscópicas, que son cualitativamente distintas y no podría ser resultante la primera de la sumatoria de las segundas. Señala:

Es un hecho básico que el mundo externo, tal como es percibido por nosotros, no parece tener la homogeneidad demandada por el mecanicismo puro. Si *en realidad* tiene las diferentes cualidades sensibles que *parece* tener, el mecanicismo puro no puede ser verdad acerca del todo del mundo externo y no puede ser la total verdad acerca de cualquier parte de él (p. 51).

Broad reconoce que, aun si se quisiera explicar las características macroscópicas a partir de las microscópicas, existe una limitación en la ciencia de su tiempo para hacer tales derivaciones; sin embargo, argumenta que no hay razón *a priori* por la que los eventos microscópicos debieran conformarse con las demandas del mecanicismo puro ni las interacciones de dichos eventos entre sí. La apariencia de mecanicismo no es razón suficiente para concluir que los eventos sean ocurrencias mecanicistas.

La tesis mecanicista que el autor combate la define como aquella teoría en que un todo podría ser deducido del conocimiento suficiente acerca del comportamiento de sus componentes. El ejemplo que consigna es el de las máquinas artificiales. Un reloj para la ilustración:

Nadie supone que el comportamiento peculiar de los relojes es una cualidad emergente de ese tipo de estructura y que no puede ser aprendido estudiando nada pero relojes. Sabemos perfectamente bien que el comportamiento de un reloj puede ser deducido del arreglo particular de sus resortes, ruedas, péndulo, etc., contenidos en él, y a partir de las leyes generales de la mecánica y la física que se aplican también a los sistemas materiales que no son relojes (p. 60).

Para predecir el comportamiento de un reloj no se requeriría haber experimentado previamente con un reloj o, para el caso, haber visto uno funcionando. Bastaría con lo aprendido del estudio de otros sistemas materiales y las leyes que los rigen, así como de las propiedades de los resortes y rodajes para hacer predicciones sobre los relojes. En otras palabras, de los componentes se puede deducir el comportamiento del todo.

En contraste, la teoría emergente de Broad asevera que un todo que está compuesto por los constituyentes A, B y C en una relación R entre sí tiene propiedades características, de tal modo que cualquier todo que tuviera A, B y C como componentes en una relación R tendría las mismas características que el primero. Además, las características del todo no pueden ser deducidas de las propiedades de A, B y C tomadas aisladamente, sino solo cuando se encuentran en R. Seguidamente, tampoco pueden ser deducidas las características de un todo cuyos componentes fueran A, B y C *pero* en una relación que no fuera R (p. 61).

Broad propone el ejemplo del cloruro de plata (AgCl_2). Lo que sabemos de este compuesto es resultado del aprendizaje cuando esta sustancia se encuentra en el estado de cloruro y no de las propiedades del cloro y de la plata respectivamente. Es cierto que se puede predicar que la propiedad de la plata es la posibilidad de combinarse establemente con el cloro para dar lugar a las propiedades del cloruro de plata. Sin embargo, la propiedad del compuesto es conocida una vez que es conocido el compuesto. De ninguna manera analizando por separado hubiéramos conocido las propiedades del compuesto (como sí, por contraste, las partes del reloj).

En el caso del comportamiento de los vivientes, el autor señala que se sabe que los componentes vitales están compuestos a su vez de componentes químicos en un arreglo particular. En este sentido, se puede hablar de un cuerpo viviente como un compuesto de segundo orden, es decir, *compuesto de compuestos*, como en el ejemplo del cloruro de plata. Nuevamente la limitación epistemológica tiene lugar puesto que en tanto no se conozca el compuesto no se podrán conocer las propiedades a partir de sus componentes (p. 67). Pero añade que las características del comportamiento del cuerpo viviente están completamente determinadas por la naturaleza y el arreglo de los componentes, en el sentido que «misma composición=mismo compuesto», posición que nuevamente nos recuerda la tesis contemporánea de la superveniencia antecedida por Alexander.

El comportamiento del componente de segundo orden tiene leyes propias, llamadas «leyes transordinales» (p. 78) que Broad considera «únicas», «finales» e «irreducibles». Estas leyes son emergentes y no son deducibles de leyes de niveles inferiores. De acuerdo con esta visión, son similares en características a las defendidas actualmente por aquellos que promueven la explicación hempeliana vía leyes, pero, al mismo tiempo, respetando la autonomía de las denominadas «ciencias especiales». Estos defensores titulan estas regularidades como «leyes blandas» o «leyes *ceteris paribus*» (Horgan & Tienson, 1990; Fodor, 1991).

Hacia el final de la obra de Broad se aborda el problema mente-cuerpo. Los eventos mentales, en su interpretación, representan otro aspecto de los niveles emergentes. En el contexto de su discusión acerca de la unidad de la mente, llega a la conclusión que se requiere una teoría central de la mente. La razón que sostiene esta resolución es que los eventos mentales no podrían ser explicados en términos solo de sus relaciones. Se requiere un «centro» mental que unifique los eventos mentales como una «mente» (p. 584). Esto lo lleva a postular, finalmente, que existe una «sustancia mental», que por momentos quiere asociar al concepto de un «ego puro». Si bien reconoce que existen razones para estar a favor y en contra del ego puro, la evidencia empírica será la que incline la preferencia para postular esta entidad. E interpreto *entidad* puesto que afirma que el ego puro es «un existente particular que es de un tipo diferente de cualquier evento, posee diferentes eventos, pero no es en sí mismo un evento» (p. 584).

Es en esta noción de ego puro que Broad parece distanciarse de la misma teoría evolucionista que afirma en la primera parte de su obra. En efecto, si bien es cierto que no asevera que hay una sustancia inmaterial o mental, como lo dirían los dualistas, su posición no se distancia tanto de los dualistas al afirmar que la mente estaría compuesta por partículas mentales (p. 600) y no queda claro cuál sería la naturaleza de esas partículas. Por ejemplo, ¿podría predicarse de ellas lo mismo que se predicaba

de los compuestos químicos, su ejemplo predilecto? El autor no aclaró en su obra esa y otras preguntas acerca del concepto de sustancia mental.

Sin embargo, se mantuvo como un emergentista en tanto que para él, la mente es un todo duradero («*enduring whole*») y finalmente sostiene que, de todas las posiciones metafísicas disponibles en el arsenal filosófico (¡discute diecisiete!), se inclina por el emergentismo materialista sobre las otras. Aunque habría que advertir que el materialismo que defiende Broad dista mucho del que hoy en día conceptualizaríamos en forma estándar. De hecho, una afirmación como la que hace al final de su obra «el idealismo no es incompatible con el materialismo» (p. 654) no pareciera ser consistente con nuestras convicciones de lo que entraña actualmente el materialismo metafísico.

2.4.4. El final de los emergentistas británicos

Con *La mente y su lugar en la naturaleza* se termina la tradición de los emergentistas británicos y son varias las razones que se han señalado para explicar su declive como teoría explicativa de la aparición y realidad de lo mental. Presentaré las dos más importantes.

Para McLaughlin (1992), el derrumbe del emergentismo británico se debió no a razones filosóficas sino empíricas y, particularmente, al advenimiento de la mecánica cuántica, que revolucionó nuestra manera de concebir el comportamiento de la materia y, en algunos casos, en contra de las suposiciones y principios de la mecánica newtoniana. Este autor cita específicamente la explicación reduccionista del enlace químico.

Los procesos químicos aludidos por los tres emergentistas británicos eran interpretados correctamente en el sentido que el producto de los componentes no podía ser la suma de lo que sería si el efecto de cada uno de los reactivos actuara independientemente. Los procesos químicos aludidos producían efectos emergentes. Pero, señala McLaughlin, esos efectos no son incompatibles con el materialismo reduccionista. La reducción de la química por la mecánica cuántica se ha tomado como un paradigma del materialismo reduccionista. Es claro que los emergentistas trabajaron bajo el modelo newtoniano, pero la mecánica cuántica amplió la concepción del mecanismo introduciendo una noción holística del mecanismo y, por tanto, según este autor, de la explicación reduccionista:

La mecánica cuántica explica reductivamente la química, pero sin apelar a principios composicionales aditivos o aun lineales y sin la postulación de nuevas fuerzas irreducibles de mayor nivel. Aún más, la mecánica cuántica ha conducido al desarrollo de la biología molecular, y los sucesos de esta disciplina (por ejemplo, el descubrimiento de la estructura del ADN) han erradicado virtualmente cualquier

forma de vitalismo de la biología. Con la evidencia actual, parece ser que todas las fuerzas fundamentales se ejercen debajo del nivel del átomo (1997, p. 10).

La posición de McLaughlin es correcta: de acuerdo con los principios de la mecánica cuántica no constituye una evidencia a favor del emergentismo la imposibilidad de reducción a partir de los principios composicionales lineales. En efecto, es posible la reducción sin asumir la división por partes. De manera que tendría que buscarse otros principios para defender el emergentismo, al menos desde la perspectiva de Broad y los otros. De otro lado, con la explicación cuántica de los enlaces químicos tampoco se podía postular más la irreducibilidad de las fuerzas químicas y, en consecuencia, los emergentistas británicos perdían algunos de sus celeberrimos ejemplos, más aún, su fundamento metafísico. Sin embargo, aun cuando la crítica de McLaughlin sea correcta, no se sigue necesariamente que mientras se pueda aceptar la reducción epistémica y óptica de la mecánica clásica a la cuántica lo mismo tendría que aplicarse para el caso de la psicología y las teorías neurocientíficas.

Un segundo problema para los emergentistas fue el apogeo del positivismo lógico temporalmente paralelo al desarrollo de los británicos. Dado que algunas de las explicaciones de los emergentistas aludían a principios metafísicos, incluidos los de la divinidad, como fue en el caso de Morgan, o justificaciones inadecuadas como la «piedad del investigador» de Alexander, los emergentistas fueron presa fácil de los críticos ultraempiristas del positivismo.

Quizás, a manera de conclusión de este periodo es que la significación del emergentismo británico fue más en el aspecto epistemológico como guía en la comprensión naturalista de lo mental. Que encontraran la relación (y hasta por momentos identificación) entre los procesos mentales y los cerebrales no es sorprendente. Después de todo, su emergentismo era evolucionista y, en consecuencia, tenía que ser consistente con otros aspectos de la realidad. En la actualidad, la actitud no-reduccionista de lo mental debe mucho a lo postulado por los emergentistas, y no porque lo mental se presente como un conjunto de propiedades misteriosas (como los *qualia*), sino sobre todo porque el descubrir los niveles de emergencia nos previene, por un lado, de suponer vacíos de continuidad en el orden natural, y por el otro, dar sentido a la causación descendente que explica coherentemente nuestra vida mental.

Una teoría emergentista para que sea útil en la normativa científica de la psicología o la biología tendría que aceptar alguna forma de principio de parsimonia y no multiplicar entes sin necesidad. De hecho, será más recomendable explicar la causación mental apelando a propiedades mentales que surgen de los componentes del sistema nervioso central que introducir, como lo hizo hacia el final Broad, «cosas» o «partículas» mentales.

2.5. Resurgimiento del emergentismo en la década de 1960

Pasaron décadas en las que el emergentismo estuvo olvidado por el efecto de las críticas. Sin embargo, su auge en nuestros días se debe al trabajo de dos autores que quiero comentar brevemente: Michael Polanyi y Roger Sperry.

El reduccionismo radical fue la respuesta al emergentismo británico y, con la teoría de la identidad mente-cerebro en vigencia en los años sesenta, aparece la figura solitaria de Polanyi, quien abogó ardientemente por la irreducibilidad de la noción de persona asociada a la defensa del emergentismo.

2.5.1. M. Polanyi

Para Polanyi existen diferentes niveles de la realidad, comenzando con el físico, el químico, el biológico y, finalmente, el consciente. Las cosas vivientes, nivel intermedio, tienen causación descendente sobre los mecanismos y partes biológicos. Y, naturalmente, los seres humanos manifiestan causación descendente cuando, por ejemplo, manifiestan conducta teleológica e intencional o escriben un texto. El resumen de la tesis de este autor la encontramos en el siguiente pasaje:

La primera emergencia, por la que la vida comienza a existir, es el prototipo de todos los estadios subsecuentes de la evolución, por la que elevadas formas de vida, con sus más elevados principios, emergen en existencia [...] El espectáculo de estadios elevados de emergencia confirma esta generalización mostrando en el nivel más alto de emergencia evolucionaria aquellos poderes mentales en los que por primera vez reconocimos nuestra facultad de conocimiento tácito (citado por Clayton, 2004, p. 18).

Toda organización es para Polanyi, emergente. Y el máximo de la organización se da en el nivel de la conciencia. De hecho, una reacción química es una forma de organización, pero la considera «pasiva», como cuando ocurre una reacción en un tubo de ensayo. En cambio, se puede contrastar la observación de una jugada de ajedrez, organización activa en la que podemos evidenciar la estrategia del maestro del ajedrez, en la razón por la que hace determinada jugada y lo que espera conseguir con esa jugada. Esta diferenciación pasiva/activa es central en la postulación y evidencia de la causación descendente que se da en el nivel cerebral, según Polanyi.

La teoría del conocimiento tácito merece una aclaración porque es uno de los sustentos de este autor en su defensa del emergentismo. El conocimiento tácito es una facultad del pensamiento que está motivada por la anticipación del descubrimiento. A cada instante estamos guiados por la sensación de la presencia de una realidad oculta a la que apuntan nuestras pistas. Este conocimiento requiere

dos niveles de la realidad, el de los particulares y el de la «comprensión del significado» (2009, p. 34).

La distinción entre el conocimiento de los particulares y la comprensión del significado en Polanyi es un eco de la tradición alemana del organicismo y, particularmente, de la concepción de la Gestalt. Los psicólogos asociados a la Gestalt afirmaban que los patrones perceptuales que organizan el campo visual se pueden explicar desde un punto de vista de la actividad eléctrica de la corteza cerebral, pero la formación de esa actividad y los patrones eléctricos no se explicaría con la mecánica de las neuronas individuales, sino en la organización del todo (Sarkar & Pfeifer, 2006, p. 233). De forma análoga, Polanyi sostiene que nuestro conocimiento de otras mentes («persona» es el término preferido) como el de un todo organizado es distinto al conocimiento de sus componentes:

El parentesco estructural entre el conocimiento de una persona y el descubrimiento de un problema, y la alineación de ambos con el conocimiento de un ladrillo, llama la atención acerca de la mayor profundidad de la persona y el problema, si se compara con la menor profundidad de un ladrillo. Se siente que las personas y los problemas son más profundos, porque esperamos aún que nos revelen de formas inesperadas en el futuro, mientras que los ladrillos no nos evocan tales expectativas. Esta capacidad de una cosa para revelarse en unas formas inesperadas en el futuro, lo atribuyo al hecho de que la cosa observada es un aspecto de la realidad que posee una significación que no se agota con nuestra concepción de un solo aspecto de la misma (2009, p. 32).

De esta cita se podría inferir que la posición de Polanyi se asemeja a actitudes pragmatistas o instrumentalistas acerca de la descripción de la psicología humana tomando diversas perspectivas, pero él insiste que la realidad organizada adquiere estatus de entidad: «[...] lo que conocemos por el entendimiento de una entidad que comprende hace una referencia *ontológica* a ella [...] la entidades que comprenden ejemplificadas por la actuación humana hábil son cosas reales, tan reales como ladrillos [...] aún más reales que ladrillos» (p. 33).

Es reconocido que varias de las suposiciones de las que parte Polanyi son contrarias a las aceptadas en la actualidad¹⁵; sin embargo, tiene el mérito de haber reiniciado el debate acerca del concepto del emergentismo contemporáneo, lo mismo que el neurólogo Sperry.

¹⁵ Para un examen de las críticas ver Clayton (2006).

2.5.2. R. Sperry

Sperry podría ser considerado el último de los precursores del emergentismo contemporáneo. Interesado en el problema de los fenómenos de la escisión cerebral y la integración de los hemisferios a través del cuerpo calloso, también desarrolló una teoría emergentista denominada *emergentismo interaccionista*.

Para él, la conciencia no tenía un rol epifenoménico, sino, en efecto, causaba la toma de decisiones en el funcionamiento cerebral. Su postura es a la vez crítica del materialismo reduccionista y del dualismo sustancial. Por ello prefiere ser denominado un interaccionista, pero monista:

Esta es la idea, en suma, que los fenómenos conscientes como propiedades funcionales emergentes del procesamiento del cerebro ejercen un rol de control activo como determinantes causales para darles forma a los patrones de flujo de la excitación cerebral. Una vez que se generan de los eventos neurales, los patrones mentales del más alto orden y los programas tienen sus propias cualidades subjetivas y progreso, operan e interactúan por sus propias leyes causales y principios que son diferentes y no pueden ser reducidas a aquellas de la neurofisiología [...] Comparados con los procesos fisiológicos, los eventos conscientes son más molares, siendo determinados por las interrelaciones configuracionales u organizacionales en las funciones neuronales [...] Las fuerzas mentales no violan, perturban o intervienen en la actividad neuronal pero la supervienen. La interacción es mutuamente recíproca entre los niveles neurales y los mentales anidados en las jerarquías del cerebro (1980, p. 201).

Es curioso que un neurólogo del siglo XX utilice el término interaccionismo que tiene, sin dudas, un sabor cartesiano. Sin embargo, el uso es debido a los trabajos de Sperry con la escisión cerebral. Al haber removido el cuerpo calloso de sus pacientes, no se podía dar ninguna descripción neuronal de la conciencia unificada que todavía conservaban. Por tanto, tendría que haber una interacción en un nivel emergente de la conciencia, desde la cual los estados conscientes ejerzan una influencia causal directa en los estados cerebrales, entre otros posibles factores causales. Esta estrategia explicativa podría ser fácilmente criticada porque bien podría interpretarse la causación descendente desde el comportamiento causal eficiente de los componentes neurales.

Si bien es cierto que el grado de sofisticación filosófica de Sperry es cuestionable, hacia el final de este periodo mencionado nos ofrece una motivación para interpretar los datos neurofisiológicos dentro de una posible fenomenología de la conciencia a partir de, nuevamente, la noción de todo-parte y una teoría de causación descendente que representa una estrategia adecuada para el desarrollo de una teoría emergente contemporánea.

3. LAS TEORÍAS EMERGENTISTAS DE LA CONCIENCIA

3.1. De la biología a la conciencia

No es solo una estrategia explicativa y —espero— persuasiva el encontrar una transición, un *continuum* desde la biología hacia la conciencia humana, sino que, desde una perspectiva evolucionista, no tiene sentido hablar de la conciencia sin tomar en cuenta la historia evolutiva que produjo los cerebros. Para el dualista no es necesario siquiera enfrentar la pregunta de este tránsito, porque casi por milagro aparecen propiedades que no tienen nada en común con el resto del orden natural. Aunque parezca insólito, podría decirse algo análogo de los fisicalistas reduccionistas, puesto que el hecho que haya o no cerebros biológicos es un detalle contingente, ya que al final todo puede subsumirse en leyes microfísicas, incluyendo las funciones cerebrales. Sin embargo, en una propuesta emergentista se tiene que estudiar la naturaleza de otras estructuras emergentes como parte de la historia evolutiva. Por ello, las teorías emergentistas que veremos a continuación han sido escogidas por distanciarse tanto de posturas dualistas como de las reduccionistas, pero fundamentalmente porque no descuidan la importante cuestión del tránsito evolutivo.

Si queremos entender los fenómenos conscientes no podemos dejar de lado la comprensión de los distintos niveles de la estructura del mundo. En este sentido, la aparición de la conciencia y de los otros fenómenos mentales es un caso más de emergencia en el que los sistemas naturales dieron origen a estructuras y propiedades novedosas. Pero la conciencia es un fenómeno emergente; en consecuencia, no se puede defender una teoría emergente del cerebro sin endosar explícitamente una versión *fuerte* de la emergencia. La versión débil no hará sino convertir eventualmente la emergencia en un epifenómeno sin ningún interés causal real.

De forma preliminar sería importante reflexionar acerca de los diferentes niveles de complejidad que aparecen como producto de procesos emergentes. Podríamos resumir al menos tres niveles de emergencia natural siguiendo la propuesta de Clayton (2004). Subrayo la idea de resumen puesto que estos niveles podrían ser subdivididos y no todos los biólogos estarían de acuerdo en una división en tres niveles¹⁶. El primero sería la aparición de la vida a partir de la materia inorgánica. La probabilidad de la emergencia de la vida es reconocida actualmente alta, dada la presencia de elementos pesados que han debido intervenir, ya sea como componentes o como catalizadores en la combinación que permitió la existencia de este nivel. Reconocemos hoy en día que las fronteras entre lo no-viviente y lo viviente no son tan claras como se creyó 50 años atrás. En efecto, hoy en día objetos como los virus

¹⁶ Emmeche, Köppe y Stjernfelt (1997), por ejemplo, distinguen hasta cuatro niveles.

tienen características que los asemejan más a cristales que a células epiteliales¹⁷. Pero si hay algo que comparten como patrones distintivos los entes vivientes son el crecimiento y desarrollo, la homeostasis, la reproducción y el intercambio energético con el entorno ambiental. Los seres vivientes, además, son susceptibles de mutaciones genéticas y ambientales siguiendo los procesos de adaptación evolutiva.

En un segundo nivel encontraríamos el monitoreo externo e interno del organismo. En el sentido biológico es una forma elemental de autoconciencia puesto que involucra no solo darse cuenta del ambiente exterior, que no es lo mismo que la percepción, y además de los propios estados internos del organismo, y la modificación de los comportamientos como resultado de la conciencia.

En un tercer nivel, que lo relacionamos con el concepto de conciencia humana explicado en la sección anterior, encontramos la autoconciencia reflexiva. Aquí el organismo se da cuenta de que se da cuenta. En este nivel encontramos fenómenos como el darnos cuenta de que estamos pensando o que conocemos nuestros propios pensamientos o que conocemos lo que es estar experimentando algunos *qualia*. Esta clasificación de Clayton tiene el valor que las experiencias del tercer nivel están ancladas en el desarrollo biológico evolutivo, al menos desde un estadio preconsciente, y el conocimiento consciente estaría construido posteriormente comenzando con el anterior.

3.2. El origen de la conciencia

Nuestras reflexiones anteriores nos llevan al problema central de este trabajo, a saber, las teorías sobre el origen del fenómeno de la conciencia. Pero deseo enfocar el problema del origen desde una perspectiva ahistórica, es decir, concentrando la discusión no tanto en lo ontogenético y filogenético sino más bien en la relación entre la conciencia y el cerebro humano. No hay un sabor no-naturalista en este enfoque, puesto que sin investigar su origen histórico, sin embargo podemos tratar de enfrentar la difícil pregunta acerca de dicha relación, a saber, ¿cuál es la relación naturalista entre la conciencia y el cerebro?

Se diría que la pregunta es ahistórica, en el sentido que nos podemos preguntar sobre cómo los estados mentales en este momento (el estado consciente, por ejemplo) se relacionan con los estados cerebrales de ahora. Esta pregunta no es otra que la ancestral pregunta del problema mente-cuerpo o mente-cerebro. Sin embargo, no todas las respuestas al problema han sido las mismas, aun cuando hayan tenido una misma motivación naturalista. Veamos las más representativas con un claro componente naturalista.

¹⁷ Ver las ilustraciones en Harrison (1989).

En los últimos 50 años se han dado, en general, tres tipos de respuestas a la pregunta por el origen de los fenómenos de la conciencia. Históricamente la primera fue la identificación de los estados mentales con los cerebrales, o la teoría de la identidad mente-cerebro (Smart, 1991 y Amstrong, 1980).

La segunda, que nació con la crítica de Putnam a la teoría de la identidad y resultó en el funcionalismo (Putnam, 1980). Cualquier versión del funcionalismo tiene en común la aseveración que estar en un determinado estado consciente es estar (realizar) un rol funcional específico relativo a algún sistema. La relación entre el cerebro y la mente es una de tipo funcional, es decir, es una especie de rol residente en el cerebro. La analogía clásica ha sido que «cerebro es a mente como *hardware* es a *software*». Lo que importa no es tanto lo que sean las mentes sino lo que hacen. Como el funcionalismo es heredero de la teoría de la identidad, el funcionalista podría identificar los estados mentales con estados cerebrales, en tanto se les tome en la relación funcional, pero bien podrían ser otros realizadores físicos, como las computadoras, eventualmente.

Una tercera posibilidad es que la relación entre la conciencia y el cerebro es una relación emergente., es decir, que los estados conscientes o propiedades son estados o propiedades de mayor nivel u orden de los estados cerebrales y sus propiedades. De acuerdo con esta visión, los estados mentales dependen del cerebro o de algún estado físico, pero no son iguales a ellos. La emergencia también ha sido entendida con el concepto filosófico de superveniencia, sosteniendo que los estados conscientes supervienen en los estados cerebrales (Kim, 1993). En las siguientes secciones presentaremos y discutiremos tres propuestas representativas de esta posibilidad naturalista.

Pero no todas las propuestas de la relación mente-cerebro o conciencia-cerebro han sido de corte naturalista. En efecto, son tres los autores contemporáneos que han combatido y obstaculizado los intentos de naturalizar la mente. Ellos son T. Nagel y C. McGinn. A estos filósofos se les conoce con el sobrenombre de «misteriocistas», apelativo que McGinn considera apropiado. Nagel inició el distanciamiento de las explicaciones naturalistas al sostener que aun cuando el fisicalismo fuera verdadero, las explicaciones basadas en un naturalismo metafísico producirían una doctrina vacía, ya que no daría cuenta de toda la riqueza de los *qualia* y fenómenos similares (1974, p. 176).

McGinn, por su lado, toma una postura aun más radical, afirmando lo siguiente: «No creo que nunca se vaya a poder especificar qué hace que el cerebro sea responsable por la conciencia, pero sí estoy seguro de que no es inherentemente milagroso» (1989, p. 349). La posición de este autor por desalentadora que parezca revela una evidente postura de limitación presente —y futura— acerca de la posibilidad de comprender el origen de la conciencia, algo así como que un delfín tratara de entender

mecánica cuántica. Quizás haya otras criaturas que puedan entender su conciencia, si la tuvieran, los humanos parece que no estamos dotados de esa capacidad.

A continuación, veamos las teorías que he escogido para discutir, a saber, las teorías emergentistas de J. Searle y D. Chalmers.

3.3. La teoría emergentista de J. Searle

El punto de partida de Searle en *El redescubrimiento de la mente*, libro en el que señala explícitamente su posición emergentista, es volver a poner a la conciencia en el lugar que nunca debió dejar, en el mundo natural (1992, p. 85). Señala que uno de los objetivos del libro es, precisamente, remover las barreras que impiden que la conciencia sea considerada un fenómeno biológico más y, que por tanto, pueda ser estudiada como cualquier otro. La conciencia, por ende, debe ser estudiada dentro del marco de dos teorías que gozan del aprecio de la comunidad científica y de toda persona instruida: la teoría atómica y la teoría de la evolución (p. 86).

Metafísicamente, Searle señala que el universo consiste enteramente de partículas que resultan en objetos pequeños, medianos y grandes, desde el nivel de los electrones hasta las galaxias, pasando por abrigos y autos. En el mundo de lo viviente, encontramos a los seres humanos junto con otros animales superiores que son parte del orden biológico como otros organismos; en otras palabras, los seres humanos son continuos con el resto de la naturaleza. Son resultado de la larga evolución biológica.

Desde el punto de vista epistemológico, las teorías mencionadas son dos formidables herramientas para la explicación. Sean sistemas muy grandes o pequeños, pueden ser explicados causalmente. Estas explicaciones pueden hacerse desde el comportamiento de los más pequeños. Entonces, muchos macrofenómenos pueden ser explicados en términos de los microfenómenos. Esta estrategia es válida tanto para seres no vivientes como los vivos. Esto es lo que Searle denomina una reducción ontológica (p. 116).

La conciencia aparece como resultado también del proceso evolutivo aunque desconocemos cómo los cerebros causan la conciencia (p. 89). Define brevemente de la siguiente la naturaleza y el origen de la conciencia:

La conciencia es una característica biológica del humano y del cerebro humano y de algunos animales. Es causada por procesos neurobiológicos y es tan parte del orden biológico natural como cualesquiera otras características tales como la fotosíntesis, la digestión o la meiosis (p. 90).

Interesantemente, la posición de Searle es la de la naturalización de la conciencia y está en contra de cualquier pretensión misteriocista de la misma. Específicamente

critica las actitudes que tendían a colocar a la conciencia fuera del orden natural como cuando afirma lo siguiente:

La exclusión de la conciencia del mundo natural fue una útil herramienta heurística en el siglo diecisiete, porque permitía a los científicos concentrarse en los fenómenos que fueran medibles, objetivos y carentes de significado, esto es, libres de intencionalidad. Pero la exclusión estaba basada en una falsedad. Estaba basada en la creencia falsa que la conciencia no es parte del mundo natural (p. 93).

Esta creencia, por supuesto, impidió un estudio naturalista de la conciencia, y en ese sentido, retrasó nuestro avance en la comprensión de la misma.

Pero, aquí aparece una asimetría. Mientras que sugiere la reducción para estudiar los fenómenos, en el caso de la conciencia no se puede proceder de ese modo. Y esto porque:

La conciencia es una propiedad emergente del comportamiento de las neuronas, y por lo tanto la conciencia es causalmente reducible a los procesos cerebrales. Pero —y esto es lo que parece tan chocante— una ciencia perfecta del cerebro no podría aún conducir a una reducción ontológica de la manera como nuestra ciencia presente puede reducir el calor, la solidez, el color, o el sonido (p. 116).

La respuesta a esta asimetría la toma de la crítica al reduccionismo de lo mental, como lo han hecho otros, de T. Nagel (1974), es decir, por razón de las experiencias fenoménicas. Pero no es una limitación epistémica la que Searle tiene en mente, es decir, puesto que no podemos ser murciélagos, no podemos saber que es sentir como un murciélago. Él considera que la diferencia es de orden ontológico. Es un asunto acerca de cuáles son las cosas que existen en el mundo, y una de ellas es la realidad de los estados conscientes siguiendo la metáfora de Nagel: «Ninguna descripción en tercera persona, objetiva, de hechos fisiológicos, puede proporcionar el carácter subjetivo, de primera persona del dolor, simplemente porque las características de la primera persona son diferentes a las características de la tercera persona» (1992, p. 117). Entonces, ¿cómo se origina esta realidad ontológica distinta?

La explicación obliga a Searle a introducir el concepto de emergencia. Según Searle la conciencia es resultado de los procesos neuronales de más bajo nivel, pero emerge de estos procesos, de manera que es una «propiedad emergente» del cerebro:

Una propiedad emergente de un sistema es una que es causalmente explicada por el comportamiento de los elementos del sistema; pero no es una propiedad de los elementos individuales y no puede ser explicada simplemente como la sumatoria de las propiedades de esos elementos. La liquidez del agua es un buen ejemplo: el comportamiento de las moléculas de H₂O explican la liquidez pero las moléculas individuales no son líquidas (1997, p. 18).

De este pasaje se puede inferir la posición que tiene Searle sobre la emergencia. En primer lugar, es una emergencia de tipo fuerte, porque no es reducible. Además la emergencia es producto de la causalidad. Nos podríamos preguntar por qué Searle tiene una visión causal de la emergencia.

Parte de la respuesta está en la visión naturalista y evolutiva del pensamiento de Searle. Son las condiciones y propiedades neurales las que causan la aparición de la conciencia y las neurales, a su vez, proceden causalmente de niveles inferiores de complejidad, lo cual nos llevaría a pensar que los eventos neurales son también emergentes de realidades de nivel inferior.

Por otro lado, de aceptar una causación emergente, Searle apunta a dar respuesta al problema clásico de interacción entre actos mentales y cerebrales. Si la conciencia es causalmente emergente del cerebro, entonces no hay generación misteriosa de ella. Sin embargo, hay que advertir, que aun cuando la emergencia sea causal, eso no descartaría la posibilidad de que la conciencia sea un epifenómeno y, por tanto, no tenga ningún rol causa efectivo, con lo cual este autor estaría atrapado en la red que, aparentemente, quiere evitar.

Sin embargo, el problema se ve más claro al analizar la posición que sustenta Searle cuando discute el concepto de superveniencia. Señala que su propuesta se enmarca dentro de la superveniencia, en el sentido que los estados mentales supervienen en los estados neurofisiológicos. Utilizando el célebre ejemplo de Putnam de los «cerebros en la cubeta», dice que dos cerebros que fueran idénticos molécula por molécula resultarían teniendo los mismos fenómenos mentales. La superveniencia de lo mental sobre lo físico «está marcado por el hecho de que los estados físicos son causalmente suficientes aunque no necesariamente causalmente necesarios para los correspondientes estados mentales» (1992, p. 125). Añade que la relación de superveniencia es causal, de esa manera evita la acusación de epifenomenalismo. Sin embargo, la propuesta da un resultado inesperado.

«Me parece obvio de todo lo que sabemos del cerebro que los fenómenos macro mentales son causados por los fenómenos micro del menor nivel. No hay nada misterioso acerca de la causación ascendente; es bastante común en el mundo físico» (pp. 125-126). Si la causación es ascendente, entonces el trabajo causal es hecho por el nivel micro. ¿Cuál sería entonces el trabajo causal del macro fenómeno? Porque no estamos hablando del origen evolutivo de la conciencia, sino de los fenómenos de la conciencia que se están generando en este preciso instante. El camino hacia el reduccionismo que el mismo Searle aborrece, estaría siendo pavimentado por él mismo.

El mismo hace un giro final inesperado: «una vez que se reconoce la existencia de las formas de causación ascendente, de lo micro a lo macro, la noción de superveniencia no hace ningún trabajo en filosofía» (p. 126).

Es en este punto que la ontología de la conciencia que postula Searle no parece ser lo irreducible que había propuesto. Porque si los eventos mentales no son reducibles a eventos cerebrales, la causación ascendente parece decir lo contrario, dejando a los eventos mentales inertes, puesto que si sostenemos que tenemos un estado mental que causa, supongamos, otro acto mental, en realidad, lo micro está causando otro micro proceso y lo mental sería simplemente un nombre heurístico que, por el momento, utilizamos hasta que una neurología madura termine el trabajo eliminativista.

La tesis de ontología autónoma de la conciencia de Searle tendría que ser robustecida con una posición emergente más radical, particularmente, si es que se quisiera considerar que es causalmente importante. Mi impresión es que este autor quiere tener una tesis fisicalista radical, fruto de su compromiso con una realidad físicamente homogénea, conviviendo con un emergentismo sustancial eficiente. Las convivencias con los opuestos terminan en divorcios, o no causan felicidad, o no causan nada, para el caso.

3.4. La teoría emergentista de D. Chalmers

La preocupación de Chalmers por el fenómeno de la conciencia lo lleva a escribir su obra más ambiciosa, *La mente consciente: en búsqueda de una teoría fundamental* (1996) en la que se argumenta que las explicaciones reduccionistas de la conciencia son imposibles y que la única forma de enfrentar este problema es abandonar un marco de referencia puramente materialista. Parte de este trabajo pretende desarrollar una teoría de la conciencia en que se pueda relacionar las leyes físicas con la experiencia fenoménica de manera sistemática. En un sentido estricto no se podría considerar la obra de Chalmers como una nítida defensa del emergentismo; sin embargo, otras obras suyas y aun esta misma son maneras de defender, por un lado, la autonomía ontológica de la conciencia y su correspondiente explicación no reduccionista y, por el otro, buscar una visión naturalista a la experiencia fenoménica. Estos objetivos podemos enmarcarlos en una visión emergentista, como veremos.

Toda teoría emergentista anclada en una teoría naturalista debe comprender la conciencia como un fenómeno emergente en el mundo natural, incluidos los actos que consideramos mentales como la volición, las actitudes proposicionales, el pensamiento. Además, estos actos deben tener causación efectiva y requieren una explicación naturalista. Ninguna teoría de la conciencia puede desconocer estos fenómenos y es precisamente lo que Chalmers aborda en una obra anterior a la mencionada (1995). Al problema de dar respuesta al origen de estos fenómenos lo denomina «el problema difícil», y, sin embargo, todas las respuestas que se han dado

al problema de la conciencia son las respuestas a los problemas «fáciles». Anota que los problemas fáciles son los siguientes:

- Habilidad para discriminar, categorizar, y reaccionar a los estímulos del ambiente;
- La integración de la información a través de un sistema cognitivo;
- La reportabilidad de los estados mentales;
- La habilidad de un sistema para acceder a sus propios estados internos;
- El foco de atención;
- El control voluntario de la conducta;
- La diferencia entre la vigilia y el sueño (pp. 200-201).

Todos estos fenómenos están asociados a la conciencia. Claro que suena extraño que los considere problemas fáciles y que además, existen respuestas. Sería muy estimulante saber, por ejemplo, de respuestas convincentes a comportamientos como el control deliberado de la conducta desde un punto de vista neurofisiológico nomás. Pero según Chalmers, por más difíciles que nos puedan parecer estos problemas no son nada comparados con los que siguen:

El verdadero problema de la conciencia es el problema de la *experiencia*. Cuando pensamos y percibimos hay un zumbido de procesamiento de información, pero también hay un aspecto subjetivo [...] El aspecto subjetivo es la experiencia. Cuando vemos, por ejemplo, experimentamos sensaciones visuales: la cualidad sentida de lo rojo, la experiencia de lo oscuro y de la luz, la cualidad de la profundidad en un campo visual. Otras experiencias siguen a la percepción en diferentes modalidades: el sonido de un clarinete, el olor de las polillas. Luego están las sensaciones corporales, desde los dolores a los orgasmos; las imágenes mentales que se evocan internamente; la cualidad sentida de la emoción, y la experiencia de una corriente de pensamiento consciente. Lo que une a todos estos estados es que hay algo que significa estar en ellos. Todos ellos son estados de la experiencia (pp. 201-202).

Esta cita, un tanto extensa, nos revela la problemática de enfrentar las explicaciones en términos de funciones o estructuras, puesto que se podría conocer la estructura o la función de alguna experiencia y, sin embargo, no saber lo que es experimentarla. Lo que lleva a Chalmers a concluir, correctamente creo, que las explicaciones de la biología o de la neurología en particular no parecen ser suficientes. La neurología puede entender las estructuras y funciones de las células, los órganos, las regiones cerebrales y hasta los organismos. Nos pueden proporcionar las neurociencias teorías que logran tener correlatos neurales de la experiencia consciente.

Pero estas explicaciones no parecen agotar la pregunta acerca de lo que son estas experiencias. Algunos teóricos desean añadir algunos ingredientes más. Y es aquí cuando el terreno se hace más arenoso y los peligros (como el epifenomenalismo) empiezan a incrementarse.

Chalmers tiene algunas propuestas. Una década después de sus publicaciones (2005 y 2006), reconoce que sus posición es muy cercana al emergentismo fuerte asociado con la causación descendente, combinación que se presenta como robusta, puesto que un emergentismo débil sería incoherente con la causación descendente (si se puede reducir todo finalmente a la física y microfísica, ¿para qué molestarse en explicar la causación mental?). Dice:

¿Existen fenómenos emergentes fuertes? Mi visión es que la respuesta a esa pregunta es sí. Creo que hay exactamente un caso de un fenómeno emergente fuerte, y es el fenómeno de la conciencia. Podemos decir que el sistema es consciente cuando hay algo que es *ser* ese sistema; esto es, cuando hay algo que siente como la perspectiva del propio sistema. Es un hecho clave de la naturaleza que esta contiene sistemas conscientes; yo soy uno así. Y hay una razón para creer que los hechos acerca de la conciencia no son deducibles de ningún número de hechos físicos (2006, p. 248).

Entonces, ¿cómo se relacionan los eventos conscientes con los eventos cerebrales? Chalmers recurre al concepto de superveniencia: «La conciencia superviene en el dominio físico del mundo, parece ser un hecho nómico que la duplicación de los estados físicos duplicará la conciencia; pero en otros mundos con diferentes leyes, un sistema físico idéntico a mí podría no tener conciencia en absoluto» (p. 248). Podemos ver que, en este caso, el autor se compromete con la tesis de la superveniencia, pero solo con la local, no con la global¹⁸.

De otro lado, la explicación última del origen de la conciencia está bajo el paraguas de sus tres principios desarrollados en Chalmers (1996):

- *El principio de coherencia estructural.* Este principio de coherencia se da entre la estructura de la conciencia y la de la subjetividad. Los contenidos de la subjetividad deben de ser entendidos como los contenidos informacionales que son accesibles a los sistemas centrales de una manera generalizada para el control global directo. Los contenidos, añade Chalmers, podrían ser expresados en un sistema de lenguaje y ser potencialmente reportables.
- *El principio de invariancia organizacional.* Este principio establece que cualquiera dos sistemas con la misma organización funcional al mínimo detalle

¹⁸ Para la diferencia ver Kim (1993, pp. 53-91).

deberán tener las mismas experiencias cualitativas. Lo que importaría en este principio es que la emergencia de la experiencia no esté determinada por una realización física específica sino por el patrón abstracto de la interacción causal entre sus componentes.

- *La teoría de la información de doble aspecto.* Los dos principios anteriores son considerados por Chalmers como no-básicos. Pero la teoría de la información de doble aspecto parte de la observación que hay un isomorfismo entre la información físicamente enraizada y ciertos espacios de información fenoménica. La información, entonces, tiene dos aspectos: uno físico y otro fenoménico. Para este auto esa diferencia de aspectos puede subyacer y explicar la emergencia de la *experiencia* a partir de lo físico. La experiencia surge en virtud de ser un aspecto de la información, cuando el otro aspecto se encuentra incorporado en el procesamiento físico. Este principio es el que permitiría hacer descripciones distintas del mismo fenómeno: una neuróloga podría observar desde el exterior (tercera persona) en términos de estructuras y funciones, mientras que nosotros podríamos tener la experiencia fenoménica desde una perspectiva solo describible en primera persona.

Estos principios son altamente debatibles y controversiales. El mismo Chalmers lo reconoce. No obstante, tiene el mérito de someterlos a la confrontación teórica y empírica.

Su visión emergente fuerte es clara en mi interpretación. No solo las microestructuras y sus dinámicas pueden producir efectos macroestructurales y macrodinámicos sino también lo macrodinámico puede volverse irreductible, en tanto propiedad emergente fuerte.

La pregunta que nos podríamos hacer es si la propuesta de Chalmers es, en el fondo, una que permita el avance en la comprensión de la conciencia a partir de la emergencia.

Considero que el mayor valor que tiene su propuesta es recordarnos lo importante de enfrentar aquellos problemas duros y no simplificarlos, como se ha hecho en el pasado. Sin embargo, las respuestas que ofrece a cambio, no son convincentes. En primer lugar, es cuestionable que se haga una diferencia tal entre la primera y la tercera persona que, de alguna manera, hace inalcanzable la posibilidad de entender el origen de las sensaciones que solemos atribuir a nuestras experiencias fenoménicas. La idea de dividir y vencer no es tan mala en ciertos casos. Se le concede que la experiencia fenoménica es el todo y que este no puede ser lo mismo que las partes neurológicas constituyentes, pero de esta certeza simplemente no se sigue que estudiar algunas de sus partes proporcione alguna intuición que nos permita entender ese difícil fenómeno.

No es el reduccionismo la alternativa para la comprensión de los fenómenos, eso nos ha quedado claro y con el tiempo la actitud no-reduccionista se ha convertido en una práctica habitual en la metodología científica, salvo los Churchland (1991 y 1994), pero la biología ciertamente nos da lecciones acerca del conocimiento de algunas funciones vitales a partir de las estructuras químicas, pues en la psiquiatría hemos hecho notables avances en la comprensión de los fenómenos psíquicos desde el estudio de algunos componentes cerebrales, lo que naturalmente no agota la realidad superior que supone la conciencia humana. No obstante, tener la actitud hostil puede llevar a lo que Dennett correctamente señala:

Hago el reclamo paralelo acerca de las pretendidas «cualidades subjetivas» o «*qualia*» de la experiencia: si no empiezas a dividir las en sus componentes (funcionales) desde el principio, y distribuir las en tu modelo, creas un monstruo, un imaginario resplandor en el ojo del homúnculo cartesiano (1995, p. 3).

Un segundo punto que preocupa en el modelo emergentista de Chalmers es su casi desprecio por la teoría evolucionista como proveedora de explicaciones potenciales para la aparición de la conciencia. En efecto, en sus afirmaciones y demostraciones tanto para la formulación de los principios mencionados como para el desarrollo de la teoría fundamental hace nula referencia a los procesos evolutivos. Pero es una presuposición básica de la biología que el proceso evolutivo fue responsable del incremento de nuestra capacidad cerebral en el nivel filogenético y la presión del ambiente en el que operaba en el nivel ontogenético. Su posición acerca de la evolución se lee en el siguiente pasaje:

Esta idea [la explicación evolutiva de la conciencia] sobreestima lo que una explicación evolutiva puede proporcionarnos. El proceso de selección natural no puede distinguir entre yo y mi gemelo zombi. La evolución selecciona las propiedades de acuerdo con su rol funcional, y mi gemelo zombi realiza todas las funciones que yo realizo tan bien como yo [...]. Se sigue que la evolución por sí sola no puede explicar por qué las criaturas conscientes en vez de los zombis evolucionaron (1996, p. 120).

Si bien es cierto que ninguna teoría ofrece explicaciones completas, y no tiene por qué hacerlo, los recursos de la teoría evolutiva son mejorables con la investigación y los datos empíricos. El estudio evolutivo de la emergencia del cerebro y de su consecuente conjunto de fenómenos conscientes constituye un enfoque naturalista al problema difícil que Chalmers nos recuerda. La neurofisiología parte del estudio de las estructuras biológicas y funciones que compartimos con otros primates, y todos esos estudios se basan en la teoría evolutiva como marco referencial. Para explicar

la superveniencia de lo mental sobre lo físico, como ejemplo de la emergencia evolutiva no se puede prescindir de la historia natural que causó que los cerebros y el sistema nervioso en general evolucionen. Prescindir de la teoría, precisamente no nos permitiría explicar que hay seres que experimentan *qualia* y no zombis que parecen hacerlo. Más difícil sería postular la existencia de zombis dentro del esquema de la teoría de la evolución, porque, en general, la teoría no contempla los casos en que los muertos se comportan como vivos.

REFERENCIAS BIBLIOGRÁFICAS

- Alexander, Samuel (1920). *Space, Time and Deity*. Volumen II. Londres: Macmillan.
- Archinov, Vladimir & Christian Fuchs (eds.) (2003). *Emergence, Causality, Self-Organisation*. Moscú: NIA-Priroda.
- Aristóteles (2007). *On the Soul [De anima]*. Traducción de J. A. Smith. Adelaide: University of Adelaide. <http://ebooks.adelaide.edu.au/a/aristotle/a8so/>
- Armstrong, David (1980). *The Nature of Mind*. Brisbane: The University of Queensland Press.
- Bickerton, Derek (2009). *Adam's Tongue: How Humans Made Language, How Language Made Humans*. Nueva York: Hill and Wang.
- Baker, Lynne Rudder (1993). Metaphysics and Mental Causation. En John Heil y Alfred Mele (eds.), *Mental Causation* (pp. 75-98). Nueva York: Oxford University Press.
- Braun, Ricardo (2009). El rol de la filosofía y la neurociencia. En Pablo Quintanilla (ed.), *Ensayos de metafilosofía* (pp. 157-178). Lima: Fondo Editorial PUCP
- Braun, Ricardo (2011). La conciencia humana y el emergentismo. *Persona*, 14, 59-185.
- Block, Ned (1995). On a Confusion about a Function of Consciousness. *Behavioral and Brain Sciences*, 18, 227-247.
- Broad, Charlie Dunbar (2001). *The Mind and its Place in Nature*. Oxford: Routledge.
- Bunge, Mario (1967). *Scientific Research. Strategy and Philosophy*. Nueva York: Springer-Verlag.
- Bunge, Mario (1977). Levels and Reduction. *American Journal Regulatory, Integrative and Comparative Physiology*, 233(3), 75-82.
- Bunge, Mario & Rubén Ardila (1988). *Filosofía de la psicología*. Barcelona: Ariel.
- Burnet, John (2005). *Early Greek Philosophy*. Londres: Adamant Media Corporation.
- Campbell, Donald (1974). La «causación-descendente» en los sistemas biológicos jerárquicamente organizados. En Francisco José Ayala y Theodosius Dobzhansky (eds.), *Estudios sobre la filosofía de la biología* (pp. 235-245). Barcelona: Ariel.

- Caston, Victor (1997). Epiphenomenalisms, Ancient and Modern. *The Philosophical Review*, 106, 309-363.
- Chalmers, David (1995). Facing up the Problem of Consciousness. *Journal of Consciousness Studies*, 2, 200-219.
- Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Nueva York: Oxford University Press.
- Chalmers, David (2006). Strong and Weak Emergence. En Philip Clayton y Paul Davies (eds.), *The Reemergence of Emergence: The Emergentist Hypothesis from Science to Religion* (pp. 244-256). Nueva York: Oxford University Press.
- Churchland, Paul (1991). Eliminative Materialism and the Propositional Attitudes. En David Rosenthal (ed.), *The Nature of Mind* (pp. 601-612). Nueva York: Oxford University Press.
- Churchland, Patricia (1994). Can Neurobiology Teach Us Anything about Consciousness? Discurso presidencial la American Philosophical Association, Pacific Division, En *Proceedings and Addresses of the APA* (pp. 23-40). Lancaster: Lancaster Press.
- Clayton, Philip (2004). *Mind and Emergence: From Quantum to Consciousness*. Nueva York: Oxford University Press.
- Clayton, Philip (2006). Conceptual Foundations of Emergence Theory. En Philip Clayton y Paul Davies (eds.), *The Emergence: The Emergentist Hypothesis from Science to Religion* (pp. 1-34). Nueva York: Oxford University Press.
- Crane, Tim (2001). The Significance of Emergence. En Carl Gillett y Barry Loewer (eds.), *Physicalism and its Discontents* (pp. 207-224). Cambridge: Cambridge University Press.
- Dennett, Daniel (1995). Our Vegetative Soul - Review of Damasio, Descartes' Error. *Times Literary Supplement*, 25 de agosto, 3-4.
- Emmeche, Claus, Simo Køppe, S. & Frederik Stjernfelt (1997). Explaining Emergence: Towards an Ontology of Levels. *Journal for General Philosophy of Science*, 28, 83-117.
- Engels, Federico (2003). *Anti-Dühring*. Moscú: Instituto del Marxismo-Leninismo.
- Flanagan, Owen (1992). *Consciousness Reconsidered*. Cambridge: The MIT Press.
- Fodor, Jerry (1991). You Can Fool Some of the People All of the Time, Everything Else Being Equal: Hedged Laws and Psychological Explanation. *Mind*, 100, 19-34.
- Fuchs, Christian (2003). Dialectical Philosophy and Self-Organisation. En Vladimir Arshinov y Christian Fuchs (eds.), *Emergence, Causality, Self-Organisation* (pp. 195-244). Moscú: NIA-Priroda.

- Galeno (1968). *On the Usefulness of the Parts of the Body*. Traducción de Margaret Tallmadge May. Nueva York: Cornell University Press. <http://home-page.mac.com/kvmagruder/hsci/06-Roman/source/galen.html>
- Goldstein, Jeffrey (1999). Emergence as a Construct: History and Issues. *Emergence*, 1, 49-72.
- Guttenplan, Samuel (1995). *A Companion to the Philosophy of Mind*. Cambridge: Basil Blackwell.
- Guthrie, William Keith Chambers (1997). *The Greek Philosophers: from Thales to Aristotle*. Londres: Routledge.
- Harrison, Stephen (1989). What do Viruses Look Like? *Harvey Lectures*, 85, 127-152. <http://crystal.harvard.edu/lib-sch/HarrisonS-91-HLect-85-127.pdf>
- Horgan, Terence (1993). From Supervenience to Superdupervenience: Meeting the Demands of a Material World. *Mind*, 10, 555-586.
- Horgan, Terence & John Tienson (1990). Soft Laws. *Midwest Studies in Philosophy*, 15, 256-279.
- Humphreys, Paul (1997). How Properties Emerge. *Philosophy of Science*, 64, 1-17.
- Jackson, Frank (1986). What Mary Didn't Know. *Journal of Philosophy*, 83, 291-295.
- Kant, Immanuel (1876). Crítica del juicio seguida de las observaciones sobre el asentimiento de *lo bello y lo sublime*. Traducción de Alejo García Moreno y Juan Ruvira. Madrid: Iruvredra. <http://www.cervantesvirtual.com/servlet/SirveObras/00364956451381094121157/index.htm>
- Kim, Jaegwon (1993). *Supervenience and Mind: Selected Philosophical Essays*. Nueva York: Cambridge University Press.
- Kim, Jaegwon (1999). Making Sense of Emergence. *Philosophical Studies*, 95, 3-36.
- Kim, Jaegwon (2000). *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge: The MIT Press.
- Levine, Joseph (2001). *Purple Haze: The Puzzle of Consciousness*. Nueva York: Oxford University Press.
- McGinn, Colin (1989). Can We Solve the Mind-Body Problem? *Mind*, 89, 349-366.
- McGinn, Colin (1991). *The Problem of Consciousness*. Oxford: Basil Blackwell.
- McLaughlin, Brian (1992). The Rise and Fall of British Emergentism. En Ansgar Berckermann, Hans Flohr y Jaegwon Kim (eds.), *Essays on the Prospect of Nonreductive Physicalism* (pp. 49-93). Berlín: Walter de Gruyter.
- McLaughlin, Brian (1997). Emergence and Supervenience. *Intellectica*, 25, 25-43.
- Mill, John Stuart (1843). *A System of Logic, Ratiocinative and Inductive*. Nueva York: Harper & Bros. <http://www.gutenberg.org/files/27942/27942-h/27942-h.html#toc53>

- Morgan, Conwy Lloyd (1923). *Emergent Evolution*. Londres: Williams and Norgate.
http://www.brocku.ca/MeadProject/Morgan/Morgan_1923/Morgan02_01.html
- Nagel, Ernest (1961). *The Structure of Science*. Indianápolis: Hacket Publishing Co.
- Nagel, Thomas (1974). What Is It Like to Be a Bat? *Philosophical Review*, 83, 7-22.
- Nagel, Thomas (1989). *The View from Nowhere*. Nueva York: Oxford University Press.
- Nagel, Thomas (2009). *The Tacit Dimension*. Chicago: The University of Chicago Press.
- Putnam, Hilary (1980). The Nature of Mental States. En Ned J. Block (ed.), *Readings in the philosophy of psychology*. Volumen 1 (capítulo 11). Cambridge: Harvard University Press.
- Reid, Robert (2007). *Biological Emergences: Evolution by Natural Experiment*. Cambridge: The MIT Press.
- Robinson, William (2007). Epiphenomenalism. En *Stanford Encyclopedia of Philosophy*.
<http://plato.stanford.edu/entries/epiphenomenalism/>
- Sarkar, Sahotra & Jessica Pfeifer (2006). *The Philosophy of Science. An Encyclopedia*. Nueva York: Routledge.
- Searle, John (1992). *The Rediscovery of the Mind*. Cambridge: The MIT Press.
- Searle, John (1997). *The Mystery of Consciousness*. Nueva York: The New York Review of Books.
- Smart, Jamieson Carswell (1991). Sensations and Brain Processes. En David Rosenthal (ed.), *The Nature of New Mind* (pp. 169-176). Nueva York: Oxford University Press.
- Sperry, Roger (1980). Mind-Brain Interaction: Mentalism, Yes; Dualism, No. *Neuroscience*, 5, 195-206. <http://people.uncw.edu/puente/sperry/sperry-papers/80s-90s/215-1980.pdf>
- Stebbing, Lizzie Susan (1950). *A Modern Introduction to Logic*. Londres: Methuen and Co.
- Stephan, Achim (1992). Emergence - A Systematic View on its Historical Facets. En Ansgar Beckermann, Hans Flohr y Jaegwon Kim (eds.), *Emergence and Reduction: Essays on the Prospects of Nonreductive Physicalism* (pp. 25-48). Berlín: Walter de Gruyter.
- Stephan, Achim (1999). Varieties of Emergentism. *Evolution and Cognition*, 5, 49-59.

CÓMO SER NATURALISTA SIN SER REDUCCIONISTA EN EL ESTUDIO DE LA MENTE Y EL LENGUAJE: INTERSUBJETIVIDAD, COMPRENSIÓN LINGÜÍSTICA Y ATRIBUCIÓN PSICOLÓGICA

César Escajadillo

Pontificia Universidad Católica del Perú

1. INTRODUCCIÓN

Al mostrar la insuficiencia de la concepción tradicional de las especies, Darwin transformó por completo la imagen del ser humano aceptada por siglos. Uno de los factores que impulsó este cambio fue la convicción de que era necesario explicar los fenómenos biológicos en términos de causas naturales —sin apelar a entidades o fuerzas situadas más allá de la naturaleza—, convicción que hasta hoy caracteriza a los programas de investigación identificados con el *naturalismo*.

Aun cuando sea prácticamente imposible ofrecer una caracterización exacta de qué se entiende por naturalismo hoy en día, es posible identificar, en el naturalismo contemporáneo, un compromiso fundamental con la siguiente afirmación: «no aceptar ninguna entidad o explicación cuya existencia o verdad contradiga las leyes de la naturaleza hasta donde las conocemos» (De Caro & Voltolini, 2010, p. 71). Esta *afirmación constitutiva del naturalismo*, como la denominan los autores, excluye de plano cualquier entidad o forma de explicación sobrenaturales; sin embargo, deja abierta la pregunta de qué entidades debemos admitir —si es que existen, además de entidades físicas, entidades no reductibles a las físicas— y cómo, o con qué métodos, debemos estudiar la realidad —si hay cabida para métodos alternativos, quizá no reductibles, a los de la ciencia natural—. Las preguntas «¿qué existe?» y «¿cómo estudiar lo que existe?» corresponden, pues, respectivamente, a los planos ontológico y metodológico del naturalismo, los mismos que atraviesan aquellos fenómenos que pueden ser abordados bajo esta perspectiva. El tema de este artículo es cómo pensar en el plano metodológico del naturalismo cuando el objeto de estudio es la mente y el lenguaje.

Desde luego, el estudio de la mente y el lenguaje es un campo sumamente amplio en el que confluyen diversos intereses y disciplinas, desde la filosofía hasta las neurociencias, pasando por la lingüística, la psicología, la ciencia cognitiva y la inteligencia artificial. Una aproximación metodológica que, en las últimas décadas, ha sido bien recibida en este campo de estudio consiste en ubicar a cada una de estas disciplinas en un nivel de explicación específico al interior de una misma *estructura explicativa unificada* (Bermúdez, 2005, p. 28). Semejante aproximación se sostiene en el principio de que cada nivel de la estructura esclarece el nivel inmediatamente superior, dando así sentido a la idea de que cada disciplina investiga lo mismo pero a niveles de explicación distintos.

Así, en el nivel más alto de la estructura encontramos el *nivel personal*, aquel que trata con la explicación y predicción de la conducta en los términos que corresponden al vocabulario de la psicología *folk* o de sentido común. Lo característico de la explicación a este nivel es que involucra estados accesibles a la consciencia y que son atribuidos a la persona como un todo —el ejemplo característico son las actitudes proposicionales: creer que, desear que, tener la intención de, etcétera—, lo que contrasta con el tipo de explicación que encontramos a *nivel subpersonal*, el cual trata con estados no accesibles a la conciencia y que solo son atribuibles a partes de la persona, como los estados computacionales y neurofisiológicos que intervienen en los procesos cuyo estudio pertenece a las disciplinas que ocupan los niveles más bajos de la estructura explicativa¹.

Ahora bien, la aproximación metodológica basada en el modelo de una estructura explicativa unificada, en cuyo centro se halla la distinción entre el nivel personal y subpersonal, da lugar a la pregunta de si existen o no relaciones verticales, de dependencia explicativa, entre los niveles personal y subpersonal; entre el nivel que corresponde a la explicación psicológica de sentido común y el que corresponde a la explicación en términos de estados y procesos de carácter no consciente. De manera más precisa, podemos formular la pregunta de la siguiente manera: ¿pueden los hechos de nivel personal —hechos relativos a la conducta de una persona— ser explicados reductivamente sobre la base de hechos de nivel subpersonal —hechos que describen la relación entre distintas partes de la persona—? En este artículo consideraremos dos respuestas a esta pregunta, ambas de corte naturalista, pero signadas

¹ Para una caracterización más completa de la diferencia entre el nivel personal y subpersonal de explicación véase Bermúdez (2005, pp. 16-39). La distinción entre el nivel personal y subpersonal fue propuesta originalmente por Dennett (1969), aunque él manifiesta que esta ya se encontraba implícita en Ryle y Wittgenstein. Dennett mismo ofreció después una distinción tripartita que complementa y modifica la anterior (1989). Para una discusión de las razones que llevaron a Dennett a modificar su distinción original, véase Hornsby (2000).

por compromisos distintos; en suma, dos respuestas que representan, cada una, modos divergentes de entender el *naturalismo*.

La primera respuesta considera que el nivel personal es un nivel de explicación autónomo cuya relación con los niveles más bajos, o subpersonales, es indirecta y no reducible. Llamaremos a esta posición *naturalismo no reduccionista*². La idea detrás de esta forma de entender el naturalismo es que la socialización humana nos ha provisto de un punto de vista —de una *segunda naturaleza*, en el decir de McDowell (1996)— desde el cual toma sentido la concepción que se tiene del ser humano como un ser natural «empapado de racionalidad»; alguien cuya conducta podemos entender y predecir solo si asumimos que se trata de un ser racional: un agente con creencias y deseos que son coherentes y consistentes entre sí. Para el defensor de esta posición, la explicación que ofrecemos de la conducta de otros agentes se sustenta en una serie de principios *normativos* —racionalidad, coherencia, consistencia— que muestran cómo los agentes *deberían* actuar en condiciones ideales, principios que no desempeñan papel alguno en la explicación de los fenómenos a nivel subpersonal. Desde un punto de vista metodológico, la implicancia de esta posición es que algunas propiedades de la mente y el lenguaje deben ser estudiadas tomando el nivel personal como un nivel de explicación autónomo, el cual no necesita ser esclarecido por medio de relaciones verticales con el nivel subpersonal.

Dicha concepción contrasta con aquella que considera que el nivel personal no es autónomo y necesita, por tanto, ser esclarecido por el tipo de explicaciones que encontramos en los niveles más bajos de la estructura explicativa, por ejemplo, en términos de mecanismos y estados que forman parte del diseño de la mente o del cerebro. Denominaremos a esta concepción, en contraste con la postura rival, *naturalismo reduccionista*. Su afirmación principal es que damos cuenta de los hechos de nivel personal cuando logramos reducirlos a hechos de nivel subpersonal, por ejemplo, cuando logramos establecer que los primeros son idénticos a, o el resultado directo de, los segundos. Ello implica, desde un punto de vista metodológico, que el estudio de las propiedades de la mente y el lenguaje consiste en lograr una caracterización adecuada del modo en que se relacionan verticalmente los niveles personal y subpersonal³.

² A la base de esta posición se encuentra tanto el naturalismo de la segunda naturaleza de McDowell (1996) como el naturalismo liberal o blando de Strawson (1985).

³ Como se podrá ver, la implicancia de negar la existencia de relaciones verticales explicativas interniveles no es que los hechos de nivel personal no interactúen causalmente con los de nivel subpersonal. La implicancia es que las causas que son explicativas de los hechos de nivel personal han de remitirnos al punto de vista dentro del cual esos hechos son *constituídos*: el punto de vista desde el cual pensamos en el ser humano como un ser racional. Así, sostendremos que lo que posibilita describir y explicar los hechos a nivel personal no es la ocurrencia de hechos de nivel subpersonal concomitantes sino el punto de vista

Consideramos que la oposición entre ambas formas de naturalismo —la cual se sigue de tomar postura respecto a la pregunta de si existen relaciones verticales entre los niveles personal y subpersonal— está en el centro de algunos de los debates más importantes que aparecen en las intersecciones de la filosofía, la lingüística y la psicología. Pensemos, por ejemplo, en el debate en torno a si las actitudes proposicionales pueden ser caracterizadas computacionalmente, a la manera de un *software*, o como símbolos abstractos en la mente/cerebro de los hablantes —símbolos en un «lenguaje del pensamiento» fodoriano—. Como dijimos, nuestro aporte a esta discusión se centrará, en lo fundamental, en el plano metodológico: nos interesa evaluar hasta qué punto algunas propiedades de la mente y el lenguaje —específicamente, la comprensión lingüística y la atribución psicológica— pueden ser explicadas reductivamente. Nuestra conclusión será que en ambos casos las pretensiones del naturalismo reduccionista son infundadas, conclusión que se sigue de observar la manera en que la interacción simultánea con otros individuos y la realidad compartida hace posible describir y explicar los hechos a nivel personal. Sobre la base de esta conclusión, veremos que la imposibilidad de dar cuenta de ambas propiedades a la manera reduccionista no implica renunciar al naturalismo cuando se estudia la mente y el lenguaje.

2. DAVIDSON Y CHOMSKY SOBRE LA COMPRENSIÓN Y LA COMPETENCIA LINGÜÍSTICAS

Nuestro punto de partida es la distinción que traza Chomsky entre dos maneras de aproximarse al estudio de la mente y el lenguaje: una, de corte internalista, comprometida con lo que Chomsky llama *naturalismo metodológico*; la otra, de corte externalista, comprometida con un *dualismo metodológico* (2000, p. 76). Como veremos, se trata de una distinción que apunta a mostrar qué compromisos deben prevalecer en el estudio empírico de la mente, en particular del lenguaje, y cuáles deben rechazarse por ser infundados o acientíficos.

Según Chomsky, una aproximación naturalista a la mente «investiga los aspectos mentales del mundo como cualesquiera otros, buscando construir teorías explicativas inteligibles con la esperanza de integrarlas a las ciencias naturales “básicas”»⁴ (2000, p. 76). El modelo de investigación científica en que se apoya esta aproximación es uno que intenta mostrar, en última instancia, cómo la descripción que hacemos

que emerge al interactuar simultáneamente con otros individuos y la realidad compartida. Dicho punto de vista, al ser irreductiblemente interpersonal o de segunda persona, asegura la autonomía explicativa del nivel personal.

⁴ En adelante, salvo se indique lo contrario, todas las traducciones son mías (nota de autor).

de ciertas propiedades mentales a nivel abstracto arroja luz sobre ciertas propiedades físicas más fundamentales, en este caso, propiedades desconocidas del cerebro humano (Chomsky 2001, p. 582). Dicha aproximación contrasta con la perspectiva del llamado *dualismo metodológico*, el cual, según Chomsky, abandona los métodos racionales de la investigación empírica al estudiar a los seres humanos «del cuello para arriba», imponiendo sobre el objeto de estudio demandas arbitrarias, de tipo a priori, que nunca serían admitidas por la ciencia natural. A juicio del autor, esta es la perspectiva que mayor influencia ha tenido en la filosofía de la mente y del lenguaje de las últimas décadas, una que resulta altamente controversial pues intenta encontrar respuestas lejos del único lugar donde se pueden obtener: «en las ciencias duras, donde la riqueza y profundidad de la comprensión provee alguna esperanza de obtener claridad sobre las preguntas» (2000, p. 77).

Pues bien, ¿a qué conclusiones nos conduce el naturalismo metodológico en el estudio de la mente y del lenguaje? Básicamente, a la posición que ve el lenguaje como un componente específico del cerebro humano. A dicho componente Chomsky lo llama *la facultad del lenguaje*. La facultad del lenguaje posee un *estado inicial* que es parte de nuestra dotación biológica innata y una serie de *estados alcanzados* durante la vida que permiten caracterizar una infinidad de expresiones lingüísticas de acuerdo con sus propiedades fonéticas y semánticas. Cada estado alcanzado por la facultad del lenguaje es resultado de un procedimiento computacional generativo al que Chomsky (2000, p. 78) llama *I-language* —*I* por ser interno e individual—, el cual asigna a las expresiones lingüísticas descripciones estructurales —compuestos con propiedades fonéticas y semánticas— que permiten producir e interpretar dichas expresiones. El *I-language* es un mecanismo de la mente/cerebro, propiamente, un estado cognitivo o mental que subyace al conocimiento que el hablante tiene del lenguaje: cómo usar el lenguaje y cómo entender diversas expresiones lingüísticas (p. 81). Chomsky reconoce que el *I-language* es algo así como un modo de hablar y entender; sin embargo, deja en claro que el conocimiento involucrado en este procedimiento no es una habilidad adquirida: se trata de la representación, en la mente/cerebro del hablante, de cierto estado alcanzado por la facultad del lenguaje (pp. 50 y 69). Todo esto, a juicio de Chomsky, no es más que práctica científica convencional según los estándares del naturalismo metodológico:

El niño, dotado de ciertas capacidades innatas, adquiere conocimiento de un lenguaje —automáticamente, y con escasa o ninguna elección sobre el tema—. El lingüista trata de averiguar qué conocimiento adquiere el niño, qué propiedades innatas de la mente/cerebro son responsables de este proceso de crecimiento del conocimiento (trata de averiguar qué sabe el niño *de manera anticipada a la experiencia*, para emplear una expresión que resulta adecuada). El lingüista

extraerá conclusiones sobre las propiedades innatas, derivadas de algún modo, para la descripción del conocimiento adquirido, en particular, para el estudio del significado, dominio que tiene el mismo estatuto que cualquier otro (p. 54).

La segunda aproximación, la del dualismo metodológico, contrasta con la anterior por mantener que no podemos estudiar la mente como se estudia cualquier otro fenómeno del mundo natural, posición que, a juicio de Chomsky, solo nos conduce al escepticismo y al misticismo en este campo de estudio. Chomsky se refiere aquí, principalmente, a la tradición filosófica de segunda mitad del siglo XX que ve a la mente y el lenguaje como fenómenos dotados de una dimensión social irreductible, tradición que se remonta a las *Investigaciones* de Wittgenstein y cuya expresión más reciente sería el paradigma davidsoniano de la interpretación radical (Davidson, 2001b, pp. 125-139). De acuerdo con Chomsky, este paradigma se erige sobre la base de dos supuestos interrelacionados, ambos falsos desde su punto de vista: primero, que en el estudio del significado —no qué son los significados, sino el modo en que un intérprete entiende lo que un hablante quiere decir— la única fuente de evidencia para el intérprete es el uso de las oraciones, de parte del hablante, en situaciones públicas reconocibles; segundo, que para explicar el modo en que el intérprete logra entender lo que el hablante quiere decir, se necesita una *teoría recursiva de la verdad*⁵ que sirva como modelo de la competencia lingüística del intérprete (2000, p. 102). Un corolario de este último supuesto, el cual Chomsky cita con desaprobación, es que «no agrega nada a esta tesis decir que si la teoría describe correctamente la competencia del intérprete, algún mecanismo en el intérprete debe corresponder a la teoría» (Davidson, 1986, p. 438, citado por Chomsky, 2000, p. 56). Con todo, Chomsky rechaza que la competencia lingüística —específicamente, la capacidad de los hablantes para entenderse por medio del lenguaje— pueda caracterizarse en los siguientes términos:

Podemos decir que la habilidad lingüística es la habilidad de converger en una teoría al paso [es decir, una teoría recursiva de la verdad] de ocasión en ocasión [...]. Pero si decimos esto, entonces debemos reconocer que no solo hemos abandonado la noción ordinaria de lo que es un lenguaje, también hemos eliminado la frontera

⁵ Se trata de la teoría que el intérprete desarrolla con el fin de determinar qué oraciones son consideradas verdaderas tanto por él como por el hablante en las mismas circunstancias. La teoría es recursiva porque apela a un conjunto limitado de reglas y axiomas para arrojar luz sobre la estructura veritativa en que se sostiene el lenguaje del hablante —el lenguaje objeto—. En vista de que Davidson suscribe la tesis central de la semántica veritativo-condicional —la tesis de que conocer el significado de una oración es saber, entre otras cosas, bajo qué condiciones la oración es verdadera o falsa (véase 2001b, p. 24)—, una teoría recursiva de la verdad ha de servir como una teoría empírica del significado para el lenguaje del hablante: una teoría que captura lo que hablante quiere decir en ocasiones específicas.

entre conocer el lenguaje y saber de modo general cómo orientarse en el mundo. Pues no hay reglas para arribar a teorías al paso, no hay reglas en un sentido estricto, solo máximas toscas y generalidades metodológicas (1986, pp. 445-446).

Al caracterizar la competencia lingüística como la habilidad de converger, de ocasión en ocasión, en una *teoría al paso* —una teoría recursiva de la verdad para el lenguaje del hablante—, Davidson está atacando la idea de que para comprender a los demás —para responder a la pregunta: ¿qué es lo que hablante quiere que el intérprete entienda al usar las palabras de cierta manera?— se necesite conocer las mismas convenciones y regularidades del idioma que otros, contar con un método de interpretación común o poseer una gramática interna similar. Su tesis es que la comprensión lingüística requiere, fundamentalmente, que cada individuo esté en comunicación con el otro: que uno *sea* entendido de la manera en que uno *espera* ser entendido, tarea que recae sobre la teoría al paso⁶. Ciertamente, el hecho de ser entendido de la manera en que uno espera ser entendido es facilitado, en circunstancias normales, por el hecho de que uno acostumbra hablar como los demás —en el sentido de *similar* que resulte más apropiado—; sin embargo, ello no implica que estemos *obligados* a hablar como otros hablan para ser entendidos. El caso del que Davidson se vale para ilustrar este punto es aquel en que el intérprete reconoce que el hablante confunde una palabra con otra cuyo sonido es parecido —como el caso de una niña que, tras visitar Disneylandia, contó a su familia de los «parques *semánticos*» que visitó—⁷, casos en que la comprensión se logra, no sobre la base de lo que las palabras significan convencionalmente, sino a partir del reconocimiento de las intenciones del hablante: qué fue lo que el hablante quiso dar a entender.

Así, la idea detrás de las *teorías al paso* —llamadas así por su carácter circunstancial y variable— es que el hecho de ser entendido de la manera en que uno espera ser entendido no supone para los hablantes obligación alguna de tipo léxico; no obliga a seguir normas del tipo: «siempre que uno quiera hablar de gatos debe usar la palabra *gato*», sino que depende más del ingenio, la suerte y la empatía. Nótese que Davidson no está diciendo que el saber lingüístico del intérprete y hablante —aquel que ambos

⁶ Davidson lo pone del siguiente modo: «Una teoría al paso es realmente como una teoría al menos en esto, en que se deriva del ingenio, la suerte y la sabiduría de un vocabulario y una gramática particulares, así como del conocimiento de los modos en que las personas expresan su punto y de reglas prácticas para reconocer qué desviaciones del diccionario son más factibles. No hay más probabilidad de enseñar o reglamentar este proceso que lo hay de enseñar o reglamentar el proceso de crear teorías nuevas para arreglárselas con información nueva en cualquier campo» (1986, p. 446).

⁷ A este tipo de lapsus involuntario se le denomina, en la lengua inglesa, como *malapropism* —del francés *mal à propos*—. Por supuesto, el lapsus también puede ser deliberado, como en el caso del título que lleva la columna que publica en *La República* el periodista Augusto Álvarez Rodrich: «frases para la *histeria* política peruana».

traen a la comunicación y es recogido por sus *teorías previas*— no sea importante para que ambos se entiendan. El punto es que la comprensión no requiere que dicho saber sea compartido de antemano. Esta posición contrasta con la de Chomsky cuando sostiene: «Otros comprenden lo que decimos porque tienen la misma naturaleza biológica y experiencia lo suficientemente similar con preferencias simples» (2001, p. 593). O cuando afirma: «puedo comprender a Jones, dentro de ciertos límites, porque mi *I-language* no es muy distinto al suyo, y porque él y yo más o menos compartimos otras propiedades desconocidas que intervienen en el intérprete total» (2000, p. 73).

Sin embargo, vemos que la inteligibilidad de esta posición se oscurece cuando reconocemos que la comprensión lingüística depende, en un sentido crucial, de la atribución recíproca de estados mentales —creencias, deseos e intenciones—, los mismos que son indispensables para determinar cómo el hablante espera ser entendido en cierta ocasión; por ejemplo, cuando dice «estoy aburrido», no con la intención de informar al intérprete de su estado anímico, sino con la intención de que este actúe de cierta manera a través del reconocimiento de su intención. Desde luego, el hablante puede decir «estoy aburrido» con la intención de informar al intérprete sobre su estado anímico actual o para vengarse por algo que el intérprete dijo el día anterior; el punto es que cuál sea la intención, o el conjunto de intenciones, que determinan tanto lo que el hablante quiere decir como el tipo de acto de habla que está realizando, así como los propósitos ulteriores que busca con él —advertir, informar, herir—, recae en la habilidad del intérprete para converger con el hablante en una teoría al paso, no en los estados de sus *I-language* respectivos.

Para apreciar mejor este punto, tomemos el caso de la *intención comunicativa griceana*, aquella cuya estructura es «trato de decirte *X* mediante *Y* para que hagas *Z*» y que guía al intérprete en la comprensión de lo que el hablante quiere que el intérprete sepa acerca de la comunicación misma. Además de ser pública o manifiesta, la intención comunicativa es recursiva porque involucra varios niveles de intencionalidad, es decir, involucra estados mentales acerca de los estados mentales propios y ajenos. Con relación a este punto, autores como Grice (1989) y Sperber y Wilson (1986) han mostrado que el reconocimiento de la intención comunicativa requiere que exista intencionalidad de tercer grado o más: es decir, además de la intención de comunicar algo, la intención de comunicar algo acerca de la comunicación misma, en una palabra, *metacomunicación*⁸. Así, solo cuando la comunicación asume una

⁸ En la propuesta de Sperber y Wilson (1986), la intención comunicativa es considerada una intención informativa de segundo orden, la cual se cumple una vez que la intención informativa de primer orden ha sido reconocida por el oyente. Desde luego, no toda forma de comunicación se basa en atribuir

forma recursiva del tipo «yo quiero₁ que tú sepas₂ y quiero que sepamos₃ juntos», el intérprete estará en condiciones de captar aquello que el hablante espera que el intérprete entienda o haga mediante el reconocimiento de su intención. En ese sentido, el carácter público y recursivo de la intención comunicativa implica que el fenómeno de la comprensión lingüística es inseparable de la metacomunicación, toda vez que el reconocimiento de la intención comunicativa requiere saber que el otro sabe y saber que ambos saben —saber que saben *juntos*—.

La réplica de Chomsky a esta posición es que el interés del naturalista metodológico no es en la comunicación como tal, sino en el estado del mecanismo interno, o subpersonal, que subyace a la comprensión que el intérprete tiene de las oraciones del hablante (2000, p. 56). Dado que el estudio del lenguaje en la línea que Chomsky suscribe se reduce al estudio del *I-language* en un momento específico, toda noción relacionada con el estudio del lenguaje y su uso deberá depender del estado del *I-language*: «el hecho es que Jones habla y entiende de la manera en que lo hace sobre la base del *I-language* que ha adquirido en el curso del crecimiento del lenguaje» (p. 72). Sin embargo, vemos que ninguna descripción subpersonal que haga referencia a mecanismos internos podría dar respuesta a la pregunta, clave en el caso del intérprete, de cómo el hablante espera que el intérprete entienda sus palabras y oraciones. Nótese que esta es una pregunta acerca de cómo atribuimos estados mentales a otros: de qué manera comprendemos lo que otros quieren decir atribuyéndoles una variedad de estados mentales. Pensar que esta pregunta puede ser respondida en términos de los mecanismos y estados que subyacen a la cognición, o con referencia a una Gramática Universal innata, es como pensar que podemos explicar la impresión que nos produce una pintura estudiando la composición del lienzo en que fue pintada.

Para refutar esta conclusión, Chomsky sostendría, en primer lugar, que a la ciencia no le interesa capturar el contenido del discurso ordinario; que hay ciertos conceptos, por ejemplo, los de creencia, significado, deseo e intención —los del vocabulario de la psicología *folk*—, que deben quedar fuera del alcance de una investigación naturalista por ser demasiado generales e imprecisos: «Los conceptos del lenguaje natural, y del sentido común en general, no son siquiera candidatos para las teorías naturalistas» (p. 22). En segundo lugar, mantendría que no es posible, como pretende el paradigma davidsoniano de la interpretación radical, una ciencia cuyo tema sea *el estudio de todo*, es decir, que el llamado *problema de la comprensión*

intenciones de este tipo, solo la denominada *comunicación inferencial* o *intencional*, cuyo carácter esencialmente cooperativo la distingue de formas de comunicación más simples.

abarque todos los aspectos relacionados con el lenguaje y su estructura que merecen ser investigados empíricamente (pp. 29, 49-50 y 69). En consecuencia:

El estudio de la comunicación en el mundo actual de la experiencia es el estudio del intérprete, pero este no es un tema para la investigación empírica por las razones habituales: no hay tal cosa como el estudio de todo. De manera similar, la ciencia no investiga otros fenómenos del mundo tal como estos se presentan en el mundo de la experiencia cotidiana. [...] La conclusión adecuada no es que debemos abandonar los conceptos del lenguaje que pueden ser estudiados productivamente, sino que el tema de la comunicación exitosa en el mundo actual de la experiencia es demasiado complejo y oscuro para ameritar atención de la investigación empírica [...] (pp. 69-70).

Sin duda, el tema de la comunicación exitosa será demasiado amplio y oscuro para quien está interesado en la pregunta de qué mecanismos en la mente/cerebro son responsables de lo que el niño parece saber, como sostiene Chomsky, «de manera anticipada a la experiencia». El punto es que nada dice que *esa* sea la pregunta a responder cuando se trata de explicar el hecho de que uno es, por lo general, entendido de la manera en que uno espera ser entendido. En este último caso, lo que se necesita es una teoría que dé cuenta de aquello que un intérprete logra a partir de la única evidencia de que puede disponer: el uso de las oraciones, de parte del hablante, en situaciones públicas reconocibles. Ahora vemos por qué no agrega nada a esta concepción decir que si la teoría describe correctamente la habilidad del intérprete, la teoría debe corresponder a algún mecanismo en la mente/cerebro del intérprete. Y es que la pregunta relevante aquí es por aquello que hace que una teoría sea interpretativa —es decir, que arroje interpretaciones correctas de lo que el hablante dice—, independientemente de cuál sea el mecanismo subpersonal involucrado en esta operación.

Al parecer, Chomsky confunde lo que en realidad son dos preocupaciones distintas: de un lado, la pregunta de cómo describimos o caracterizamos, mediante una teoría recursiva de algún tipo, la habilidad del intérprete para comprender a los demás; y, de otro lado, la pregunta de qué propiedades innatas de la mente/cerebro están a la base, en tanto hacen posible, el saber que los hablantes poseen⁹. La segunda cuestión es independiente de la primera dado que, como hemos visto, la comprensión no requiere que los hablantes compartan las mismas propiedades innatas. En ese sentido, Chomsky podría estar en lo correcto al afirmar que el lenguaje puede ser estudiado a la manera de un órgano del cuerpo, mas no en que la comprensión que

⁹ McDowell (1994) distingue, en un sentido análogo, entre lo que hace posible y lo que constituye los fenómenos a nivel personal, entre relaciones facultativas y constitutivas, respectivamente.

tenemos de los demás, vale decir, la cuestión de cómo accedemos a sus pensamientos para determinar el sentido de sus oraciones, se siga de los estados alcanzados por los *I-language* de los interlocutores.

Por otro lado, sostener que el tema de la comunicación interpersonal no tiene cabida dentro de una investigación naturalista es aferrarse a una noción innecesariamente estrecha de lo que significa ser naturalista. Aquí es donde podemos apreciar las diferencias entre los dos modelos de naturalismo que señalamos al inicio, el reduccionista y el no reduccionista.

El primero considera que una adecuada explicación de los hechos o patrones de conducta de nivel personal implica tratar esos hechos como idénticos a, o el resultado directo de, estados y procesos subpersonales —en otras palabras, como hechos que versan sobre la relación entre partes de la persona: hechos relativos al diseño de la mente o la actividad del cerebro—. Esta sería una manera de garantizar la continuidad entre el nivel personal y subpersonal reduciendo los hechos de nivel más alto a los de nivel más bajo. A contracorriente de esta posición, el modelo no reduccionista considera que los hechos de nivel personal —por ejemplo, cómo el hablante espera ser entendido en cierta ocasión— están a la vista solo con relación a la concepción que tenemos de los individuos como agentes o seres racionales, poseedores de estados mentales, concepción que emerge en el contexto que suministra la interacción comunicativa con otros. Así, al intentar explicar reductivamente los hechos de nivel personal sobre la base de hechos de nivel subpersonal, observamos que el paso de un nivel a otro conlleva cambiar de tema, no su esclarecimiento; esto debido a que los hechos de nivel personal están ligados *constitutivamente* a la concepción que tenemos de los demás como seres racionales. Que la relación sea constitutiva quiere decir, en buena cuenta, que allí donde carecemos de la concepción adecuada no podemos identificar los hechos que logramos describir con su ayuda. Desde luego, lo dicho hasta aquí no quiere decir que no pueda haber conexiones interesantes, de tipo causal, entre los hechos de nivel personal y los de nivel subpersonal. Lo que el naturalista no reduccionista niega es que sea forzoso hallar dichas correlaciones para dar cuenta de los hechos de nivel personal. Como veremos en el próximo apartado, esto quiere decir que aquello que explica los hechos en cuestión ha de remitirnos por necesidad al punto de vista dentro del cual esos hechos son constituidos.

Hasta aquí, hemos visto que no hay razón alguna para sostener que el estudio empírico de la mente y el lenguaje deba prescindir de aquellas nociones que Chomsky considera demasiado generales y oscuras, ni que sea acientífico emplearlas, como hace el modelo de la interpretación radical, para iluminar la comprensión y competencia lingüísticas.

3. TEORÍAS AL PASO Y ATRIBUCIÓN PSICOLÓGICA: LA PERSPECTIVA DE SEGUNDA PERSONA

Una consecuencia que se desprende del argumento de Davidson en contra de la idea de que la comprensión lingüística requiere que tanto intérprete como hablante posean de antemano un saber lingüístico similar —que conozcan las mismas convenciones y regularidades del idioma; que posean un método de interpretación común o una gramática interna universal— es que no hay atribución de estados mentales, ni de significados, con anterioridad a la comunicación interpersonal. Pues en lo que a la comprensión de los demás respecta, lo importante es poder determinar qué es lo que el hablante quiere que el intérprete entienda —que crea o haga— a través del reconocimiento de su intención, fundamentalmente: qué es lo que el hablante quiere que el intérprete sepa, y sepa que ambos saben, sobre sus estados mentales y la comunicación misma.

Desde luego, hay casos en que el lenguaje es usado para fines distintos de la comunicación —por ejemplo, para aclarar nuestros propios pensamientos—, así como casos en que atribuimos estados mentales a otros individuos sin estar en comunicación directa con ellos —al escuchar casualmente una conversación o leer una novela—. El punto no es que para todo uso del lenguaje se necesite estar en comunicación con otro, es más bien que la pregunta «¿qué significa o quiere decir X?», así sea acerca de lo que uno mismo quiere decir, nos obliga a identificar algún tipo de intención —semántica, comunicativa, etcétera—, búsqueda que es inseparable del *hecho general de la comunicación*, del punto de vista que asumimos cuando se está en comunicación con otros. En ese sentido, argumentaremos ahora que no hay atribuciones de estados mentales, ni de significados, independientemente de lo que se conoce como *la perspectiva de segunda persona*: el punto de vista que surge a partir de la relación entre dos o más individuos que interactúan simultáneamente entre sí y el mundo compartido —aquel que está la base del conocimiento *intersubjetivo* o de las otras mentes—¹⁰.

La segunda consecuencia, implícita en el párrafo anterior, es que una teoría al paso —la teoría que el intérprete usa para entender las oraciones del hablante— ha de servir, a su vez, como modelo para explicar la capacidad de *atribución psicológica* —también llamada capacidad de *mentalización* o *teoría de la mente*—, la capacidad para atribuir estados mentales a otros individuos, incluido uno mismo,

¹⁰ En los orígenes de esta perspectiva está el *otro generalizado* de Mead (1974), condición necesaria para la aparición del gesto y símbolo significantes, así como el modelo davidsoniano de la *triangulación* (Davidson, 2001c, pp. 205-220), el cual establece las condiciones mínimas para la aparición de pensamiento y lenguaje. El carácter comprometido, no distante, de este punto de vista ha sido desarrollado ejemplarmente por Strawson (2008[1974], pp. 1-28) con relación al caso de las actitudes reactivas personales —la ira, la culpa y la gratitud—.

con el fin de hacer inteligible su conducta. En ese sentido, mostraremos ahora que la teoría al paso es la teoría que nos permite leer la mente de los demás, en tanto la interacción recíproca establece las condiciones mínimas para empezar a atribuir estados mentales a otros, tener conocimiento del mundo y autoridad sobre nuestros propios pensamientos¹¹. En suma, veremos que la atribución psicológica depende en gran medida de la comunicación porque la teoría que el intérprete desarrolla para hacer inteligible la conducta del hablante emerge sobre el trasfondo de una visión de la realidad que es compartida, visión que es indispensable para pensar en los otros como seres *racionales*: seres cuya conducta puede ser explicada y predicha sobre la base de estados mentales. Con ello se espera mostrar que lo que posibilita hablar de hechos a nivel personal —de patrones de conducta característicamente racionales— no es tanto la existencia de hechos subpersonales concomitantes como el punto de vista que emerge en la interacción simultánea con otros individuos y la realidad compartida, y que las causas que son explicativas de los hechos de nivel personal deben remitir al punto de vista bajo el cual estos hechos son constituidos: el punto de vista que está a la base de la concepción que se tiene de las personas como seres racionales.

Un primer aspecto a tener en cuenta es la estrecha relación entre la atribución de estados de creencia, deseo e intención, de un lado, y, de otro, el concepto de verdad objetiva: la diferencia entre lo que se piensa y lo que es el caso. Una manera de trazar esta relación es pensando en la teoría al paso como la teoría que el intérprete usará para empezar a atribuir estados mentales al hablante sobre la base de una visión compartida de la realidad, o lo que es lo mismo, a partir de una manera común de clasificar las cosas —en términos de acierto y error—. En efecto, podemos pensar en la teoría al paso como la teoría que el intérprete desarrolla para dar sentido a la idea de que tanto él como el hablante comparten una visión de la realidad; de eventos y objetos que tienen lugar en una realidad que es independiente de lo que cada uno piensa sobre ella. Así, no comprender a otro es no ser capaz de dar sentido a su conducta sobre la base de una teoría semejante, es fallar en reconocer que cada uno responde a objetos y eventos situados en el mundo compartido. Pues a medida que el intérprete va dando sentido a la conducta del hablante atribuyéndole pensamientos similares o distintos a los de él, va descubriendo también que la visión de las cosas que tanto él como el hablante poseen es una que ambos comparten, y saben que comparten, en tanto juzgan que hay similitud en su manera de clasificar las cosas. Veamos esto a continuación.

¹¹ Básicamente, a partir de este punto nuestra tarea consiste en dar relleno filosófico a la siguiente afirmación: «cuando usamos los conceptos de creencia, deseo y demás, debemos estar preparados, conforme la evidencia se acumula, para ajustar nuestra teoría a la luz de consideraciones de consistencia total: el ideal constitutivo de la racionalidad controla parcialmente cada fase en la evolución de lo que debe ser *una teoría en desarrollo*» (Davidson, 2001a, p. 223, las cursivas son mías).

¿Por qué se necesita descubrir que cierta visión de las cosas es compartida con otros para atribuirles pensamientos? Básicamente, porque ciertos estados mentales —aquellos con contenido proposicional— requieren poder diferenciar, como mínimo, entre lo que se piensa del mundo y lo que es el caso, entre el acierto y el error. Dado que esta es una diferencia que el individuo no establece por sí solo, reconocer que dicha diferencia existe independientemente de uno implica ser capaz de reconocer cuándo otros la poseen y aplican. Precisamente, es esto lo que el intérprete debe poder decir del hablante para atribuirle estados mentales, que es alguien que clasifica las cosas o distingue entre el acierto y error. Sin embargo, el intérprete podrá decir esto únicamente a partir del modo en que él mismo distingue entre el acierto y el error; en otras palabras, ha de partir de sus propios estados mentales, de lo que él mismo piensa, para dar sentido a la idea de que el hablante también clasifica las cosas y tiene, por tanto, pensamientos acerca del mundo. De ahí que si el intérprete descubre que el hablante tiene pensamientos, entonces el intérprete habrá descubierto que su visión de las cosas es similar a la del hablante por ser el hablante alguien que clasifica las cosas de manera similar a como lo hace el intérprete. Nótese que la noción de similitud en que intérprete y hablante se apoyan para dar sentido a la idea de que ambos clasifican los mismos objetos de la realidad es una que se deriva del *juicio* que cada uno tiene acerca de su propia manera de clasificar las cosas y la manera en que el otro las clasifica. En otras palabras, para dar sentido a la idea de que ambos individuos comparten una visión de las cosas, es necesario que cada uno *perciba* la similitud en cuestión, que ambos sepan que su visión de las cosas es una que *ambos* comparten. En definitiva, es necesario que cada individuo comparta, y sepa que comparte con alguien más —al percibir como similar—, una visión de las cosas. Recapitulando:

- a) Tener estados mentales requiere poder diferenciar entre el acierto y el error.
- b) Para atribuirte estados mentales, debo poder señalar de qué manera tú distingues entre el acierto y error.
- c) Solo podré señalar de qué manera tú distingues entre el acierto y el error si parto de mi propio caso, de la similitud que encuentro entre tu manera de clasificar las cosas y la mía.
- d) Para decir que tu manera de clasificar las cosas es similar a la mía, no basta con que yo diga que es similar, tiene también que ser similar para ti: tú debes poder decir que mi manera de clasificar las cosas es similar a la manera en que tú las clasificas.
- e) Para atribuirte estados mentales a ti, tú y yo debemos compartir, y saber que ambos compartimos, una manera de clasificar las cosas.

Para ilustrar este razonamiento, tomemos el caso de la creencia. Para empezar, el intérprete no podría atribuir una creencia al hablante sin saber, él mismo, bajo qué condiciones la creencia en cuestión es verdadera o falsa, pues tener una creencia requiere poder señalar, aunque sea de manera implícita, cuándo la creencia es errónea. Asimismo, el intérprete no podrá atribuir creencia alguna sin asumir que el otro sabe, al igual que él, bajo qué condiciones la creencia es verdadera o falsa. Desde luego, para descubrir esto el intérprete no tiene otro punto de referencia que no sea su propio caso. Así, a medida que el intérprete reconoce que el criterio que él usa es también el del hablante, se da cuenta de que su visión de las cosas es la visión que él comparte con el hablante, al ser este último alguien que encuentra su propia manera de clasificar las cosas similar a la del intérprete. Por tanto, la visión de la realidad que el intérprete tiene, y que él sabe que tiene, es la visión que él sabe que comparte con el hablante; una que el hablante sabe que tiene, y sabe a su vez que el intérprete tiene, en la medida que logran comprenderse mutuamente sobre la base de esa visión.

La idea entonces es que para entenderte a ti, para atribuirte estados mentales, tú y yo debemos compartir, y saber que compartimos, una visión de las cosas. La visión que tú y yo compartimos, y sabemos que compartimos, se sigue de reconocer que ambos clasificamos las cosas de manera similar —en tanto cada uno percibe la similitud en cuestión—. De acuerdo con el modelo epistémico propuesto por Davidson (2001c, pp. 205-220), lo que esta forma de interacción genera es un triángulo conceptual que articula tres variedades de conocimiento proposicional: acerca de uno mismo —subjetivo—, acerca de los demás —intersubjetivo— y acerca de una realidad que es compartida con los demás —objetivo—. Estas tres variedades de conocimiento aparecen de manera simultánea y son *interdependientes*, lo que significa que no puede haber una sin las otras dos. Así, para tener conocimiento de nuestra propia mente se necesita tener conocimiento de la mente de los demás y de la realidad que compartimos con ellos; para tener conocimiento de las otras mentes se necesita conocer nuestra propia mente y la realidad compartida; finalmente, para tener conocimiento de la realidad compartida es necesario conocer nuestra propia mente y la mente de aquellos con quienes nos comunicamos.

Ahora bien, una particularidad del conocimiento intersubjetivo, o de las otras mentes, es que nos remite de manera explícita y simultánea a los tres vértices del triángulo conceptual descrito por Davidson. En efecto, como intérprete no necesito saber cómo tú te representas el mundo para yo poder describirlo correctamente, ni tampoco saber cómo tú te representas mi mente para conocer mis propios estados mentales. Sin embargo, sí he de saber cómo es el mundo para ti y cómo tú piensas en tus propios estados mentales si puedo atribuirte estados de creencia, deseo e intención. Es esta particularidad del conocimiento intersubjetivo a la que nos referimos

cuando señalamos que la atribución de estados mentales es siempre de *segunda persona* —o que nos remite al punto de vista que emerge al interactuar con otros y el mundo—, pues este proceso requiere atribuir el concepto de verdad objetiva, y este es un concepto que no se tiene sin saber que se tiene y sin saber que es compartido con otros (2001c, p. 121). Nótese que la atribución de estados mentales es de segunda persona no porque sean estados que uno atribuye a alguien más, sino porque la comprensión que tenemos de los demás depende de compartir, y saber que compartimos con ellos, una manera de clasificar las cosas. En definitiva, no puedo decir que tú clasificas las cosas, y que posees tales o cuales pensamientos, a menos que la similitud en nuestros juicios sea manifiesta para ti, para mí y para *ambos*.

Esta concepción de la atribución psicológica ha sido criticada por quienes consideran que debe haber lo que Dennett (1989, pp. 42 y 294) llama un *hecho profundo* acerca de la atribución mental, un hecho psicológico profundo, enteramente independiente del intérprete, que muestre cuál es el estado mental que el hablante posee realmente. Así, por ejemplo, se ha argumentado que la perspectiva de segunda persona de la atribución psicológica implica que el hablante podría no saber, por lo general, en qué estado mental se encuentra si el reconocimiento de cuál es su estado mental requiere que esté en comunicación con otro individuo. Según esta línea de pensamiento, bastaría que no esté en comunicación con alguien más para que el hablante no sepa, por lo general, qué es lo que dice o piensa. La intuición que motiva esta objeción es que el éxito al momento de atribuir estados mentales a otros depende de que el intérprete descubra algo que solo era conocido por el hablante, algo así como un objeto mental interno. Sin embargo, si la atribución consistiera en descubrir lo que el hablante esconde, entonces, para conocer los pensamientos de otros, tendríamos que estar en contacto con los mismos objetos internos con los que ellos están familiarizados, al punto de que podría ser imposible llegar a saber de qué pensamientos se trata.

Según la concepción alternativa que hemos desarrollado, atribuir estados mentales a otros individuos equivale a reconocer que hay oraciones verdaderas que describen su manera de pensar, oraciones verdaderas que les atribuyen estados de creencia, deseo, intención, entre otros. Las oraciones son verdaderas no porque existan objetos mentales internos a los que corresponden esas oraciones; son verdaderas en virtud del proceso por el cual descubrimos, como intérpretes, que la visión de las cosas que compartimos con el hablante es una que ambos sabemos que compartimos, en suma, verdaderas en virtud de una *norma interpersonal* que genera la comunicación. El hecho mental a capturar, si se puede llamar así, está, no en la cabeza del hablante, sino en la relación que se establece al interactuar simultáneamente con el hablante y el mundo. Desde luego, siempre habrá más de una manera correcta de registrar o

capturar el hecho en cuestión, así como hay diferentes maneras de indicar la temperatura o el peso de un objeto empleando diferentes sistemas de medida (Davidson, 2001c, pp. 214-215). La diferencia es que en el caso de las oraciones que capturan el modo de pensar de los demás, no tenemos otro estándar para juzgar la corrección o verdad de las oraciones que atribuyen estados mentales que no sea un *estándar interpersonal*. Que sea interpersonal quiere decir que el criterio de verdad y falsedad al que apelamos en este caso no antecede a la comunicación, pues debe tomar en cuenta tanto lo que el intérprete encuentra razonable, lo que hablante encuentra razonable, así como lo que es razonable según la visión de las cosas que intérprete y hablante comparten. Las oraciones que usamos para reportar la temperatura o el peso de un objeto no dependen de un estándar semejante. Son verdaderas o falsas independientemente de lo que cualquiera piense o diga. En ese sentido, la diferencia entre hablar de los objetos físicos y de estados psicológicos pasa por reconocer que el estándar normativo que utilizamos en el segundo caso es constitutivo de los estados y hechos en cuestión. En el caso de los estados mentales, dicho estándar es provisto por la *teoría al paso*, de ahí su importancia para el estudio de la atribución psicológica¹².

Si lo dicho hasta aquí es correcto, entonces el problema se presenta no para quienes consideran que la atribución psicológica es de segunda persona, sino para quienes consideran que debe haber un hecho mental profundo, independiente del proceso de interpretación, que muestre en qué estado mental se encuentra realmente el hablante. La intuición detrás de esta manera de pensar en la atribución psicológica es que solo un hecho semejante podría garantizar la autoridad que cada quien tiene sobre sus propios pensamientos: el hecho de que uno sabe, por lo general, lo que uno piensa de un modo en que otros no lo pueden saber. Sin embargo, podemos ver que los problemas con esta concepción sobrepasan las soluciones. Por un lado, la existencia de un hecho mental independiente de toda interpretación, oculto en la cabeza del hablante, haría de la comunicación un accidente feliz, por no decir algo imposible. Este problema se podría resolver aduciendo que el hecho mental a capturar es aquel que el hablante establece por sí solo, pero esta solución no nos lleva muy lejos, pues, para empezar, supondría que cada hablante es infalible respecto a sus propios estados mentales —lo cual es falso—; y, en segundo lugar, haría que la identidad de los pensamientos recaiga completamente en la subjetividad del hablante, lo que haría

¹² Un ejemplo que ilustra esta importancia es el fenómeno de la irracionalidad. Una creencia es irracional no porque resulte extraña o poco común, sino porque resulta inconsistente con otras cosas que el agente cree o debería creer. El descubrimiento, de parte del intérprete, de qué cuenta como una creencia razonable y qué cuenta como una creencia irracional lo obliga, de entrada, a triangular con el hablante y el mundo para determinar a qué categoría pertenece la creencia en cuestión. El holismo en este caso es inevitable, pues la misma creencia podría pertenecer a una categoría distinta en otro contexto.

imposible en este caso hablar de pensamientos dado que la noción se trivializaría por completo. Según hemos visto, la condición para atribuir estados mentales a otros individuos es pensar en ellos como seres que poseen conocimiento de sus propios estados mentales —es decir, autoridad de la primera persona—, así como conocimiento de los estados mentales ajenos y de la realidad, partiendo siempre de una manera de clasificar las cosas cuya similitud es manifiesta para dos o más individuos. Como vemos, la estrategia aquí es de todo o nada.

¿Quiere decir esto que uno no sabe lo que piensa o quiere decir hasta que es interpretado por alguien más? La pregunta «¿qué es lo que yo pienso o quiero decir?» no es una que, por lo general, se me plantee a mí mismo. El hecho de que, por lo general, no se me plantee a mí mismo es una condición esencial para que yo pueda ser interpretado, para que otros me puedan atribuir pensamientos. En suma, para saber cuáles son mis propios pensamientos no es necesario que yo sea interpretado, basta con que sea *interpretable*. El hecho de ser interpretable implica, desde luego, que otros son intérpretes de mí y que yo soy intérprete de ellos. Así, en los casos en que la pregunta sí se me plantea, por ejemplo, cuando trato de aclarar mis propios pensamientos, puedo tratar de responderla tal y como haría en el caso de alguien más, es decir, pensando en mí *como si* fuera alguien más —nótese el *proviso*—. Puesto en los términos de Dennett (1989), puedo utilizar la actitud intencional para explicar y predecir el comportamiento de otros y también para explicar y predecir mi propio comportamiento. Dado que eso es cuanto está implicado en la noción de *agencia* —la idea de un ser cuya conducta cobra sentido a la luz de dicha actitud—, la autoridad que cada quien tiene sobre sus propios pensamientos no solo permanece intacta según la perspectiva de segunda persona de la atribución psicológica; también es explicada por esta perspectiva¹³.

¹³ A nivel empírico, dicha perspectiva se ve reforzada por los estudios que subrayan el rol de la interacción social en el desarrollo de la comprensión que el infante tiene de las otras mentes y del mundo (Carpendale & Lewis, 2004). De acuerdo con esta posición, la comprensión que el niño tiene de las otras mentes y del mundo es «construida» en la interacción social a medida que el infante aprende a relacionarse con otros de manera triádica, distinguiendo entre sí mismo, los demás y el mundo de objetos físicos que comparte con los demás. El triángulo epistémico que genera este tipo de interacción es uno que se forma a partir de la *atención conjunta*, cuando el infante y adulto reconocen que sus puntos de vista confluyen sobre los mismos objetos del entorno. La evidencia apunta, en ese sentido, a que la capacidad de leer la mente de los demás y comprender el mundo social que nos rodea son características propias de aquellos individuos capaces de triangular entre sí y el mundo compartido, lo cual depende, ciertamente, de que la comunicación entre estos devenga intencional en algún punto. Ahora bien, una objeción a esta tesis es que la evidencia podría servir para sacar la conclusión opuesta, a saber, que la comunicación se vuelve intencional porque los estados mentales son complejos desde el inicio y no complejos en virtud del tipo de comunicación que los individuos logran a partir de la atención conjunta. La objeción no da en el clavo debido a que, como hemos visto, hablar de estados mentales

Estos puntos permiten ver con más claridad en qué consiste la autonomía explicativa del nivel personal, aquel en el cual apelamos a creencias y deseos —entre otros estados mentales— para explicar o hacer inteligible la conducta del agente. Las creencias y deseos explican una acción al mostrar qué fue lo que el agente vio de deseable en la acción, esto es, al *racionalizar* la acción del agente —se trata de un caso en que explicar algo es hallarlo racional—. La caracterización que uno ofrece de la conducta del agente en términos de creencias y deseos es una que está atada, y no por mera casualidad, a una visión de la realidad que es manifiesta tanto para el agente como para el intérprete. En el caso de la explicación psicológica de sentido común, esto quiere decir que las razones del agente explican su acción al mostrar cuál era la perspectiva del mundo, y de sí mismo, que tuvo el agente al momento de actuar: qué creía el agente o vio de deseable en la acción que llevó a cabo. Ahora bien, si lo que explica la acción del agente no son las razones que este tuvo al momento de actuar sino causas subpersonales que escapan a su reconocimiento —como sería el caso de los estados computacionales o neurofisiológicos que subyacen a la cognición y otras operaciones mentales—, entonces lo explicado deja de ser la acción de un agente, pues no estaríamos frente a un caso en que tratamos de dar cuenta de algo hecho intencionalmente —la definición estándar de *acción*— sino de algo que ocurre como resultado de fuerzas que carecen por completo de intencionalidad. Al prescindir de las motivaciones conscientes o inconscientes del agente, es la acción del agente lo que desaparece, lo que en buena cuenta significa no pensar en los demás como agentes o seres racionales.

En ese sentido, podemos decir que la autonomía del nivel personal se deriva del carácter esencialmente intersubjetivo, o de segunda persona, de la atribución y explicación psicológicas. Puesto en otros términos, podemos decir que lo que determina la existencia de hechos a nivel personal —lo que nos autoriza a hablar, a este nivel, de patrones de conducta característicamente *racionales*— es el punto de vista que emerge al interactuar simultáneamente con otros individuos y la realidad compartida, no la ocurrencia de hechos subpersonales concomitantes. La importancia de este punto de vista no puede ser minimizada si, como creemos, la perspectiva de segunda persona es *constitutiva* de los hechos que describimos y explicamos con ayuda del vocabulario de la psicología *folk* o de sentido común.

complejos, como las actitudes proposicionales, requiere dar sentido a la idea de que dos criaturas responden a los mismos objetos o eventos en el mundo. El reconocimiento de que hay similitud en sus respuestas o que clasifican las cosas de la misma manera es un rasgo que depende enteramente de sus juicios, de lo que cada uno encuentra similar. Es sobre la base de dicha similitud percibida que podemos empezar a atribuir pensamientos a otros con el fin de hacer inteligible su conducta.

Podemos ver ahora por qué la pretensión del naturalismo reduccionista consistente en hallar relaciones verticales, de dependencia explicativa, entre los niveles personal y subpersonal —relaciones que iluminen las propiedades de la mente y el lenguaje— aparece como infundada. La posibilidad de hallar relaciones verticales explicativas interniveles supone que podemos correlacionar, uno a uno, los hechos de nivel personal con los de nivel subpersonal. Pero si la perspectiva que asumimos al describir un hecho a nivel personal es constitutiva del hecho descrito, entonces no hay posibilidad de esclarecer o explicar ese hecho con referencia a algo —algún mecanismo o proceso subpersonal— que no está ligado esencialmente a una visión de la realidad que es manifiesta para dos individuos como mínimo. Nuestro hablar de hechos a nivel subpersonal toma su fuerza no de esa visión, sino del punto de vista que no incluye referencia alguna a los individuos que la poseen: el punto de vista que versa sobre la relación entre diferentes partes de la persona y que representa los fenómenos como instancias del modo en que las cosas suceden según leyes. De modo que no puede haber relaciones verticales explicativas allí donde los hechos a correlacionar responden a perspectivas que son, además de distintas, *distintivas*, perspectivas que representan variedades no reductibles de conocimiento —como son el conocimiento subjetivo, intersubjetivo y objetivo—. El carácter constitutivo de la perspectiva de segunda persona y del ideal de racionalidad que va aparejado a ella determinan, pues, la ausencia de continuidad explicativa entre los niveles personal y subpersonal.

4. CONCLUSIÓN

Hemos abordado el papel de la intersubjetividad —específicamente, el papel que desempeña la comunicación interpersonal— en la atribución de pensamiento y significado, es decir, su importancia para el tipo de explicación que requiere atribuir a otros individuos estados mentales con contenido. La importancia de la comunicación en este proceso reside, básicamente, en que una adecuada caracterización de la conducta verbal y psicológica del agente conlleva situar sus acciones sobre el trasfondo de una visión de la realidad que ha de ser manifiesta para dos o más individuos, condición que se cumple en la comunicación, al interactuar simultáneamente con otros individuos y la realidad compartida. La perspectiva que asumimos sobre los otros cuando explicamos sus acciones de esta manera es de segunda persona y, como tal, independiente del tipo de inteligibilidad que encontramos a nivel subpersonal. En ese sentido, se ha sostenido que algunas propiedades de la mente y el lenguaje —viz, la comprensión lingüística y la atribución psicológica— no pueden ser iluminadas en términos de relaciones verticales entre los niveles personal y subpersonal; antes bien, su estudio está circunscrito, fundamentalmente, al nivel en el cual pensamos en el ser humano como un agente o ser racional.

REFERENCIAS BIBLIOGRÁFICAS

- Bermúdez, José Luis (2005). *Philosophy of Psychology: A Contemporary Introduction*. Londres: Routledge.
- Carpendale, Jeremy I. M. & Charlie Lewis (2004). Constructing an Understanding of Mind: The Development of Children's Social Understanding within Social Interaction. *Behavioral and Brain Sciences*, 27, 79-151.
- Chomsky, Noam (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, Noam (2001). Language and Problems of Knowledge. En Aloysius Martinich (ed.), *The Philosophy of Language*. Tercera edición (pp. 581-599). Oxford: Oxford University Press.
- Davidson, Donald (1986). A Nice Derangement of Epitaphs. En Ernest LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (pp. 433-458). Oxford: Blackwell.
- Davidson, Donald (2001a). *Essays on Actions and Events*. Segunda edición. Oxford: Clarendon Press.
- Davidson, Donald (2001b). *Inquiries into Truth and Interpretation*. Segunda edición. Oxford: Clarendon Press.
- Davidson, Donald (2001c). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- De Caro, Mario & Alberto Voltolini (2010). Is Liberal Naturalism Possible? En Mario de Caro y David MacArthur (eds.), *Naturalism and Normativity* (pp. 69-86). Nueva York: Columbia University Press.
- Dennett, Daniel (1969). *Content and Consciousness*. Londres: Routledge.
- Dennett, Daniel (1989). *The Intentional Stance*. Cambridge: The MIT Press.
- Grice, Paul (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.
- Hornsby, Jennifer (2000). Personal and Sub-personal; A Defence of Dennett's Early Distinction. *Philosophical Explorations*, 3(1), 6-24.
- McDowell, John (1994). The Content of Perceptual Experience. *The Philosophical Quarterly*, 44(175), 190-205.
- McDowell, John (1996). *Mind and World*. Cambridge: Harvard University Press.
- Mead, George Herbert (1974). *Mind, Self, and Society. From the Standpoint of a Social Behaviorist*. Chicago: The University of Chicago Press.
- Sperber, Dan & Deirdre Wilson (1986). *Relevance: Communication and Cognition*. Cambridge: Harvard University Press.

- Strawson, Peter (1985). *Skepticism and Naturalism: Some Varieties*. Nueva York: Columbia University Press.
- Strawson, Peter (2008[1974]). *Freedom and Resentment and Other Essays*. Nueva York: Routledge.
- Tomasello, Michael (2008). *Origins of Human Communication*. Cambridge: The MIT Press.

LA SISTEMATICIDAD DE LA COGNICIÓN HUMANA EN CUESTIÓN¹

Antoni Gomila

Universidad de las Islas Baleares

1. INTRODUCCIÓN

El argumento de la sistematicidad puede considerarse como el argumento más importante a favor de la arquitectura cognitiva clásica: la idea de que la cognición consiste en procesos inferenciales de estados mentales simbólicos (Fodor, 1975 y 1987, Fodor & Pylyshyn, 1988). El argumento puede resumirse así: 1) la cognición es sistemática; 2) la única/la mejor explicación posible de la sistematicidad de la cognición es una arquitectura cognitiva clásica, de reglas y representaciones. Por tanto, los modelos conexionistas o bien son implementaciones de un modelo clásico (si respetan la sistematicidad) o bien son incapaces de explicar la sistematicidad de la cognición. A lo largo de los años, los conexionistas han tratado de resistir el argumento atacando la segunda premisa, es decir, tratando de mostrar que también es posible dar cuenta de la sistematicidad de la cognición, sin postular unas representaciones simbólicas que se combinen formalmente, sino a partir de sus redes de nodos interconectados. El debate subsiguiente se ha centrado sobre todo en la cuestión de cómo concebir la estructura constituyente de la cognición, que todo el mundo parecía considerar requerida para dar cuenta de la sistematicidad.

En este trabajo, queremos adoptar una estrategia distinta: vamos a rechazar la primera premisa, es decir, vamos a defender que la cognición no es sistemática en general. Ciertamente la cognición humana, en algunos casos, exhibe cierta sistematicidad, pero ese no es una propiedad general de la cognición. Si estamos en lo cierto,

¹ Este trabajo ha recibido el apoyo del Ministerio español de Ciencia y Competitividad, a través del proyecto FFI2009-13416-C02 y de la Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia-II PCTRM 2007-10 (proyecto 11944/PHCS/09). Fue presentado en el IV Coloquio «Lenguaje, Cognición y Contexto», celebrado en Písaq, del 18 al 20 de julio de 2012. Agradezco a los participantes en el workshop sus preguntas y discusiones y a los organizadores la posibilidad de formar parte de este proyecto.

entonces ni el cognitivismo clásico ni el conexionista proporcionan la mejor explicación de la cognición, aunque pueden tener interés localmente. Los enfoques corpóreos de la cognición, como la teoría de sistemas dinámicos y la percepción ecológica, ofrecen un marco más prometedor para comprender la cognición «desde abajo» (Gomila & Calvo, 2008; Spivey, 2007), al ofrecer un modelo de los fundamentos sensoriomotores e interactivos de la cognición (Kello y otros, 2010). Desde este punto de vista, la cuestión de la sistematicidad deviene la cuestión de cómo es posible que la cognición pueda convertirse ocasionalmente en sistemática, cuando lo hace, a partir de un nivel básico de funcionamiento que es más bien no sistemático, no composicional, interactivo y complejo.

Para cuestionar entonces la primera premisa, la afirmación de que la cognición humana se caracteriza por su sistematicidad, comenzaremos, en la siguiente sección, por plantear cómo debe entenderse esta noción de sistematicidad, para pasar a continuación a presentar una serie de argumentos en contra de que se trate de una propiedad general de la cognición. En la sección «La sistematicidad depende del lenguaje» mostraremos que la sistematicidad cognitiva depende de la lingüística. Y en la siguiente sección «Los seres no verbales no muestran sistematicidad cognitiva» presentaremos la evidencia que indica que la habilidades cognitivas de los sistemas cognitivos no verbales no son sistemáticas.

Sobre la base de estas evidencias, concluiremos que la cognición no es sistemática en general. Sin embargo, reconoceremos que las mentes verbales son capaces de procesos sistemáticos (como el razonamiento o la solución de problemas). Esto obliga a plantear la cuestión de qué modo mentes no sistemáticas pueden llevar a cabo procesos sistemáticos; o más precisamente, pueden convertirse en sistemáticas, dado que los procesos sistemáticos característicos se desarrollan tarde en la infancia. Nuestra respuesta a esta cuestión será que las mentes devienen capaces de procesos sistemáticos por el hecho de convertirse en mentes verbales. Así como los ejemplos de sistematicidad cognitiva son parasitarios de la sintaxis, la idea es que adquiriendo tal sintaxis combinatoria llegamos a ser capaces de tal sistematicidad cognitiva. Según esta propuesta, el enfoque del «lenguaje del pensamiento» de Fodor invierte el orden de dependencia entre el lenguaje y el pensamiento cuando trata de explicar las isomorfías entre ambos. El cognitivismo clasicista, de hecho, ilustró la noción de sistematicidad cognitiva con ejemplos lingüísticos para afirmar que del mismo modo que la explicación en el caso del lenguaje es una sintaxis y semántica combinatoria, algo parecido debe ocurrir en el caso del pensamiento, que concebían como más básico que el lenguaje. En nuestra opinión, la adquisición del lenguaje es la clave para la aparición de la sistematicidad cognitiva, y no al contrario.

En resumen, en lugar de postular una arquitectura de símbolos y reglas formales, como el nivel básico de explicación de la cognición, en nuestra opinión el nivel básico es dinámico e interactivo, y gracias a la adquisición del lenguaje, con su sintaxis combinatoria, aparecen procesos sistemáticos. Esta propuesta retoma el enfoque sugerido por Dennett (1993), pero puede ser vista actualmente como una versión de la teoría de procesos duales de la cognición (Spelke, 2003; Carruthers, 2005; Evans & Frankish, 2009) —una versión que postula que el nivel básico es dinamicista, corpóreo, interactivo—. En este nivel básico no vamos a encontrar estados discretos, amodales y suficientemente abstractos para poder ser llamados «simbólicos», y mucho menos que se correspondan con significados lingüísticos. El lenguaje, en nuestra opinión, reestructura este nivel básico, haciendo posible entonces procesos sistemáticos, entre otras propiedades.

2. EL ARGUMENTO DE LA SISTEMATICIDAD

En «El lenguaje del pensamiento» (Fodor, 1975), el argumento de la sistematicidad jugó un papel central en la propuesta fodoriana a favor de un «lenguaje del pensamiento (de aquí en adelante, LOT, por sus siglas en inglés) y una arquitectura cognitiva clásica, según la cual los procesos cognitivos son transiciones inferenciales de estructuras simbólicas según reglas formales. No fue el único argumento que utilizó este autor, pero sin duda es un argumento más asumible que su otro argumento central, basado en un fuerte innatismo conceptual. En su «Why there still has to be a language of thought?» (1987), ambos argumentos fueron mejor formulados. En el trabajo conjunto con Pylyshyn (1988), el argumento de la sistematicidad se convirtió en la piedra de toque contra los, entonces recientes, modelos conexionistas. Generó una enorme controversia y estimuló tanto el desarrollo del conexionismo como de su interpretación. Ahora que está teniendo lugar el desarrollo enérgico de una serie de enfoques alternativos al cognitivismo —sea clásico o conexionista—, con su énfasis común en la corporeidad de la cognición y su naturaleza interactiva (Gomila & Calvo, 2008), y cuando se cumplen 25 años de la publicación del importante trabajo de Fodor y Pylyshyn, es importante reevaluar su argumento.

La estructura básica del argumento de la sistematicidad puede resumirse de este modo:

- (1) La cognición es sistemática.
- (2) La mejor (única) explicación de la sistematicidad es la arquitectura cognitiva de reglas formales y estructura simbólica composicional. Por tanto,
- (3) la mejor explicación de la cognición consiste en la estructura composicional.

En Fodor y Pylyshyn (1988), el argumento se usa para plantearle un dilema al conexionismo: o bien no puede explicar la cognición (si sus modelos no incluyen la estructura composicional), o bien no es más que una implementación de un modelo clásico (si la incluye). La reacción más extendida a este reto consistió en rechazar la premisa 2, al tiempo que se asumía como válida la 1. No es nuestro propósito volver a ese debate, sino más bien adoptar una estrategia distinta frente al argumento: la de rechazar la premisa 1. Para tal propósito, es necesario comenzar por profundizar en lo que se entiende por sistematicidad.

La noción de sistematicidad en cuestión nunca fue definida explícitamente, sino aclarada por medio de ejemplos lingüísticos. De hecho, en Fodor (1987), la noción de sistematicidad cognitiva es introducida en primer lugar por medio de ejemplos lingüísticos. La estrategia de su razonamiento puede reconstruirse de este modo:

- a) Hay una cierta propiedad de nuestras capacidades lingüísticas que se debe al hecho de que los lenguajes tienen una semántica combinatoria.
- b) El pensamiento también tiene esta misma propiedad. Por tanto,
- c) también el pensamiento debe tener una semántica combinatoria (p. 148).

En nuestra opinión, la premisa b) es incorrecta: a veces el pensamiento tiene esta propiedad de sistematicidad, pero no siempre ni en general: solo cuando depende del lenguaje (y no al revés, como pretende Fodor). Pero en este punto, la cuestión es de qué modo el lenguaje nos aclara la noción de sistematicidad en cuestión. La respuesta es clara: nuestro conocimiento lingüístico no consiste en un listado de expresiones con su correspondiente significado (como sí sería el caso cuando nos manejamos con un libro de frases para turistas), sino que existe una conexión intrínseca entre las diferentes oraciones que podemos entender o producir. Dicho al modo de Fodor, si uno comprende la frase «*John loves Mary*» entonces también entiende la frase «*Mary loves John*». Para Fodor, el pensamiento también es sistemático en este sentido en general, porque en su opinión el lenguaje expresa el pensamiento y no podría expresar un contenido semántico sistemático si el correspondiente pensamiento no lo fuera en primer lugar; además, rechaza la posibilidad de que el pensamiento solo sea sistemático en el caso de criaturas lingüísticas como nosotros.

Es importante darse cuenta de que es la sistematicidad conceptual-inferencial lo que importa: las conexiones intrínsecas entre los que puede ser pensado. Es posible formular nociones alternativas de sistematicidad, pero no son las que importan desde el punto de vista del argumento. Así, por ejemplo, se podría hablar de sistematicidad de diseño para señalar la conexión necesaria entre diferentes partes del cuerpo, como por ejemplo, que si un organismo tiene dos pulmones entonces también tendrá

dos riñones; o que si tiene cierto tipo de sensores, entonces tendrá cierto tipo de actuadores o viceversa («si camina, entonces puede ver», digamos). Claramente no es esta la noción que interesa a Fodor y Pylyshyn. Pero más dudas puede despertar una noción de sistematicidad conductual, en la cual la conexión sistemática se establece entre tipos de conductas (por ejemplo, si se tiene la conducta de huir, también se tiene la de luchar) o una noción de sistematicidad de habilidades que señale la conexión intrínseca entre poder hacer diversas cosas (por ejemplo, entre cantar y hablar o entre apuntar con el dedo y pulsar). Resulta evidente que la mejor explicación de estas interdependencias no es en ningún caso un modelo clásico de reglas y representaciones constituyentes, por lo que habrá que evitar la tentación de defender la sistematicidad cognitiva sobre la base de dependencias que no sean estrictamente conceptuales-inferenciales.

Para evitar el reproche frecuente de que, en el argumento de Fodor, no se establecía una distinción suficiente entre la propiedad de la sistematicidad, y su explicación en términos de estructura constituyente, McLaughlin ha tratado de proporcionar una caracterización independiente de la sistematicidad (2009). Según su propuesta, la sistematicidad conceptual puede reformularse en términos de conexiones legaliformes: si un sujeto puede pensar cierto pensamiento, entonces también puede pensar muchos otros. De nuevo, la idea básica es la interconexión entre contenidos, en oposición a lo que podría verse como una lista, limitada, de pensamientos independientes unos de otros. La novedad es que, en este caso, se evita apelar a ejemplos lingüísticos, y al hacerlo, se evita el supuesto implícito de la estructura composicional, clara en el caso del lenguaje. Para hablar de la sistematicidad del pensamiento, en resumen, hará falta poder establecer que existe una conexión necesaria entre los diferentes pensamientos que un sistema pueda manejar.

La ventaja de esta formulación es que no utiliza implícitamente la noción de estructura composicional en la caracterización de la sistematicidad, como suele reprocharse desde el lado conexionista. También proporciona una clara indicación de cómo poner a prueba empíricamente si un sujeto cognitivo manifiesta sistematicidad cognitiva: dada una habilidad conceptual, se trata de examinar si existen otras habilidades interconectadas. Por supuesto, pueden plantearse cuestiones sobre la individuación de las habilidades conceptuales, pero al menos proporciona un método con el que proceder, en lugar de simplemente dar por supuesto que tales interdependencias conceptuales existen. En las dos secciones siguientes, presentamos una serie de evidencias que convergen en una respuesta negativa a la cuestión de si la cognición es sistemática.

3. LA SISTEMATICIDAD DEPENDE DEL LENGUAJE

Como ya se mencionó, los ejemplos que utiliza Fodor para ilustrar la noción de sistematicidad cognitiva son todos lingüísticos. Su ejemplo más característico es «*brown cow*» (vaca marrón): para ser capaz de pensar pensamientos sobre «*brown cow*», un pensador debe ser capaz de pensamientos relativos a «*brown*» y a «*cow*». Otro ejemplo recurrente, ya citado, es «*John loves Mary*» - «*Mary loves John*». En esta sección sostendremos que su discusión de estos ejemplos pasa por alto un aspecto muy importante de ellos, y es el hecho de que en estos casos la sistematicidad cognitiva está guiada, o es relativa, a la estructura sintáctica de las expresiones correspondientes.

En efecto, Fodor parece asumir que los patrones de sistematicidad que encuentra en sus ejemplos lingüísticos se reproducen de modo equivalente en el nivel conceptual: entender el pensamiento «que Juan quiere a María» implica entender también el pensamiento «que María quiere a Juan», sin darse cuenta de que la caracterización de tales contenidos implica relaciones sintácticas (Hinzen, 2012). El modo de demostrarlo consiste en considerar la multiplicidad de combinaciones conceptuales que deberían ser posibles, si el modo de combinación de las unidades conceptuales fuera independiente del lenguaje, y contrastarlo con las que efectivamente se dan.

Así, por ejemplo, si las combinaciones conceptuales no fueran restringidas por las reglas sintácticas, un hablante de inglés debería poder pensar el concepto complejo «vaca marrón», tanto como «marrón vaca», pero ocurre que no, que solo se dan las combinaciones conceptuales sintácticamente permitidas: la primera en el caso de los hablantes de español, la segunda en el de los hablantes de inglés. Del mismo modo, debería darse una relación sistemática entre «*John loves Mary*» y «*Mary is loved by John*», ya que la estructura conceptual que expresan ambas oraciones es la misma, si la hipótesis de Fodor fuera cierta, pero lo que encontramos es que esa interrelación solo aparece a partir de los cinco años de edad (Karmiloff & Karmiloff-Smith, 2002), cuando los niños aprenden la estructura de pasivo; antes de esa edad entienden «*Mary is loved by John*» como sinónima de «*Mary loves John*», y no de «*John loves Mary*», lo que sugiere, de nuevo, que la representación conceptual está estructurada sintácticamente.

Estos ejemplos apuntan a que las interdependencias conceptuales-inferenciales están estructuradas por las relaciones sintácticas del lenguaje del hablante, que las limitan y las estructuran. Algo parecido puede decirse del fenómeno de la opacidad (esa propiedad semántica que se da cuando falla la substitución *salva-veritate* de expresiones correferenciales). Así, es posible entender «Juan canta» y «Mary piensa», sin entender «María piensa que Juan canta» cuando Juan no está cantando: el dominio de la atribución de creencias falsas no ocurre hasta los cuatro años, dos años

después de la comprensión de las oraciones sintácticas activas simples. De hecho, hay abundante evidencia de que la adquisición de la estructura del complemento sintáctico, de la oración subordinada, es la clave que abre la posibilidad de la atribución de creencias falsas (De Villiers & De Villiers, 2009), lo que apunta de nuevo al papel de la sintaxis en la estructuración de los contenidos conceptuales del pensamiento.

Por otra parte, la combinación conceptual es característicamente sensible al contexto, lo cual también pone en cuestión el modo fodoriano de entender la sistematicidad de ejemplos como «vaca marrón». Según la caracterización de este autor, una vez que un pensador dispone de dos conceptos, como «vaca» y «marrón», puede combinarlos de modo algebraico: el significado de la combinación es la intersección de las referencias de cada uno de ellos. Sin embargo, no es esto lo que ocurre en realidad: abundan los ejemplos en los que el significado del compuesto no puede derivarse (algebraicamente) del significado de sus componentes. Así, entender la expresión «guitarra española» no se consigue haciendo la intersección del conjunto de las guitarras con el conjunto de las cosas españolas; ni una «taza de té» se puede entender combinando algebraicamente las cosas que son tazas con las cosas que son tés. El significado compuesto puede depender del contexto léxico, de las analogías elegidas por los hablantes a la hora de entender la proposición, o del tipo de significados involucrados, todo lo cual también depende, en último término, de las categorías sintácticas implicadas. Lo que no es cierto, en cualquier caso, es que, si alguien entiende el concepto X y el concepto Y, se sigue necesariamente que entiende la combinación XY, como requiere la sistematicidad (Estes & Gluksberg, 2000; Costello & Keane, 2000).

Además, la evidencia empírica también apoya la conclusión de que la sistematicidad cognitiva, en la medida que se da, está limitada o estructurada por el desarrollo lingüístico, porque el desarrollo léxico, el uso de etiquetas léxicas en la comunicación lingüística, modifica el modo en que son captados los conceptos que tales etiquetas expresan. Tales efectos han sido demostrados experimentalmente por los trabajos de Lupyan y colaboradores (2007, 2008a, 2008b, 2010). Según estos estudios, la adquisición de etiquetas léxicas juega un papel en el aprendizaje conceptual, haciéndolo más rápido; dando coherencia a los conceptos, reforzando los rasgos perceptivos correlacionados involucrados; haciéndolos más discretos y definidos, y por tanto, menos recordables sus ejemplares concretos; incluso en tareas de búsqueda visual, las etiquetas léxicas facilitan el procesamiento de estímulos familiares homogéneos, lo que se ha denominado el efecto de «agrupamiento» (*grouping*). El efecto es más significativo cuanto más perceptivamente variables sean los correspondientes conceptos, es decir, cuando menos semejantes en el nivel perceptivo sean los ejemplares miembros de la categoría. La etiqueta léxica induce considerar a los ejemplares como

más internamente semejantes, y más diferentes con los de otras categorías, que sin etiqueta. Esto requiere resaltar más unas propiedades e ignorar otras, según el concepto en cuestión; por ejemplo, el color no es importante para «silla», pero sí en el caso de «tomate». En general, cuanto más abstracto sea el concepto, más difícil resulta reconocer ejemplares concretos. Así, la categorización léxica conlleva el coste de mayores errores de reconocimiento de casos particulares, lo que indica en realidad, un cambio representacional, un modo distinto de codificar los estímulos percibidos.

Por último, la dependencia de la sistematicidad cognitiva del desarrollo lingüístico también se evidencia en el desarrollo sintáctico: mediante la integración progresiva de los conceptos adquiridos en una estructura predicativa. En un reciente estudio (Barceló, Corominas-Murtra & Gomila, 2012) analizamos el patrón de desarrollo sintáctico de tres niños hablantes de tres lenguas distintas, recogido en el Childes (<http://childes.psy.cmu.edu>), entre los veinte y treinta meses de edad, y encontramos un mismo patrón de «explosión sintáctica» en torno a los veintisiete meses. Antes de esa edad, sus preferencias son de dos o tres palabras, con muy pocas palabras apareciendo en diversidad de oraciones; a partir de esa edad, que parece indicar el momento de adquisición de las categorías sintácticas y sus reglas de combinación elementales, los niños comienzan a combinar de manera abierta las diferentes palabras adquiridas, lo que señala el inicio de la sistematicidad. Antes de ese momento no disponemos de evidencias independientes de sistematicidad cognitiva, lo que de nuevo apunta a la dependencia sintáctica de la sistematicidad cognitiva.

Así pues, la sistematicidad del pensamiento parece seguir la sistematicidad sintáctica, porque las propiedades de la primera reflejan las de la segunda, en lugar de al revés. La sistematicidad, por tanto, parece deberse más bien a la adquisición de un sistema recursivo de dependencias jerárquicas que adquirimos con el lenguaje.

4. LOS SERES NO VERBALES NO MUESTRAN SISTEMATICIDAD COGNITIVA

Dada la tendencia de Fodor a basarse en ejemplos lingüísticos, no está claro de entrada cómo proceder para considerar la cuestión de si las criaturas no verbales muestran o no sistematicidad cognitiva. Como apuntamos, McLaughlin nos ofrece un modo más útil a este respecto, para operacionalizar la noción en el caso de tareas no verbales, las cuales tendremos que valorar si la cognición es sistemática. Dada una habilidad conceptual, tenemos que considerar si hay otras habilidades que estén nómicamente conectadas con ella, de tal modo que no se pueda tener una sin tener la otra, o si más bien cabe verlas como independientes. Aun así, no es perfectamente claro cómo proceder, ni cuánta evidencia hace falta para establecer una conclusión: ¿basta un único ejemplo negativo para concluir que la cognición no verbal no es sistemática?

Si no, ¿cuántos hacen falta? ¿O cabría decir más bien que la carga de la prueba recae sobre quien sostiene que la cognición no verbal es sistemática?

Afortunadamente, no necesitamos entrar a fondo en estas cuestiones, porque el consenso en la psicología del desarrollo cognitivo y la psicología comparada, a nuestro modo de ver, constituye un aval a la conclusión de que los seres no verbales no muestran sistematicidad cognitiva. En el caso de la psicología del desarrollo cognitivo (por ejemplo, Spelke, 2003), el acuerdo más extendido afirma que, aunque los bebés muestran cierta comprensión conceptual de conceptos básicos (como objeto, espacio, agencia, numerosidad, causa, entre otros) en los primeros meses de vida, mucho antes de lo que Piaget sostuvo, esos conceptos no son combinados sistemáticamente ni desplegados flexiblemente en sus interacciones conductuales, hasta mucho más adelante —precisamente, tras producirse el desarrollo lingüístico oportuno—.

Ilustraremos esta concepción con un estudio, especialmente revelador, sobre la representación del espacio (Hermez-Vázquez, Spelke & Katnelson, 1999) que comparó la ejecución de niños prelingüísticos con adultos en una tarea de desorientación y reorientación espacial. Esta tarea, usada primeramente con ratas (Cheng, 1986), consiste en lo siguiente: tras familiarizar a los sujetos con un espacio rectangular cerrado, donde eran capaces de localizar la ubicación de una única fuente de comida en un rincón, se les sacaba y desorientaba. Algo más tarde, se les devolvía a la cámara rectangular y se registraba dónde buscaban el alimento en primer lugar: esta conducta permitió inferir qué información espacial memorizada usaban las ratas para reorientarse. Si estas se fijan en la forma de la cámara, por ejemplo, es decir, si buscan en el rincón con una cierta configuración de lado largo-lado corto, como tal configuración es ambigua, deben acabar el 50% de las veces en el rincón opuesto. Pero el rincón de la comida se especificaba unívocamente mediante un olor o el brillo de la pared corta. Si las ratas localizaban enseguida la fuente de comida, eso significaba que recordaban una de estas pistas unívocas. El resultado fue que las ratas buscaron igualmente en los rincones opuestos, a pesar de la presencia de información unívoca para resolver la tarea, lo que sugiere que su memoria espacial se basa en la geometría del espacio.

No es sorprendente que la reorientación se base en el recuerdo de pistas geométricas: la configuración espacial es duradera y raramente involucra simetrías engañosas. Otras pistas disponibles pueden ser transitorias y, por tanto, menos fiables: los olores, los colores, la disposición de objetos, etcétera. Las ratas no son la única especie que se basa en la geometría para orientarse, muchas otras especies hacen lo mismo (O'Keefe & Burgess, 1996), por ejemplo, los niños humanos. En un primer estudio, el grupo de Spelke adaptó la tarea de reorientación para usarla con niños de dieciocho a veinticuatro meses (Hermer-Vázquez & Spelke, 1994 y 1996). En lugar

de una fuente de alimentación, la esquina crítica escondía un juguete: los niños veían cómo se escondía, y después de desorientarlos (levantándolos y girándolos), tenían que buscarlo. En una condición la cámara era totalmente blanca, sin ninguna otra pista que la ambigua información geométrica. En una segunda, una pared era azul, lo que rompía la simetría y ofrecía una pista unívoca para la búsqueda. Pero los niños actuaron del mismo modo en las dos condiciones: como las ratas, buscaron el mismo número de veces en el rincón correcto y en el simétrico. También se usaron otras pistas para desambiguar el rincón donde el juguete estaba escondido, como la textura o un patrón dibujado sobre la superficie, sin que los niños las tuvieran en cuenta, a pesar de que todas estas pistas —color, textura, dibujo— son tenidas en cuenta por ellos en otras tareas. La conclusión que se desprende es que, aun cuando los niños puedan tener los conceptos geométricos y los de color o textura, no son capaces de combinarlos, de integrarlos en una única representación conceptual. En cambio, los adultos sí fueron capaces de identificar unívocamente el rincón en cuestión, gracias a tal integración conceptual (Hermer-Vázquez & Spelke, 1994).

Aún más importante, en un estudio posterior pudieron mostrar que esta capacidad de integración conceptual está relacionada con el desarrollo del lenguaje espacial. Es en torno a los cuatro años de edad que los niños comienzan a usar expresiones tales como «en el lado azul» o «detrás de la pared», mientras que los términos «izquierda» y «derecha» no aparecen hasta los seis. Por tanto, es de esperar que hasta los seis años no se pueda resolver unívocamente la tarea de reorientación si la pista desambiguadora es que se trata del «rincón de la izquierda». La demostración de este efecto se consiguió utilizando una tarea doble con adultos, a los que, además de la tarea de reorientación, se les pedía que realizaran simultáneamente una tarea verbal, a partir del razonamiento de que si el procesamiento lingüístico estaba comprometido con esta segunda tarea, y no podía contribuir a la de reorientación, la actuación de los adultos sería como la de los niños prelingüísticos. Y eso fue lo que ocurrió (Hermer-Vázquez, Spelke & Katnelson, 1999).

Ejemplos similares pueden encontrarse en la psicología comparada, con la excepción del efecto reestructurante del lenguaje. En general, los estudios de cognición primate indican que los primates no humanos son buenos en tareas perceptivas, pero no tan buenos en tareas cognitivas como el razonamiento (Gómez, 2004; Premack, 2004, Penn, Holyoak & Povinelli, 2008). Por el contrario, los únicos ejemplos de comprensión sistemática en estas especies consisten en casos de chimpancés a los que se les enseñó a usar símbolos. Por tanto, puede decirse que las mentes no verbales muestran habilidades cognitivas altamente especializadas y encapsuladas.

5. SOBRE LA PRIORIDAD DEL LENGUAJE

Hemos sostenido que la cognición humana no es sistemática en general, sino solo cuando el pensamiento está estructurado lingüísticamente. Esto sugiere la hipótesis complementaria de que la sistematicidad cognitiva es debida al lenguaje. Dado que el lenguaje no está disponible al nacer, se sugiere naturalmente la hipótesis de que las mentes humanas devienen sistemáticas al hacerse verbales. En otros términos, las mentes infantiles no son particularmente diferentes de las de otras mentes animales, excepto por su potencial para desarrollarse y por su predisposición para adquirir el lenguaje. Las mentes humanas devienen verbales en la comunidad social en la que se desarrollan y, de este modo, pueden manifestar la sistematicidad de sus procesos: las estructuras sintácticas que adquieren, estructuran sus pensamientos, y las etiquetas léxicas reconfiguran sus conceptos. Esta es la hipótesis que hemos defendido en el libro «Verbal minds: language and the architecture of cognition» (Gomila, 2012), sobre la base de una revisión exhaustiva de la nueva evidencia acumulada los últimos diez años. De este modo, proponemos invertir el orden de dependencia entre lenguaje y pensamiento: en lugar de tomar la sistematicidad del lenguaje como derivada de la del pensamiento, tal como propone Fodor, nuestra propuesta consiste en ver la sistematicidad del pensamiento, cuando se da, como derivada de la del lenguaje.

Ya hemos señalado que, de hecho, en el orden del desarrollo temporal, la sistematicidad lingüística precede a la cognitiva. En esta sección queremos considerar una objeción de principio a que se da también la dependencia indicada: el otro argumento de Fodor en favor del LOT, su nativismo extremo (1975). Según su planteamiento, es imposible adquirir un lenguaje sistemático si uno no es ya capaz previamente —de modo innato— de tal sistematicidad. Así, por ejemplo, en su crítica a una propuesta parecida de Dennett, alega:

Pero no se nos dice de qué modo una mente inicialmente asistemática podría aprender un lenguaje sistemático, dado que este último permite expresar proposiciones que esa mente no podría entender. De qué modo, por ejemplo, una mente que puede concebir «que Juan quiere a María» pero no «que María quiere a Juan» puede aprender un lenguaje que es capaz de expresar ambos. Ni es claro qué podría hacer que el lenguaje fuera sistemático si no es la sistematicidad de los pensamientos que permite expresar. Por ello, la idea de que la mente aprende la sistematicidad del lenguaje únicamente barre el problema de debajo la alfombra de la entrada, a debajo de la alfombra de la sala de estar. En conjunto, pienso que es mejor que demos por supuesto, y como no negociable, que la sistematicidad y la productividad están engarzadas en la «arquitectura» de la representación mental y no en las vaguedades de la experiencia (1998, pp. 26-27, la traducción es mía).

Este tipo de argumentación recuerda la que se planteó en la geología de los años veinte del siglo pasado, en contra de la teoría de Wegener de la deriva continental: él fue acusado de «resolver» el problema de la deriva continental, creando un problema todavía mayor, el de la fuerza gigantesca requerida para mover los continentes en primer lugar. Parafraseando a Fodor, la geología decidió dar por supuesto, como no negociable, que la tectónica de placas era parte constitutiva de la «arquitectura» de la corteza terrestre. Dicho de otro modo, el modo de Fodor de resolver el problema de la sistematicidad de los procesos cognitivos superiores consiste en sostener que toda la cognición humana es sistemática para evitar la pregunta por cómo deviene sistemática la mente. El problema con su estrategia es que no encuentra apoyo en la evidencia: hay diferencias entre las mentes no verbales y las verbales, y la sistematicidad es una de ellas (Penn, Holyoak & Povinelli, 2008). Aunque explicar cómo es que esto es así puede constituir un reto difícil, es inevitable afrontarlo, si queremos avanzar en nuestra comprensión.

Además, la dificultad explicativa que plantea Fodor no es tal: lo es únicamente si se da por supuesto el modelo de aprendizaje conceptual como formulación y puesta a prueba de hipótesis, como el mismo reconoció finalmente (2008). En el caso de la adquisición del lenguaje, este modelo supone concebir el aprendizaje espontáneo de la primera lengua del mismo modo en que se aprende una lengua extranjera de adulto: traduciéndola a la lengua que uno ya domina. En tal caso, efectivamente, la primera lengua requiere ya de una lengua originaria en la que formular tales hipótesis. Pero este no es el único modelo disponible de adquisición de la primera lengua; al contrario: toda la evidencia señala que no aprendemos a hablar del mismo modo en que un adulto aprende una lengua extranjera (Gomila, 2011).

En cualquier caso, la cuestión señalada por Fodor sigue siendo relevante: explicar cómo una mente no sistemática puede adquirir un lenguaje sistemático (y cómo, al hacerlo, adquiere la capacidad de la sistematicidad). Este es el reto que las teorías de procesos duales de la mente tratan de resolver. Por ejemplo, distinguiendo entre representación implícita y explícita: «aprender palabras proporciona etiquetas internas explícitas para ideas que antes eran meramente implícitas, y esta ganancia en explicitud tiene consecuencias cognitivas» (Gentner, 2003, p. 225). En realidad, esta es la primera lección de la psicología cognitiva: «la forma más frecuente de recodificación [...] consiste en traducir a un código verbal» (Miller, 1956, p. 89) lo que tiene efectos de economía cognitiva, y de facilitación de diversos procesos, incluyendo la memoria a largo plazo o el razonamiento. En nuestra opinión, el elemento clave, no obstante, es la estructura de dependencia jerárquica lo que proporciona el andamiaje formal que reestructura el tipo de contenidos que podemos pensar y permite su combinación sistemática.

6. CONCLUSIÓN: UN ENFOQUE DUAL DE LA COGNICIÓN HUMANA

Son múltiples los investigadores en el área de la psicología del pensamiento que han convergido en la idea de un doble régimen de funcionamiento de la cognición humana: por una parte, un sistema de procesos intuitivos, rápidos, automáticos, inconscientes, implícitos, en paralelo, de tipo asociativo; y por otra, un sistema de procesos reflexivos, lentos, voluntarios, conscientes, explícitos, seriales, basados en reglas. Así, se han propuesto teorías duales para la memoria (Schacter, 1987), el razonamiento (Kahneman, 2002; Evans & Over, 1996; Stanovich, 1999; Sloman 1996 y 2002), la representación del conocimiento (Anderson, 1993, Dienes & Perner, 1999), el aprendizaje (Berry & Dienes, 1993; Reber, 1989) y la teoría de la mente (Gomila, 2001 y Reddy, 2008). Creemos que la distinción entre procesos sistemáticos y no sistemáticos también coincide con esta dualidad.

Los procesos del primer grupo serían los cognitivamente básicos, evolutivamente antiguos, y no fácilmente controlables (por ejemplo, mediante instrucciones verbales). Los del segundo requieren atención y esfuerzo, llevan más tiempo y requieren de control consciente. Así, si tomamos el caso del aprendizaje de habilidades, podemos distinguir entre habilidades de bajo nivel para la navegación, como desplazarse en línea recta, o de alto nivel, como pilotar un barco. La cuestión crucial, como vimos al final de la sección anterior, es: ¿cómo están relacionadas?, ¿cómo se relacionan los dos niveles de la cognición? En particular, ¿constituyen sistemas separados, pero dentro de una misma arquitectura?, ¿o compiten entre sí para tomar el control del sistema? Incluso: ¿cómo es que el segundo nivel aparece en el curso del desarrollo, si con el primer sistema sería suficiente?

Actualmente no hay consenso sobre cómo responder a estas preguntas. Por una parte, hay quien concibe la arquitectura cognitiva básica como una serie de múltiples sistemas modulares, de «conocimiento nuclear», innato, y que el segundo nivel, posibilitado por el lenguaje, permite integrar esos diferentes procesos, haciendo posible un mayor nivel de control, y de flexibilidad, a través de la simulación mental mediante el habla interna (Frankish, 2004; Carruthers, 2005; Evans & Frankish, 2009). Pero también es posible concebir el nivel básico en términos interactivos y dinámicos. Según esta otra versión, el nivel básico de la arquitectura cognitiva involucra la capacidad, independiente del lenguaje, de la categorización perceptiva (como condición de posibilidad de adquirir el lenguaje en primer lugar), pero no es preciso concebirla en términos simbólicos, al contrario: el desarrollo del paradigma de la cognición corpórea muestra que es mejor concebir la categorización perceptiva como un proceso continuo de diferenciación, como la activación de patrones distribuidos de redes neuronales, relativos al contexto, y acoplados a aspectos del ambiente

(Spivey, 2007; Kello y otros, 2010). Desde este punto de vista, no es preciso postular un lenguaje del pensamiento para dar cuenta de esta capacidad cognitiva básica, ni concebirlo en términos de la activación de símbolos mentales. Luego, las etiquetas léxicas transforman estas contingencias sensoriomotoras robustas en unidades significativas que pueden ser combinadas sistemáticamente gracias a las estructuras sintácticas.

La cognición humana, por tanto, aunque surge de la interacción sensoriomotora, corporal, con el mundo, no se agota en ella; o mejor dicho, parte de esta experiencia interactiva tiene que ver con símbolos sociales, que son internalizados y dan lugar a mediadores simbólicos internos de tales interacciones. Los procesos superiores, abstractos, discretos, proposicionales, controlados, flexibles, parecen depender de tales mediadores, que resultan, entonces, un nuevo nivel de organización cognitiva. Una función cognitiva del lenguaje, en consecuencia, puede consistir en hacer posible una mayor abstracción, independiente del contexto y flexible.

El punto de consenso importante, en cualquier caso, para las teorías de procesos duales radica en ver al lenguaje como el desarrollo crítico para hacer posible este nivel superior de organización cognitiva. Así, cuando consideramos las habilidades cognitivas no verbales de los humanos adultos, no encontramos grandes diferencias con las de otros primates. Se asemejan en que utilizan mecanismos rápidos y frugales, inconscientes (que pueden llamarse módulos, si se quiere). Las mentes verbales, por el contrario, son lentas, esforzadas, conscientes, inferenciales, flexibles. Podría decirse que, a través del lenguaje, nuestras mentes devienen de propósito general, mientras que las mentes animales son especializadas (Premack, 2004). Los humanos pueden recombinar elementos mentales más allá de su experiencia sensoriomotriz. Mientras que las mentes no verbales pueden representarse lo que perciben, los humanos pueden representarse lo que imaginan. Otro modo de decirlo es que el lenguaje proporciona las estructuras proposicionales de los pensamientos. De nuevo, esto invierte el orden de dependencia entre lenguaje y pensamiento: es el lenguaje el que convierte al pensamiento en sistemático, al proporcionar a nuestra mente un sistema combinatorio de representación de «infinitud concreta». Además, permite la meta-representación y formas crecientes de control y flexibilidad cognitiva.

REFERENCIAS BIBLIOGRÁFICAS

- Anderson, John (1993). *Rules of the Mind*. Nueva Jersey: Lawrence Erlbaum.
- Aizawa, Kenneth (2003). *The Systematicity Arguments*. Dordrecht: Kluwer Academic Press.
- Barceló-Coblijn Lluís, Bernat Corominas-Murtra & Antoni Gomila (2012). *From Syntactic Trees and Small-World Networks. Complexity in the Ontogeny of Syntactic Networks of Three Different First Languages: Dutch, German and Spanish*. Trabajo presentado en el taller In Search of New Ways to Understand Language Variation, Change and Acquisition. Miramar Palace, San Sebastián, 25 y 26 de junio.
- Berry, Diane & Zoltán Dienes (1993). *Implicit Learning: Theoretical and Empirical Issues*. Nueva Jersey: Lawrence Erlbaum.
- Broadbent, Donald Eric (1971). *Decision and Stress*. Nueva York: Academic Press.
- Carruthers, Peter (2005). *The Architecture of the Mind*. Oxford: Oxford University Press.
- Cheng, Ken (1986). A Purely Geometric Module in the Rat's Spatial Representation. *Cognition*, 23(2), 149-178.
- Costello, Fintan & Mark Keane (2000). Efficient Creativity: Constraints on Conceptual Combination. *Cognitive Science*, 24(2), 299-349.
- De Villiers Jill G. & Peter A. de Villiers (2009). Complements Enable Representation of the Contents of False Beliefs: The Evolution of a Theory of a Theory of Mind. En Susan Foster-Cohen (ed.), *Language Acquisition* (capítulo 9). Palgrave: Macmillan.
- Dennett, Daniel (1993). Learning and Labeling. *Mind and Language*, 8(4), 540-548.
- Dienes, Zoltán & Joseph Perner (1999). A Theory of Implicit and Explicit Knowledge. *Behavioral and Brain Sciences*, 22(5), 735-808.
- Estes, Zachary & Sam Gluksberg (2000). Interactive Property Attribution in Concept Combination. *Memory & Cognition*, 28(1), 28-34.
- Evans, Jonathan & David Over (1996). *Rationality and Reasoning*. Nueva York: Psychology Press.
- Evans, Jonathan & Keith Frankish (eds.) (2009). *In Two Minds: Dual Processes and Beyond*. Oxford: Oxford University Press.
- Fodor, Jerry (1975). *The Language of Thought*. Nueva York: Harvester Press.
- Fodor, Jerry (1987). Why There Still Has to Be a Language of Thought? En *Psychosemantics* (pp. 135-154). Cambridge: The MIT Press.
- Fodor, Jerry (1998). *Concepts. Where Cognitive Science went wrong*. Oxford: Clarendon Press.
- Fodor, Jerry (2008). *LOT2. The Language of Thought Revisited*. Oxford: Oxford University Press.

- Fodor, Jerry & Zenon Pylyshyn (1988). Connectionism and Cognitive Architecture. *Cognition*, 28, 3-71.
- Frankish, Keith (2004). *Mind and Supermind*. Cambridge: Cambridge University Press.
- Gentner, Dedre (2003). Why are we so Smart. En Dedre Gentner y Susan Goldin-Meadow (eds.), *Language in Mind* (pp. 195-235). Cambridge: The MIT Press.
- Gómez, Juan Carlos (2004). *Apes, Monkeys, Children, and the Growth of Mind*. Cambridge: Harvard University Press.
- Gomila, Antoni & Paco Calvo (2008). Directions for an Embodied Cognitive Science: Towards an Integrated Approach. En Paco Calvo y Antoni Gomila (eds.), *Handbook of Cognitive Science: an embodied approach* (pp. 1-25). Londres: Elsevier.
- Gomila, Antoni (2001). La perspectiva de segunda persona: mecanismos mentales de la intersubjetividad. *Contrastes*, 6, 65-86.
- Gomila, Antoni (2011). The Language of Thought: Still a «Game in Town»? *Teorema*, 30(1), 145-155.
- Gomila, Antoni (2012). *Verbal Minds: Language and the Architecture of Cognition*. Hadley: Elsevier.
- Hermer-Vazquez, Linda & Elizabeth Spelke (1994). A Geometric Process for Spatial Reorientation in Young Children. *Nature*, 370, 57-59.
- Hermer-Vazquez, Linda & Elizabeth Spelke (1996). Modularity and Development: The Case of Spatial Reorientation. *Cognition*, 61(3), 195-232.
- Hermer-Vazquez Linda, Elizabeth Spelke & Alla Katnelson (1999). Sources of Flexibility in Human Cognition: Dual-task Studies of Space and Language. *Cognitive Psychology*, 39(1), 3-36.
- Hinzen, Wolfram (2012). Narrow Syntax and the Language of Thought. *Philosophical Psychology*, 23(1), 1-23.
- Kahneman, Daniel (2002). *Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice*. Discurso de aceptación del premio Nobel. <http://nobel-prize.org/economics/laureates/2002/kahneman-lecture.html>
- Karmiloff, Kyra & Annette Karmiloff-Smith (2002). *Pathways to Language*. Cambridge: Harvard University Press.
- Kello, Christopher y otros (2010). Scaling Laws in Cognitive Sciences». *Trends in the Cognitive Sciences*, 14(5), 223-232.
- Lupyan, Gary, David Rakison & James McClelland (2007). Language is Not Just for Talking: Redundant Labels Facilitate Learning of Novel Categories. *Psychological Science*, 18(2), 1077-1083.

- Lupyan, Gary (2008a). The Conceptual Grouping Effect: Categories Matter (and Named Categories Matter More). *Cognition*, 108(2), 566-577.
- Lupyan, Gary (2008b). From Chair to «Chair»: A Representational Shift Account of Object Labeling Effects on Memory. *Journal of Experimental Psychology: General*, 137(2), 348-369.
- Lupyan, Gary, Sharon Thompson-Schill & Daniel Swingley (2010). Conceptual Penetration of Visual Processing. *Psychological Science*, 21(5), 682-691.
- McLaughlin, Brian (2009). Systematicity Redux. *Synthese*, 170, 251-274.
- Miller, George (1956). The Magical Number 7, Plus or Minus 2: Some Limits on Our Capacity for Processing Information. *Psychological Review*, 63(2), 81-97.
- O'Keefe, John & Neil Burgess (1996). Geometric Determinants of the Place Fields of Hippocampal Neurons. *Nature*, 381(6581), 425-428.
- Penn, Derek, Keith Holyoak & Daniel Povinelli (2008). Darwin's Mistake: Explaining the Discontinuity between Human and Nonhuman Minds. *Behavioral and Brain Sciences*, 31(2), 109-78.
- Premack, David (2004). Is Language the Key to Human Intelligence? *Science*, 303(5656), 318-320.
- Reber, Arthur (1989). Implicit Learning and Tacit Knowledge. *Journal of Experimental Psychology: General*, 118(3), 219-235.
- Reddy, Vasudevy (2008). *How Infants Know Minds*. Cambridge: Harvard University Press.
- Schacter, Daniel (1987). Implicit Memory: History and Current Status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13(3), 501-518.
- Sloman, Steven (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin*, 119(1), 3-22.
- Sloman, Steven (2002). Two Systems of Reasoning. En Thomas Gilovich, Dale Griffin y Daniel Kahneman (eds.), *Heuristics and Biases* (pp. 379-396). Cambridge: Cambridge University Press.
- Spelke, Elizabeth (2003). What Makes us Smart? Core Knowledge and Natural Language. En Dedre Gentner y Susan Goldin-Meadow (eds.), *Language in Mind. Advances in the Study of Language and Thought* (pp. 277-312). Cambridge: The MIT Press.
- Spivey, Michael (2007). *The Continuity of Mind*. Oxford: Oxford University Press.
- Stanovich, Keith (1999). *Who is Rational? Studies of Individual Differences in Reasoning*. Nueva Jersey: Erlbaum.

EL VÍNCULO DE APEGO COMO ESCENARIO PARA EL DESARROLLO DE LA COGNICIÓN SOCIAL TEMPRANA

Carla Mantilla

Pontificia Universidad Católica del Perú

El estudio ontogenético de la cognición social se ha visto enriquecido en las últimas décadas con la aparición de modelos que enfatizan el origen relacional de las capacidades para experimentar y procesar información del mundo social (Carpendale & Lewis, 2010). Uno de los temas clásicos y centrales en el estudio de la cognición social es la llamada *teoría de la mente*, entendida como la capacidad para explicar las acciones propias y ajenas a partir de atribuir a los agentes estados mentales tales como deseos y creencias (Premack & Woodruff, 1978). Su emergencia, esperada aproximadamente entre los tres y los cinco años de edad, marca el inicio paradigmático del conocimiento psicológico del mundo social. Es justamente el estudio de la teoría de la mente, el que ha sido tomado como punto de referencia por las nuevas corrientes de estudio de la cognición social en el desarrollo para señalar dos problemas. De un lado, la desatención a las formas de cognición social tempranas y preverbales apoyadas en otros hitos del desarrollo como por ejemplo la imitación (Meltzoff & Moore, 1977), las *protoconversaciones* (Trevarthen & Aitken, 2001), la *interafectividad* (Stern, 1985) y la capacidad para *la atención conjunta* (Carpendale & Lewis, 2010). De otro lado, la tendencia a mantener una visión individualista del infante, y por tanto a no contemplar los escenarios intersubjetivos en los que la cognición social temprana se desarrolla (Carpendale & Lewis, 2010; Fonagy & Target, 1997).

El estudio filogenético del ciclo vital humano representado en la actualidad por la llamada psicología evolucionista del desarrollo, sugiere que la necesidad de un periodo de infancia suficientemente extenso fue seleccionado en la evolución para asegurar la transmisión cultural y el aprendizaje y desarrollo de las habilidades psicosociales en los individuos (Bjorklund & Pellegrini, 2002). Asimismo, dicha corriente sostiene que determinadas características del cuidado infantil y más específicamente de los cuidadores principales (en tanto representantes de este primer ambiente social infantil),

fueron seleccionadas en la evolución para facilitar el despliegue de las capacidades asociadas a la cognición social y sus correlatos biológicos (Fonagy, 2008 y Gergely, 2007). En esta línea, Schore (2000) llama la atención acerca de la inmadurez cerebral con la que el bebé nace y la necesidad de un entorno inicial que provea las condiciones de seguridad para que el cerebro culmine su desarrollo. El rol facilitador de este ambiente sería central sobre todo durante los dos primeros años de vida, momento en el cual los circuitos neuronales encargados de las funciones de regulación —básicamente la regulación afectiva y la comunicación emocional—, centrales para la organización del sí mismo, están consolidándose.

El vínculo de apego, entendido como un sistema diádico de regulación afectiva que ofrece al infante un refugio de seguridad ante situaciones de amenaza y un punto de partida para la exploración del mundo (Siegel, 1999), se presenta como el escenario ideal para recibir al neonato, brindando el tipo de estímulos que el cerebro necesita para su desarrollo, y generando las condiciones para que el infante explore, distinga y se ubique en un mundo físico y en un mundo relacional. En ese sentido, lo consideramos un referente conceptual importante para comprender los mecanismos y procesos ambientales que el programa genético asociado con el desarrollo de la cognición social requiere para su despliegue en el desarrollo.

En tradición intersubjetivista de la filosofía de la mente, Davidson (2001) plantea que todo conocimiento y acto de comunicación intencional tienen como condición la *triangulación*, es decir, la relación de mutua implicancia que existe entre el conocimiento del entorno, de uno mismo y de los demás; aquello que nos permite reconocernos compartiendo un mundo con otros individuos. En este sentido, la capacidad para comprender la propia mente y las demás mentes, elemento central de la cognición social, supondría la capacidad para triangular. Pensamos que una condición inicial para dicho despliegue articulado de estos tres dominios de la experiencia, el del sí mismo, el conocimiento de las mentes y el mundo circundante, supondría la capacidad inicial o previa del infante para diferenciar estos tres elementos, es decir, diferenciarse de los demás y del mundo no animado (Thompson, 2006 y Stern, 1985). La condición sobre la cual se levantaría la *triangulación* estaría dada pues, por la presencia de un triángulo originario no articulado en sus vértices, en tanto el infante no sería aún capaz de reconocerse a sí mismo compartiendo un mundo con otros. En el marco descrito por la psicología del desarrollo contemporánea, la capacidad de percibir de forma diferenciada el mundo físico, el propio sí mismo y los demás sujetos, estaría presente al menos de forma implícita, en los primeros meses de vida. Como parte de su dotación, el infante cuenta con mecanismos para percibir de forma diferenciada el mundo animado del inanimado (Thompson, 2006), y es de la mano de sus cuidadores que aprenderá información importante acerca de la naturaleza

del ambiente físico y del mundo social. Con relación al mundo interpersonal, desde que nace, el bebé responde de forma distinta ante los rostros mostrando una clara preferencia por estos en comparación con otros estímulos (Carpendale & Lewis, 2010). Su interacción con las personas supone el despliegue de acciones como la imitación (Meltzoff & Moore, 1977), la coordinación de movimientos y de sonidos en el juego libre y en las *protoconversaciones* (Trevvarthen & Aitken, 2001), la capacidad de entonar o sintonizar afectivamente (Stern, 2004), etcétera. El bebé usa, interactúa y espera cosas distintas de las personas en comparación de los objetos físicos, entre otros motivos porque las primeras son agentes o seres capaces de causar sus propias acciones.

Pero ¿en qué momento del desarrollo puede evidenciarse que estos tres dominios de la experiencia se articulan de forma tal que el sujeto se sepa compartiendo un mundo común con los demás?, en otras palabras ¿qué hito del desarrollo evidencia la capacidad de *triangulación* descrita por Davidson la cual resulta de suma importancia para comprender fenómenos humanos tan complejos como la intersubjetividad y la atribución psicológica? Si bien dicho filósofo analítico planteó las condiciones lógicas necesarias para hablar de *triangulación*, es decir, la aparición simultánea de estos tres modos de conocimiento en el desarrollo y la irreductibilidad o imposibilidad de que cualquiera de ellos derive de los otros dos, no consideró evidencia importante proveniente de la investigación centrada en la cognición social temprana que diera pistas más finas sobre su construcción y complejización en el desarrollo. Asimismo, Davidson otorgó un rol central a la creencia como estado mental paradigmático en juego en las actividades interpretativas del ser humano y la comunicación intencional. Otros como Tomasello y otros (2005), proponen atender al rol de la percepción y la atención como estados mentales paradigmáticos del niño preverbal participante de una situación triangular.

A partir de revisar evidencia de la psicología del desarrollo actual y de la observación de infantes proveniente del psicoanálisis, nosotros proponemos dos hitos en el desarrollo de la *triangulación*. Un primer hito lo situamos entre los nueve meses y el primer año de vida momento en el cual el infante es capaz de atender sostenidamente a un foco atencional junto a otra persona, manifestando una forma de *triangulación* implícita y no reflexiva (Eilan, 2005; Carpendale & Lewis, 2010). Posteriormente, nuevas capacidades cognitivas y afectivas como, la conciencia reflexiva o extendida, el lenguaje, la atribución de deseos y luego de creencias, la capacidad *metarrepresentacional*, la distinción entre apariencia y realidad, el juego simbólico, etcétera, proporcionarán las condiciones necesarias para desplegar formas más complejas de conocimiento del mundo subjetivo y relacional, es decir, formas reflexivas o explícitas de *triangulación*. Por tanto, un segundo hito estaría ubicado

a partir de los tres años momento en el que el niño manifiesta de forma más explícita y reflexiva el conocimiento del mundo subjetivo y del intersubjetivo y las capacidades interpretativas mentales que están a la base, a lo que llamaremos *mentalización* (Fonagy, 2008). Vamos a poner especial atención al del triángulo que vincula al sí mismo con los otros sujetos, aquel que nos explica la intersubjetividad y nos permite señalar de forma clara el rol central que el vínculo afectivo de apego tiene en el desarrollo del conocimiento de la mente y de las demás mentes.

1. EL VÍNCULO DE APEGO Y LA MENTALIZACIÓN

El apego es un sistema de representaciones mentales que sintetiza información acerca de la disponibilidad del cuidador principal para servir de refugio ante situaciones de amenaza y de base segura para explorar el mundo. Este sistema tiene la función de ayudar al pequeño a evaluar la accesibilidad del cuidador, para luego buscar la proximidad y garantizar la sensación de seguridad. El sistema de apego se construye a partir de las experiencias vinculares de los dos primeros años de vida, etapa en la cual, el adulto es un regulador de la experiencia del infante (Bowlby, 1969). Este utiliza las funciones psicológicas del adulto para organizar su propia experiencia, en especial, su capacidad para regular las emociones. De tal forma, las estrategias de regulación del adulto, es decir, su forma de experimentar y lidiar con las situaciones angustiantes, las emociones negativas y los momentos plácidos, se verán reflejadas en las formas de regulación que el infante comience a exhibir en su desarrollo (Siegel, 1999).

Además de la regulación afectiva, otros procesos psicológicos centrales se organizan gracias al vínculo de apego, en particular, la memoria autobiográfica, los procesos motivacionales y las formas de interacción con el mundo (Siegel, 1999 y Fonagy, 2008). Por ejemplo, la experiencia con un cuidador sintonizado y sensible a las necesidades del infante, rápido en sus respuestas y capaz de alinearse a su estado afectivo, facilitará la futura motivación por buscar en el entorno ayuda o apoyo durante situaciones problemáticas, así como la capacidad para autocalmarse. Asimismo, la atención del adulto a las necesidades del infante, la forma de validar sus demandas y su capacidad para verlo como un ser psicológicamente motivado, generará en el niño una imagen de sí mismo como sujeto querido, valorado y reconocido.

Al conjunto de esquemas o representaciones mentales generadas en las experiencias vinculares relacionadas con el apego se les denomina modelo operativo interno, el cual resume un estilo particular de experimentar y actuar que se pondrá en marcha en futuras situaciones de separación, amenaza o desregulación (Siegel, 1999). De tal modo, los estilos de apego expresan formas de regulación afectiva, interacción con el

mundo y narrativas autobiográficas. Se han identificado cuatro estilos de apego: uno seguro, uno inseguro evitativo, otro inseguro ambivalente y finalmente uno desorganizado (Fonagy, 2000 y Siegel, 1999).

Pero, ¿cómo hablar de un registro de experiencias en una etapa de la vida en la que aún no existe un registro explícito de los eventos? El tema de la memoria aquí es de particular importancia. En los momentos en que el sistema representacional del apego se construye, las experiencias vividas no pueden ser registradas a modo de eventos, aún no se han desarrollado los sistemas de memoria episódica, lo que opera es un registro de tipo implícito, sensorial y emocional (Schore, 2000). Recordemos que los sistemas de memoria explícitos y declarativos se desarrollan a la par de la región cerebral del hipocampo, la cual madura hacia los dos años y medio de vida (LeDoux 2001 y Schore, 2000). Antes de esto los centros de registro se encuentran básicamente ligados al cerebro primitivo, por ejemplo a la amígdala. Al respecto LeDoux (2001) señala que durante estas fases iniciales del desarrollo se genera una suerte de *self implícito*, el cual es el basamento más primario de nuestra subjetividad al contener preferencias, miedos, pautas afectivas y de reacción forjadas en los primeros momentos de la vida. En este sentido, el vínculo de apego tiene un rol importante en la construcción de las representaciones implícitas del sí mismo. Stern (2004), añadirá que es en esta etapa del desarrollo en la que también se construyen las formas implícitas de estar con otros individuos.

Ahora nos vamos a centrar en las características del cuidador que facilitan la construcción de un estilo seguro de apego en el niño y el desarrollo de su capacidad para conocer el mundo, su mente y las otras mentes. Es importante señalar que el estilo de apego seguro supone la construcción de un modelo operativo interno que opera de la siguiente forma en el infante: ante situaciones de separación el infante buscará la proximidad con el cuidador, recuperando la calma y retomando la actividad exploratoria. En el adulto, este estilo se evidencia en su capacidad para autocalmarse y buscar apoyo en situaciones angustiantes. Asimismo, los adultos con apego seguro tienden a comprender sus acciones y las del resto en base a la atribución de estados mentales y son capaces de tolerar que sus explicaciones tengan un grado de incertidumbre y que representen una perspectiva posible de los acontecimientos (Siegel, 1999).

En este punto resulta pertinente preguntarnos ¿qué características del cuidador facilitan la construcción de un estilo seguro de apego? En primer lugar, se plantea que el cuidador principal sea suficientemente sensible para captar las señales del bebé y responda a ellas en un tiempo prudente y con eficacia. Para ello, es fundamental que el adulto esté emocionalmente disponible y sea capaz de alinear su estado mental con el del niño (Siegel, 1999). Este alineamiento se resume en la siguiente situación: el adulto capta el estado emocional del niño y le trasmite su entendimiento a partir

de canales no verbales, que a su vez transforman la experiencia del niño reduciendo la angustia y aumentando la comodidad. La sensibilidad supone además, la capacidad para captar los momentos en los que el infante necesite soledad y retiro, por tanto, un cuidador sensible debe saber acercarse y retirarse oportunamente. En general, las relaciones de apego tienen la característica de ser como una danza circular sincronizada en la que se alternan momentos de compromiso y descompromiso (Siegel, 1999).

Gergerly (2007 y 2008) plantea que el vínculo de apego es un sistema de *pedagogía humana* que abre la puerta al desarrollo de la comprensión mental, y cuya tarea inicial es facilitar el aprendizaje sobre las emociones, elemento central de la vida social y de la experiencia subjetiva. Esta pedagogía supone que el adulto muestre al niño el valor informativo y comunicativo que las emociones tienen, el impacto que generan en la propia experiencia, en las demás personas y las distintas formas de expresarlas y regularlas.

La sensibilidad del cuidador se relaciona estrechamente con su capacidad para interpretar psicológicamente al infante, es decir, con su tendencia a atribuirle estados mentales y motivaciones propias. En este sentido, Fonagy (2008) y Fonagy y Target (1997) sostienen, que un predictor del estilo seguro de apego es la capacidad de *mentalización* de los cuidadores. La hipótesis es la siguiente: padres sensibles y *mentalizantes* estarán más dispuestos a captar e interpretar las distintas señales del bebé de un modo eficaz, ello generará la sensación de seguridad en el infante, el despliegue de su interés por la mente del adulto y la tranquilidad para explorar el entorno físico, mental y social.

Pero, ¿qué significa y qué implica exactamente la *mentalización*? Se trata del conjunto de capacidades empleadas en la interpretación de los estados mentales propios y ajenos. Implica percibir, imaginar e interpretar explícita e implícitamente las acciones propias y las de los demás en términos de estados mentales (Fonagy, 2008). Refiere a una actividad que sucede a varios niveles de conciencia y se apoya en distintos mecanismos. A nivel implícito se evidencia en acciones no reflexivas como la intuición, el reflejo y el alineamiento emocional, es decir, la capacidad de transmitir al infante con gestos y acciones aquello que está siendo captado. Por ejemplo, gesticular y emitir determinados sonidos cuando se percibe que el niño está feliz, triste o frustrado por algo. A nivel explícito, supone verbalizar una explicación acerca del estado psicológico del niño, tratando de brindarle una comprensión de su situación. Por ejemplo, cuando en una situación en la que un niño tiene miedo de dormir solo y alarga el momento de irse a acostar pidiendo otro cuento, el adulto le comenta que quizá sienta temor de separarse y estar solo, y lo invita a explorar junto con él su sentir ofreciéndole alternativas para calmarse.

En resumen, el concepto de *mentalización* interrelaciona cuatro aspectos. Uno metacognitivo, en tanto supone una actividad interpretativa sobre los estados mentales. Otro atribucional, en tanto implica asimismo la atribución de significados o hipótesis implícitas o explícitas que usamos para comprendernos y comprender a los demás. Uno relacional en tanto se trata de una capacidad puesta en marcha en el mundo social de seres intencionales y, finalmente, contempla un aspecto individual en tanto se trata de un conjunto de habilidades y procesos que varía y se observa en diferentes grados y niveles de complejidad (Holmes, 2010).

La *mentalización*, fundamental para nuestra supervivencia como especie y nuestra adaptación en el desarrollo, se despliega en una línea del desarrollo que podemos rastrear a partir de cuatro hitos y que revela cuatro momentos en la comprensión de la agencia, entendida como la forma de experimentar el sí mismo (*self*) como causante de las acciones (Fonagy, 2008). Dicha agencia que en un momento inicial es básicamente física o corporal, evoluciona a formas mentalísticas aumentando la posibilidades del infante de comprender que las acciones son causadas por estados mentales propios a cada agente.

1. Agencia teleológica. En la segunda mitad del primer año de vida el infante muestra capacidades para establecer relaciones causales entre las acciones de los agentes y el mundo. Puede distinguir que las acciones provocan consecuencias, es decir, que son medios para conseguir fines específicos. Por tanto, se percibe como un agente capaz de lograr una meta eligiendo la acción más eficiente entre varias alternativas. Los niños esperan que los humanos se comporten como si fueran máquinas predecibles y no toman en cuenta los estados mentales para explicar sus acciones. Hacia el final de esta etapa aparece la referencia social, en la cual el niño observa los gestos del adulto para encontrar información que le permita acercarse o alejarse de ciertos objetos. Asimismo, comienza el señalamiento de objetos con alternancia de miradas hacia el adulto y la atención compartida.
2. Agencia intencional. En el segundo año se desarrolla una comprensión mentalística de la agencia, el infante comprende que las acciones son causadas por estados mentales tales como los deseos. Comienza el juego simbólico compartido y junto a él se incrementan las habilidades colaborativas. La regulación de emociones se puede comprender en esta etapa como un reflejo de la relación con el cuidador principal. Se desarrolla la capacidad para razonar no egocéntricamente acerca de los deseos y sentimientos de los demás. Sin embargo, aún no se puede representar estados mentales independientemente de la realidad física, por tanto la diferenciación entre la realidad interna y externa, apariencia y realidad, aun no está consolidada.

3. Agencia representacional. Entre los tres y los cuatro años la comprensión *mentalística* de la agencia incluye estados mentales como las creencias. Pueden ser capaces de distinguir entre lo que perciben en un agente y lo que podría ocurrir en la mente del agente. Se logra una comprensión de las emociones del otro a través de la empatía. Pueden atribuir falsas creencias a sí mismos y a los demás, acceden a intercambios sociales nuevos como bromear, engañar y hacer trampa.
4. Agencia autobiográfica. Hacia los seis años se suman avances en las habilidades para recordar y contar experiencias y actividades intencionales de forma coherente y organizada causal y temporalmente. Se complejiza la teoría de la mente con nuevas habilidades como la generación de representaciones de segundo orden, la capacidad de dar cuenta de sentimientos mezclados o encontrados, es decir, atribuir conflicto o ambigüedad, desplegar formas de engaño más sutiles como mentiras blancas, poseer mayor conciencia de que las propias expectativas pueden influir o sesgar la interpretación de un evento ambiguo. En esta etapa emerge la comprensión de sí mismos a partir de la elaboración de narrativas autobiográficas.

Sintetizando, podemos decir que el tránsito hacia la *mentalización* implica pasar de una comprensión teleológica de los agentes basada en la aplicación de una racionalidad en la que se predice la acción según los medios observados que se consideren más eficaces para lograr ciertas metas, a una comprensión *mentalística*. Esta comienza con la atribución de deseos y emociones en un marco no egocéntrico, pero aún dependiente de la realidad externa, ya que todavía no hay conciencia de que los estados mentales son representaciones de escenarios posibles del mundo. La relación realidad y apariencia, mundo interno y mundo externo, no está aún delimitada. Fonagy (2008) y Fonagy y Target (1997), plantean que en esta etapa, el niño vive simultáneamente dos maneras de experimentar esta relación: el *modo aparente* (*pretend mode*) y el *modo de equivalencia psíquica*. En el modo de equivalencia, la relación entre mente y mundo está equiparada, es especular en tanto los estados mentales son reflejos del mundo y viceversa, no hay perspectivas. El modo de apariencia ocurre en la situación de juego, aquí, sorprendentemente el niño opera representacionalmente, puede creer que él es un bombero y que la cortina es una llama que apagará con un objeto que imagina es una manguera. Sin embargo, necesita mantener separadas la realidad interna de la externa. Para que la comprensión mentalística se desarrolle, ambos modos se integran de forma que resulta posible que los estados mentales sean representaciones que se relacionan con el mundo pero no de manera equivalente ni desacoplada. Por ello, en el siguiente paso, la comprensión *mentalística* se sofisticada e incluye la atribución

de creencias y perspectivas diferentes acerca del mundo. Finalmente la *mentalización* llega al nivel más avanzado, en que la actividad *metarrepresentacional* permite al niño tener estados mentales sobre sus estados mentales, atribuir ambigüedad y conflicto, ser irónico y desarrollar narrativas sobre sí mismo y los demás.

El desarrollo de la *mentalización* supone la presencia de cuidadores capaces de reflejar el estado mental del niño tanto a nivel implícito como explícito, a ellos se sumarán otros agentes socializadores como los pares y los maestros que serán de importancia fundamental para la mayor complejización y el afianzamiento de esta capacidad a lo largo de su desarrollo.

Una hipótesis interesante planteada por autores como Fonagy (2008), Fonagy y Target (1997), Siegel (1999) y Gergely (2008 y 2007), es que el tipo de experiencia vincular de apego tendrá consecuencias en el desarrollo de las habilidades relacionadas con la *mentalización*. El sustento empírico para defender esta hipótesis proviene tanto del trabajo clínico psicoanalítico como de la investigación científica en desarrollo humano, en ambos dominios se observa que ciertos estilos de apego se asocian con dificultades en la capacidad para mentalizar, priorizándose estrategias más básicas en la regulación de estados afectivos displacenteros, como por ejemplo la huida o el ataque.

La investigación se ha centrado en niños y adultos que presentan un estilo desorganizado de apego (Main & Goldwyn, 1998). En los niños clasificados con este estilo, el sistema de apego y la búsqueda de proximidad se activan no ante una amenaza externa sino ante la presencia del propio cuidador; de tal forma se genera una paradoja irresoluble en la cual quien debe proveer la calma es quien atemoriza (Fonagy, 2008 y Siegel, 1999). El pánico paraliza al niño y este activa estrategias como el congelamiento, la huida y el ataque. En adultos el estilo desorganizado se visualiza en sus estrategias para regular los afectos, las cuales no toman en cuenta los estados mentales y priorizan la descarga motora o la acción. Si consigo mismos cancelan las capacidades interpretativas, con los demás las exacerbaban. Llegan incluso a confundir, como en el modo prementalizador de equivalencia psíquica, lo que hay en su mente con la realidad, siendo proclives a ideaciones paranoides que los llevan a actuar impulsivamente en contra de sí mismos o de los demás. Se trata de personas que han sufrido severas experiencias traumáticas en las relaciones tempranas, en particular, en el vínculo con el o los cuidadores principales.

Otro hallazgo de la investigación revela una tendencia a la transmisión transgeneracional de los estilos de apego mediada por la capacidad de *mentalización* (Fonagy, 2008). Padres *mentalizadores* tienden a generar confianza y cercanía mental con sus hijos y ello facilita el vínculo seguro. Los déficits en la *mentalización* de los padres, generarían diferentes escenarios en el desarrollo del apego en sus hijos. En el caso de los padres con historia de trauma psíquico, con probables estilos desorganizados

de apego, se tenderá a revivir con sus hijos las situaciones traumáticas y desplegar defensas psicológicas como la negación del estado mental de estos, la disociación, y la necesidad de acallar la posible emergencia del trauma con actos de abuso y violencia hacia los niños. Se trata de adultos aterrorizados de sus hijos que resultan aterrorizantes para sus hijos, ello genera las condiciones para transmitir el patrón desorganizado de apego. Evidentemente, estamos hablando de población clínica, en el caso de los adultos generalmente con trastornos limítrofes de personalidad, patología en la cual, se exhiben enormes déficits en las capacidades regulatorias de afectos e impulsos, tendencia a acciones compulsivas, autolesión, y gran inestabilidad emocional y relacional. En ellos se evidencia una dificultad para reflexionar sobre los estados mentales y representar mentalmente la experiencia. Ahora bien, los efectos de la experiencia vincular de apego en el desarrollo de la *mentalización* se observan también en los demás estilos organizados de apego.

En el caso de los estilos inseguros de apego, también se observa el impacto en las capacidades mentalizadoras aunque de manera menos severa. El estilo inseguro evitativo en niños es facilitado por figuras poco sensibles en la captación de señales y poco disponibles emocionalmente para el intercambio y la interacción con los niños. Por ello, en lugar de buscar la proximidad, el infante se muestra indiferente frente al cuidador cuando hay situaciones de desregulación. Prefiere seguir haciendo lo suyo y prescindir de la cercanía y el consuelo pues ha aprendido a no esperar un contacto sensible por parte de su figura de apego. De adulto, manifestará una tendencia a prescindir de los afectos como elemento informativo en la interpretación, mostrará distancia emocional, preferencia por lo práctico y racional y rechazo al recuerdo o tendencia a idealizar el pasado.

Con relación al estilo inseguro ambivalente, se observa que en situaciones de separación o desregulación, los niños buscan la proximidad de forma desesperada, llorando y mostrando rabia, y no consiguen consolarse ni calmarse con la cercanía del cuidador. En la adultez, este estilo llamado preocupado, evidencia una dificultad para liberarse de las experiencias tempranas difíciles y la tendencia a vivir atormentados por recuerdos que no pueden ser superados. Estas vivencias obstaculizan la captación de señales en los demás y generan la atribución poco fidedigna de contenidos mentales. Cuando estos adultos se convierten en padres, tienen a ser intrusivos en su afán de ser sensibles y no cometer los mismos errores que sus figuras de apego, y otras veces se distancian movidos por el mismo afán. Resultan difíciles de predecir para sus hijos y por ello, facilitan en estos la exageración de las estrategias para llamar su atención y una gran dificultad para regresar a la calma.

Vemos pues que los estilos de apego de los padres y sus capacidades mentalizadoras, se tienden a transmitir transgeneracionalmente.

En este marco nos preguntamos ¿las capacidades básicas de la cognición social, en particular las relacionadas con la interpretación de las acciones a partir de la atribución de estados mentales, son elementos garantizados en el desarrollo, únicamente dependiente de la activación de un programa genético? Pensamos que la evidencia proveniente de las investigaciones psicoanalíticas sobre el vínculo afectivo, y la psicología evolutiva, apoyan la tesis sobre la importancia que los escenarios relacionales tienen en el desarrollo de esta capacidad y en sus posibles déficits. Si nuestra mente es producto de la evolución, los escenarios que facilitan su desarrollo también lo serían.

2. ATENCIÓN CONJUNTA Y MENTALIZACIÓN: HITOS DE LA TRIANGULACIÓN

La *mentalización* o capacidad para interpretar la mente y las otras mentes a partir de atribuir estados mentales a los agentes, se consolida entre los tres y los cinco años de edad lo cual coincide con la aprobación de las pruebas de la falsa creencia, es decir, con la adquisición de una teoría de la mente (Fonagy & Target, 1997). La capacidad para representar estados mentales y el reconocimiento de que estos suponen formas de percibir el mundo que varían de sujeto a sujeto, son condiciones para la *triangulación*. Entendemos por *triangulación*, siguiendo a Davidson (2001), la idea de que todo conocimiento es producto de una relación triangular en la que el sujeto conoce e interpreta sus estados mentales, en un mundo que es objetivo y del cual se ha formado perspectivas y creencias junto a otros individuos a los que entiende a partir de atribuirles estados mentales. Siguiendo esta línea de razonamiento podríamos decir que la *mentalización*, la teoría de la mente y la *triangulación* se consolidan en simultáneo. Ahora bien, si la *mentalización* sigue una línea de desarrollo que implica la atribución de estados mentales previos a la creencia, como por ejemplo deseos o sentimientos (Thompson, 2006 y Fonagy, 2008) o incluso, en un nivel más básico, la atribución de metas (Tomasello y otros, 2005), nos preguntamos si podemos encontrar formas más básicas de *triangulación* que marchen a la par de los hitos de la *mentalización*.

A un nivel básico, pensamos que la *triangulación* requeriría el reconocimiento de uno mismo como distinto a otros sujetos y diferente del mundo físico. Podríamos suponer que a un nivel implícito, no reflexivo (Gergely, 2007; LeDoux, 2001; Stern, 2004 y 1985) el infante es capaz de diferenciar su experiencia de la de las personas de su entorno, en particular la madre (Stern, 1985), y reconocer un mundo físico como diferente al animado (Thompson, 2006). Esto, según las perspectivas contemporáneas psicoanalíticas sobre la mente (Stern, 2004) y planteamientos de la psicología del desarrollo centrada en el estudio de la cognición social temprana (Carpendale & Lewis, 2010; Thompson, 2006), existiría desde los primeros meses de vida. El mundo animado, tiene particularidades únicas dentro de las cuales el elemento

más saltante es la emoción y la contingencia de los intercambios sociales (Thompson, 2006; Gergely, 2007). Si bien en un momento inicial, el infante se interesa más en la sincronía entre sus gestos y los del cuidador, poco después de los tres meses comienza a motivarse más por los intercambios menos predecibles y cambiantes, sin curso fijo. Ello coincide con la mayor capacidad para discriminación visual de detalles en los rostros, sobre todo el de la madre.

En esta etapa el bebé humano muestra expectativas diferentes para los objetos físicos y para las personas, de estas, espera justamente intercambios que tengan a la emoción, la coordinación y la contingencia como protagonistas. Ningún otro objeto del mundo circundante del niño tiene la capacidad de ser un agente espontáneo y dinámico, capaz de dirigirse a metas. El infante espera que los agentes sociales se comporten de esa manera y muestra sorpresa cuando ello no ocurre, esto no sucede cuando se trata de objetos físicos. Esta expectativa se ve con claridad en experimentos como el del *still face*¹ (Tronick, 1989). Asimismo, dicha característica particular del mundo interpersonal se muestra en los episodios de intercambio cara a cara con el adulto, en dónde el infante es capaz de coordinar secuencias de sonidos y gestos en lo que se ha llamado *protoconversaciones* (Trevorthen & Aitken, 2001). Las capacidades imitativas que el bebé muestra desde las primeras horas de nacido, sugiere a los investigadores sobre el tema (Meltzoff & Moore, 1977; Stern, 1985), la posibilidad de que el bebé se perciba como un agente diferente del resto, capaz de reproducir en su experiencia algo que viene como estímulo de fuera. Asimismo, Stern (1985) señala que el bebé es capaz de percibirse como alguien que inicia y dirige su acciones, en ese sentido, se experimenta como un centro de experiencia diferente a otros sujetos capaces también de autoiniciar sus acciones. En este sentido Thompson (2006) agrega que la cognición social y la autoconciencia emergen simultáneamente y se complejizan a partir de la percepción de nuevas acciones y actividades acaecidas en el contexto relacional.

Luego de la primera mitad del año de vida, el conocimiento y cercanía emocional entre la figura de apego y el bebé han configurado una forma de comprensión y comunicación interpersonal basada en el afecto, el cual se considera, el medio y contenido primario de la comunicación. A ello se llama interafectividad, e implica tanto comunicar como captar un estado afectivo utilizando canales *transmodales* como

¹ El experimento en cuestión supone la grabación de secuencias de intercambio libre cara a cara entre madres y bebés. La consigna experimental supone que la madre en un momento dado se detenga y muestre un rostro neutro y ningún gesto de comunicación. Lo observado en los bebés son respuestas de sorpresa, intentos por «llamar la atención» de la madre con gestos y sonidos, luego fastidio y llanto. Los resultados revelan la expectativa de los infantes por la contingencia del intercambio con los adultos, en especial con sus madres.

el gesto, el movimiento y el sonido. La interafectividad se plantea como la forma más temprana de compartir experiencias internas (Stern, 2004).

De esta reseña podríamos decir que bastante antes del primer año de vida, los tres vértices de la *triangulación* son reconocidos y distinguidos por el infante a nivel implícito, pero ¿en qué momento podemos hablar de la articulación de los tres en una escena en que el infante perciba un mundo objetivo compartido con otros individuos? Sugerimos situar un primer hito de la *triangulación* alrededor entre los nueve meses y el primer año de vida, momento en el cual el infante puede compartir un foco atencional deliberadamente con otras personas. Nos referimos a la capacidad para la atención conjunta (Eilan, 2005; Carpendale & Lewis, 2010; Thompson, 2006; Stern, 2004 y 1985). Atender a un objeto del mundo, un juguete por ejemplo, se acompaña de otros fenómenos como la referencia social y el señalamiento (Stern, 1985 y Thompson, 2006). Ello implica que el infante intencionalmente dirige la atención del adulto señalando un objeto que despierte su interés y motive una intención por verlo y sostener la atención sobre él. Hay la percepción de seres intencionales al nivel de atribuirles la capacidad de atender a lo mismo, a sostener la atención y engancharse en una actividad conjunta que generalmente supone una secuencia de interacción que va sofisticándose hasta presentar turnos y reconocer roles específicos. Tomasello y otros (2005) denominan a este momento el *enganche triádico*, pues adulto e infante monitorean perceptualmente al otro y al objeto coordinando miradas e interactúan como agentes dirigidos a metas. Asimismo, a nivel de atención y percepción son capaces de reconocer la objetividad del mundo pues, al coordinar acciones con otros en torno a objetos, prueban y corrigen el curso de las acciones buscando dar con la manera correcta, eficiente, digamos objetiva, de realizar una meta o anticipar la acción del otro. Para estos autores estas capacidades humanas relacionadas con la motivación por compartir estados psicológicos y participar en acciones colaborativas requieren un contexto relacional para desarrollarse, pues no son elementos garantizados. En tanto escenario para explorar el mundo mental y físico con seguridad, pensamos que el vínculo de apego es en particular aquella relación que facilita tanto el reconocimiento como la articulación de los tres vértices del triángulo. Anteriormente, enfatizamos en las condiciones de seguridad que este vínculo ofrece tanto para la exploración del ámbito físico como del mental y social.

Existe evidencia desde la neurociencia (Fonagy, 2008; Fonagy, Gergely, Jurist & Target, 2002) que describe la vinculación entre las capacidades mentalizadoras y el apego. A nivel descriptivo ambas se asocian de forma positiva: individuos con apegos seguros muestran un desarrollo óptimo de la *mentalización*, apegos desorganizados implican fallas en la *mentalización*. Sin embargo, a nivel neurológico la relación es caprichosa a primera vista. En los mapeos cerebrales funcionales, lo que se observa

es un apagamiento de los centros cerebrales asociados con la *mentalización* en momentos en que se activan los circuitos cerebrales vinculados a la búsqueda de proximidad (los del apego). En primera instancia, pareciera que este hallazgo desacopla la relación entre apego y *mentalización*, sin embargo, una interpretación más fina supone que el desarrollo de formas mentalizadoras para interpretar la agencia, sucede en momentos calmos, en los que no hay que buscar con premura la cercanía para regular la experiencia, se trata de momentos en los que se facilita un nivel óptimo de activación (*arousal*) y de atención en el infante. Los investigadores, por tanto, sugieren que la exploración del mundo mental es facilitada por un contexto vincular que no tenga grandes sobresaltos, es decir, que implique la presencia de un cuidador percibido constantemente como una base segura. De esta forma, la exploración de las mentes y del mundo se desenvuelve, atenta, libre y confiadamente. Del mismo modo, se explica por qué en apegos patológicos desorganizados, en los que la figura de cuidado es también la persona temida de la cual hay que protegerse, el nivel alto de activación (*arousal*) producto de la angustia vivida obstaculiza que el nivel de atención óptimo para la exploración del mundo mental se produzca. Asimismo, la exploración de la mente resulta muy amenazante: implica tener que atribuir al cuidador intenciones malignas con relación a uno mismo y encontrar en su mente representaciones distorsionadas acerca de las propias motivaciones. Así, las capacidades mentalizadoras se detienen para poder sobrevivir y preservar a la figura de apego, por ello, el desarrollo del entendimiento de la agencia se detiene a un nivel teleológico, de equivalencia psíquica y de modo aparente. Se concluye de lo anterior que la exploración del mundo mental y físico se facilita cuando la relación de apego es percibida como segura.

De acuerdo con Tomasello y otros (2005), el enganche triádico paradigmático en la capacidad de atender conjuntamente a un objeto se sofisticaba con el desarrollo de mayores estrategias colaborativas en que el infante propone y modifica planes, conoce mejor su rol y el rol de los demás, y puede hacer inferencias más sofisticadas para comprender el proceder del adulto. La adquisición del lenguaje entra aquí a potenciar todas estas habilidades. Esto coincide en términos ontogenéticos con la atribución de deseos y sentimientos descrita por Fonagy (2008) como hito del desarrollo de la *mentalización*. Proponemos entonces que a nivel de comunicación intencional y habilidades interpretativas, el primer marco triangular aparece con la habilidad de la atención conjunta al interior de la seguridad del vínculo afectivo. El paso a la *mentalización* totalmente desarrollada se acompaña de la adquisición de mayores capacidades metarrepresentacionales apuntaladas por el lenguaje, y de un reconocimiento más potente del mundo objetivo y del lugar que el infante y los demás ocupan en él. La posibilidad de atribuir a uno mismo y a los demás perspectivas diferentes sobre un mundo común, el cual se puede nombrar, con el cual se puede

jugar y al que se puede transformar internamente con la imaginación, describen la versión más lograda de la *triangulación* y la *mentalización* y por tanto, el conocimiento reflexivo del mundo subjetivo, interpersonal y del entorno. Lo que Fonagy (2008) llama agencia representacional. En esta etapa, se desarrollan además los registros mnémicos declarativos, episódicos, simbólicos y explícitos de la memoria, contribuyendo a la construcción de la memoria autobiográfica y la historización del niño al interior de una familia, una comunidad y un entorno. Aquí, pares, padres y maestros serán la proyección de ese vínculo primero de seguridad, es decir, la pedagogía humana (Gergely, 2007) responsable de facilitar el desarrollo de lo que Tomasello y otros (2005) llaman la pequeña diferencia que hace la gran diferencia entre los seres humanos y otros animales, en síntesis, las capacidades para compartir contenidos sobre el mundo y tener una intencionalidad colaborativa.

3. MENTALIZACIÓN: ¿MÓDULO INNATO Y/O LOGRO DEL DESARROLLO?

La pregunta sobre el tipo de mecanismo que explica la aparición de las habilidades para la interpretación de la propia mente y de las demás mentes en el desarrollo, se puede contestar al menos de tres maneras desde la filosofía de la mente y la psicología del desarrollo (Carpendale & Lewis, 2010). Al hablar de *mentalización*, según hemos dicho, nos referimos a un conjunto de habilidades que nos permiten interpretar la propia mente y las demás mentes a partir de atribuir, a nivel explícito o implícito, estados mentales. Si bien la *mentalización* se consolida hacia entre los tres y cinco años, es antecedida por formas más elementales de atribución, básicamente de metas y deseos. Sin embargo, en la visión más tradicional sobre el tema, la teoría de la mente es el nombre del constructo que hoy podemos relacionar con la *mentalización* (Fonagy, 2008), la lectura de mentes y la autoatribución de estados mentales (Carruthers, 2006). Los tres modelos que describen el tipo de mecanismo que explica la teoría de la mente son la teoría-teoría (Wellman, 1990 y Perner, 1991), la teoría de la simulación (Goldman, 2006 y Gordon, 1986) la teoría del módulo innato (Carruthers, 2006).

En breve, la primera entiende que la teoría de la mente responde al desarrollo de una serie de habilidades inferenciales a partir de las cuales el niño, como un pequeño científico, elabora una teoría acerca del funcionamiento de la mente a partir de la observación y el establecimiento de relaciones causales entre los agentes y sus acciones. El segundo modelo, el de la simulación, sostiene que el individuo imagina un escenario en donde reproduce las circunstancias ajenas en situaciones contrafácticas. Se trata de una actividad imaginativa mediante la cual el sujeto se pone en las circunstancias del otro para comprenderlo. El tercer modelo supone la existencia

de un módulo innato, es decir, un dispositivo de procesamiento y motivación de dominio específico y de base neural, llamado módulo de la lectura de mentes, que se activa en el desarrollo por programación genética (Carruthers, 2006). Hay quienes han señalado que dichos modelos tendrían un compromiso con visiones internalistas de lo mental en tanto desatienden el contexto relacional en el que se desarrollan las habilidades interpretativas. (Carpendale & Lewis, 2010).

En este sentido, se señala que la caracterización del «niño científico» propuesto por la teoría-teoría, es la de un individuo que mantiene un estado contemplativo y por ende distante con relación al mundo social. La teoría de la simulación, hoy apuntalada por las investigaciones sobre el sistema de neuronas espejo (Iacoboni, 2009), pareciera reducir el fenómeno de la interpretación a la existencia de un simulador encarnado que automáticamente permite el conocimiento de las otras mentes a partir de una perspectiva de primera persona. Finalmente, la teoría del módulo innato, resta importancia a las experiencias relacionales de la infancia en su capacidad para alterar la activación del módulo en cuestión. En los tres casos, señalan Carpendale y Lewis (2010), se parte del estudio de la mente del niño como si esta estuviera separada de las otras mentes, y a las cuales llega a comprender con el paso del tiempo. Los autores sugieren, más bien, que las capacidades para la comprensión mental se desarrollan en escenarios relacionales y surgen en simultáneo.

La *mentalización* se construye tanto a partir de mecanismos de simulación, en tanto es una actividad imaginativa, como de inferencia, pudiendo ser esta automática como en el reflejo, o reflexiva como en la explicación. Además, es compatible con una visión modularista de la mente/cerebro, en tanto se sustenta en mecanismos especializados ubicados en circuitos neuronales específicos que interactúan con otros módulos especializados, por ejemplo, con el del sistema de apego (Fonagy, 2008). En definitiva, las habilidades mentalizadoras han sido producto de una serie de adaptaciones a lo largo de la evolución, que nos han permitido sobrevivir a entornos sociales cada vez más complejos. Entornos en los que hubo que generar escenarios mentales acerca de las posibles acciones de los competidores o de aquellos con los que se debía cooperar. En este sentido, se trata de un concepto que rescata elementos centrales de las posturas descritas.

Ahora bien, ¿la activación e implementación idónea de ese módulo inscrito en el código genético de nuestra especie y constitutivo de nuestra arquitectura cognitiva, es poco sensible al tipo de experiencia en el desarrollo? Parecería que hemos llevado este tema al terreno de la vieja discusión entre naturaleza y crianza. Sin embargo, nuestra postura pretende entender la naturaleza vía la crianza.

Carruthers (2006), defensor de una visión modularista masiva de la mente, sostiene que la naturaleza no se arriesga, en tanto que si bien el cerebro necesita de cierto

tipo de estímulos para el desarrollo del módulo asociado a la lectura de mentes, no resulta crucial ni determinante la cualidad de dicha estimulación. En su visión, el desarrollo y activación del módulo señalado depende en efecto de estímulos específicos como la presencia de un cuidador (en particular de la presencia de rostros humanos, voces, etcétera), pero importa solamente que esa presencia esté garantizada más no el tipo de experiencia particular que esa presencia ofrezca. La historia personal de cada bebé no parece ser aquí de gran importancia y podría estar sugiriéndose que las habilidades para regular las emociones y tener lo que Gergely (2008) llama un sí mismo psicológico, es decir, un sí mismo que se experimenta como agente de sus acciones y se percibe de forma relativamente estable a lo largo del tiempo y de las situaciones, son elementos dados y garantizados. La evidencia clínica y la psicopatología del desarrollo parecen contradecir radicalmente esta idea, señalando que se trata de logros determinados en gran parte por el tipo de experiencia vincular temprana, pues es esta, la pedagogía encargada de viabilizar el proyecto de la madre naturaleza, es decir, la activación y desarrollo del módulo relacionado a la comprensión mental.

El apego por ejemplo, resume una experiencia particular vincular en un modelo operativo interno, es decir, un conjunto de representaciones acerca de la disponibilidad del cuidador y las ventajas de la cercanía, que se generaliza a otras relaciones a lo largo de la vida. Esta experiencia vivida configurada en una representación mental, funciona como un filtro personal que sirve para interpretar los estímulos relacionados con el contacto interpersonal. En el modelo operativo está resumida una parte importante de la memoria autobiográfica a nivel implícito. Hemos visto como los apegos seguros generan expectativas positivas acerca del contacto interpersonal, facilitan la búsqueda de proximidad, ella a su vez permite explorar en calma el mundo mental, encontrar modelos eficaces para regular las emociones que redundarán en el desarrollo de la capacidad de autocalmarse, y comenzar el despliegue de las habilidades mentalizadoras. Sin embargo, esta historia, que ocurre aproximadamente en el 55% de la población (Bakermans-Kranenburg & Van IJzendoorn, 2009), se distingue de la del restante 45% en que se presentan los tres estilos de apegos restantes.

¿Qué pasa aquí con el despliegue del módulo innato? ¿Se requiere solo de la existencia de un estímulo llamado cuidador o también importa la cualidad de esa presencia? Como hemos mencionado, la evidencia sugiere que las habilidades para la *mentalización* no se desarrollan siempre al mismo nivel que en sujetos con apegos seguros, y el ejemplo más contundente es lo que ocurre en los apegos desorganizados (Fonagy & Target, 1997; Fonagy, 2008; Gergely, 2007). En estos últimos, las habilidades interpretativas están seriamente comprometidas evidenciándose una suerte de detención en su desarrollo producto de la presencia de situaciones vinculares adversas. En ese sentido, la obstaculización en el despliegue del módulo de la *mentalización*

en el desarrollo, parece deberse también a formas de adaptación y supervivencia ante situaciones que no son las esperadas. Por ejemplo, en personas que presentan este tipo de apego, las estrategias prioritarias utilizadas para la regulación afectiva son las de ataque y huida, evolutiva y filogenéticamente más primitivas que las mentalizadoras (Gergely, 2007). Asimismo, predominan formas de comprensión de la agencia similares a las de un niño en estadios previos a la *mentalización*, en los que se presentan dificultades en la diferenciación de la realidad externa de la interna, y en la capacidad para comprender las acciones psicológicamente (Fonagy & Target, 1997). De tal manera, se instala una tendencia a resolver las tensiones vía la descarga motora y física como se observa en conductas como la autolesión, violencia física, y otros actos impulsivos. La dificultad para representar mentalmente los estados afectivos y darles un significado que module el curso de la acción, deja al individuo expuesto al desborde. En relación con la interpretación de las otras mentes, se genera un estado de alerta constante en las relaciones interpersonales producto de las experiencias traumáticas en el vínculo temprano en que el cuidador era errático y violento. El resultado es una exacerbación de la necesidad de «leer» a los demás, en tanto en la infancia debieron anticiparse constantemente a las acciones del cuidador para poder sobrevivir a lo impredecible de su actuar. Sin embargo, la ansiedad y la necesidad de disminuir la tensión no generan las condiciones requeridas para captar de modo fiel las claves perceptuales que los lleven a atribuciones más realistas y fieles. Se da una *pseudomentalización*, que no es otra cosa que una estrategia de defensa para poder sobrevivir al intercambio interpersonal. Se observa, pues, un escenario en donde se detiene la capacidad interpretativa de la propia agencia y se exagera hiperactivamente la lectura de mentes con fines defensivos (Fonagy, 2008). La disfuncionalidad producto de esta situación y la incidencia de este tipo de problemáticas en poblaciones clínicas, ha generado la creación de formas adaptadas de psicoterapia psicoanalíticas que apuntan al desarrollo de las capacidades mentalizadoras detenidas en el desarrollo (Allen, Fonagy & Bateman 2008). La experiencia vincular que ofrece la psicoterapia constituye un nuevo escenario relacional suficientemente seguro en donde poder explorar el mundo mental.

Todo lo anterior apoya la idea de que la *mentalización* (capacidad central de la cognición social), no es un elemento garantizado en el desarrollo sino que depende dramáticamente de las condiciones relacionales de la vida temprana, básicamente de la seguridad del vínculo de apego. Esta capacidad puede detenerse en su desarrollo debido a situaciones vinculares específicas y logra restablecerse si se generan condiciones similares a las que debieron facilitar su despliegue. Es decir, depende decisivamente de experiencias relacionales muy particulares pues estas son la vía para su actualización en el desarrollo.

REFERENCIAS BIBLIOGRÁFICAS

- Allen, John, Peter Fonagy & Anthony Bateman (eds.) (2008). *Mentalizing in Clinical Practice*. Arlington: American Psychiatric Publishing.
- Bakermans-Kranenburg, Marian & Marinus van IJzendoorn (2009). The First 10 000 Adult Attachment Interviews: Distributions of Adult Attachment Representations in Clinical and Non-clinical Groups. *Attachment and Human Development*, 11(3), 223-263.
- Bjorklund, David & Anthony Pellegrini (2002). *The Origins of Human Nature: Evolutionary Developmental Psychology*. Washington DC: American Psychological Association.
- Bowlby, John (1969). *Attachment and Loss*. Volumen 1: *Attachment*. Nueva York: Viking Penguin.
- Carpendale, Jeremy & Charlie Lewis (2010). The Development of Social Understanding. A Developmental Perspective. En Willis Overton (ed.), *The Handbook of Life-span Development*. Volumen 1: *Cognition, Biology and Methods* (pp. 584-627). Nueva Jersey: John Wiley & Sons.
- Carruthers, Peter (2006). *The Architecture of the Mind*. Oxford: Oxford University Press.
- Carruthers, Peter (2009). How We Know Our Minds. The Relationship between Mind Reading and Metacognition. *Behavioral and Brain Sciences*, 32, 121-182.
- Cavell, Marcia (2006). *Becoming a Subject*. Nueva York: Oxford University Press.
- Davidson, Donald (2001). *Subjetivo, intersubjetivo, objetivo*. Madrid: Cátedra.
- Eilan, Naomi (2005). Joint Attention: Communication and Mind. En Naomi Eilan, Christoph Hoerl, Teresa McCormack y Johannes Roessler (eds.), *Joint Attention: Communication and Other Minds* (pp. 1-33). Oxford: Oxford University Press.
- Flinn, Mark & Carol Ward (2005). Ontogeny and Evolution of the Social Child. En Bruce Ellis y David Bjorklund (eds.), *Origins of the Social Mind. Evolutionary Psychology and Child Development* (pp. 19-44). Londres: The Guilford Press.
- Fonagy, Peter (2008). The Mentalization Approach to Social Development. En Fedric N. Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (pp. 3-56). East Sussex: The Analytic Press.
- Fonagy, Peter, György Gergely, Elliot Jurist & Mary Target (eds.) (2002). *Affect Regulation, Mentalization and the Development of the Self*. Nueva York: Other Press.
- Fonagy, Peter & Mary Target (1997). Attachment and Reflective Function: Their Role in Self Organization. *Development and Psychopathology*, 9, 679-700.
- Gergely, György (2007). The Social Construction of the Subjective Self: The Role of Affect-Mirroring, Markedness, and Ostensive Communication in Self-development. En Linda Mayes, Peter Fonagy y Mary Target (eds.), *Developmental Science and Psychoanalysis. Integration and Innovation* (pp. 42-82). Londres: Karnac Books.
- Gergely, György (2008). The Development of the Unreflective Self. En Fredric Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (pp. 57-102). East Sussex: The Analytic Press.

- Goldman, Alvin (2006). *Simulating Minds. The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.
- Gordon, Robert (1986). Folk Psychology as Simulation. *Mind and Language*, 1(2), 156-171.
- Holmes, Jeremy (2010). *Exploring in Security. Towards an Attachment Informed Psychoanalytic Psychotherapy*. East Sussex: Routledge.
- Iacoboni, Marco (2009). *Las neuronas espejo. Empatía, neuropolítica, autismo, imitación, o de cómo entendemos a los otros*. Madrid: Katz.
- LeDoux, Joseph (2001). *The Synaptic Self. How Our Brains Become Who We Are*. Nueva York: Viking Press.
- Main, Mary & Ruth Goldwyn (1998). *Adult Attachment Scoring and Classification System* [manuscrito]. Berkeley: Universidad de California.
- Meltzoff, Andrew & Keith Moore (1977). Imitation of Manual and Facial Gestures by Human Neonates. *Science, New Series*, 75-88. http://ilabs.washington.edu/melt-zoff/pdf/77Meltzoff_Moore_Science.pdf
- Perner, Josef (1991). *Understanding the Representational Mind*. Cambridge: The MIT Press.
- Premack, David & Guy Woodruff (1978). Does the Chimpanzee Have a Theory of Mind? *Behavioral and Brain Sciences*, 1, 515-526.
- Schore, Allan (2000). Attachment and the Regulation of the Human Right Brain. *Attachment & Human Development*, 2(1), 23-47.
- Siegel, Daniel (1999). *La mente en desarrollo. Cómo interactúan las relaciones y el cerebro para moldear nuestro ser*. Bilbao: Desclée de Brouwer.
- Stern, Daniel (1985). *El mundo interpersonal del infante. Una perspectiva desde el psicoanálisis y la psicología evolutiva*. Buenos Aires: Paidós.
- Stern, Daniel (2004). *The Present Moment in Psychotherapy and Everyday Life*. Nueva York: Norton & Company.
- Thompson, Ross (2006). The Development of the Person: Social Understanding, Relationships, Conscience, Self. En Nancy Eisenberg (ed.), *The Handbook of Child Psychology. Volumen 3: Social, Emotional and Personality Development* (pp. 26-41). Nueva Jersey: John Wiley & Sons.
- Tomasello, Michael, Melinda Carpenter, Josep Call, Tanya Behne & Henrike Moll (2005). Understanding and Sharing Intentions: The Origins of Cultural Cognitions. *Behavioral and Brain Sciences*, 28(5), 75-735.
- Trevarthen, Colwyn & Kenneth Aitken (2001). Infant Intersubjectivity: Research, Theory and Clinical Applications. *Journal of Child Psychology and Psychiatry*, 42(1), 3-48.
- Tronick, Edward (1989). Emotion and Emotion Communication in Infants. *American Psychologist*, 44(2), 112-119.
- Wellman, Henry (1990). *The Child's Theory of Mind*. Cambridge: The MIT Press.

INTERSUBJETIVIDAD Y ATRIBUCIÓN PSICOLÓGICA¹

Diana Pérez

Universidad de Buenos Aires – Consejo Nacional
de Investigaciones Científicas y Técnicas

Silvia Español

Facultad Latinoamericana de Ciencias Sociales –
Consejo Nacional de Investigaciones Científicas y Técnicas

1. INTRODUCCIÓN

En la literatura psicológica y filosófica reciente es usual toparse con el concepto de intersubjetividad. También es usual toparse, especialmente en el ámbito de la psicología del desarrollo, con la distinción entre intersubjetividad primaria, secundaria y terciaria. Parece bien establecido —como fruto de innumerables trabajos empíricos— que cada uno de estos modos de contacto intersubjetivo aparece en un cierto momento evolutivo claramente establecido y que la secuencia de hitos en el desarrollo implicados en estos tres modos de intersubjetividad está bastante clara. Sin embargo, desde nuestro punto de vista, no se ha logrado explicar satisfactoriamente el tránsito de un modo de intersubjetividad a otro. En parte, esto ocurre porque tampoco está del todo claramente especificado en qué consiste cada uno de estos modos de interacción intersubjetiva. En este trabajo no vamos a ocuparnos del problema del tránsito de la intersubjetividad primaria a la secundaria, sino que hemos englobado ambas bajo el término *formas básicas de intersubjetividad* y nos ocuparemos del tránsito desde estas hacia la intersubjetividad terciaria. Nuestra apuesta es que se puede lograr una caracterización apropiada de este pasaje si se adopta el marco de la perspectiva de segunda persona en la comprensión de la intersubjetividad y se busca una explicación

¹ Una versión inicial de este trabajo fue presentada en el *IV Coloquio Internacional Lenguaje, Contexto y Cognición*, realizado por Mente y Lenguaje, en Pisac, del 18 al 20 de julio de 2012. Agradecemos a los participantes del encuentro por sus valiosas preguntas y sugerencias y especialmente a César Escajadillo por su comentario crítico.

genética del desarrollo de la intersubjetividad que incorpore más explícita y detalladamente algunos elementos desatendidos en la literatura reciente, específicamente el juego de ficción y la capacidad narrativa. En este trabajo nos detendremos especialmente en el primero.

2. LAS FORMAS BÁSICAS DE LA INTERSUBJETIVIDAD

Dos son los supuestos básicos implicados en el clásico problema de las otras mentes: (i) que nuestros estados mentales nos resultan transparentes, tenemos un acceso directo, privilegiado a ellos; en cambio, (ii) que los estados de los otros nos resultan opacos y no accedemos a ellos de una manera directa. El «problema» —cómo nos damos cuenta de que los demás son seres con mente y le atribuimos deseos, intenciones, creencias si no tenemos un acceso directo a sus mentes— recibió dos respuestas clásicas: la simulacionista y la teórico-inferencial. Tales respuestas han sido denominadas también perspectiva de primera y de tercera persona debido a que, en el primer caso, se resalta el acceso privilegiado a la primera persona del singular («yo») en la atribución de estados mentales a los otros (y su proyección simulada en los demás); mientras que en el segundo el énfasis explicativo está puesto en la posibilidad de atribución, mediante procesos inferenciales o teóricos, a un otro objetivado y distante, un «él». Pese a sus diferencias y aparente oposición, ambas suponen un mismo punto de partida: la idea de que la mente de los otros es opaca y no puede ser percibida directamente o, en otras palabras, la creencia de que no es posible tener una experiencia directa del otro.

En los últimos diez años se ha desarrollado una nueva aproximación teórica que reformula el punto de partida mismo desde el cual suelen plantearse estas dos teorías, que ha dado en llamarse «perspectiva de segunda persona»². La perspectiva de segunda persona pone el énfasis en la relación «yo-vos» y resalta que los modos básicos de estar y comprender a los otros, y a uno mismo, se basan en un conjunto de habilidades para la comprensión recíproca que se desarrolla y expresa en contextos interactivos, por ende, públicos y sociales. Asume que en la vinculación con el otro intervienen componentes que tienen una directa expresión pública, como expresiones faciales o posturales y diversos patrones conductuales. Y considera que los aspectos expresivos —como el tono de voz o la configuración facial— son percibidos como directamente significativos y constituyen la base de una reacción correspondiente que tiene pleno sentido dentro de la situación de una interacción. Lo que un sujeto hace tiene una respuesta sensible en el otro, de tal suerte que acciones y reacciones tienen

² Para argumentos en su favor véanse Gallagher (2001), Gomila (2002) y Hutto (2008), entre otros.

los rasgos de la reciprocidad. Se destaca también que la atribución mental en contextos interactivos no solo es recíproca sino que además los participantes se percatan de su mutua atribución, lo cual condiciona el contenido de la atribución. Lo mental no es por tanto considerado algo privado sino público, expresivo y dinámico.

La distinción entre las perspectivas de primera, tercera y segunda persona apunta a destacar los posibles modos de acercamiento a cierto fenómeno —el de la comprensión cotidiana de las mentes humanas que adoptan los teóricos, tanto dentro de las diversas aproximaciones filosóficas como psicológicas. Paradigmáticamente, quienes han destacado el rol de la empatía, como Collingwood, o de la simulación, como Gordon, Heal y Currie en la literatura filosófica, son defensores de una perspectiva de primera persona, tanto como Harris, que pone en primer plano la imaginación desde la psicología del desarrollo. Quienes han destacado el rol de las teorías en la comprensión de las mentes, como Lewis o Fodor, y, en general, los funcionalistas en el ámbito filosófico o Gopnik y Wellman en el ámbito de la psicología, son defensores de la tercera persona. Y, finalmente, quienes han destacado el carácter expresivo de la conducta humana y el rol del cuerpo en las interacciones psicológicas, como Wittgenstein y Gallagher en el ámbito filosófico o Reddy y Gómez en el ámbito psicológico, pueden considerarse defensores de una perspectiva de segunda persona³.

Ahora bien, los teóricos de la segunda persona suelen recuperar la teoría de la intersubjetividad de Trevarthen para quien su propuesta «[...] no es nada menos que una teoría de cómo las mentes humanas, en los cuerpos humanos, pueden reconocer los impulsos del otro de forma intuitiva, con o sin elaboraciones cognitivas o simbólicas [...]» (1998, p. 17). Aunque algunos autores toman solo la idea general de intersubjetividad, la mayoría incorpora su clásica distinción ontogenética entre intersubjetividad primaria y secundaria (Trevarthen, 1979 y 2000; Trevarthen & Hubley, 1978; Trevarthen & Reddy, 2007).

³ Es importante destacar que los defensores de la segunda persona no han cuestionado la idea de una asimetría en la *atribución* de sensaciones, ideas o sentimientos a nosotros mismos y a los otros, pese a que la asunción del carácter público y observable de los estados mentales de los otros torna plausible la posibilidad de aprehenderlos directamente. Tal idea suele ejemplificarse con el clásico caso del concepto *dolor*. Se dice que los criterios o las fuentes de evidencia a los que apelamos a la hora de afirmar juicios de dolor en uno u otro caso difieren: (i) atribuimos estados psicológicos, como el dolor, a los otros sobre la base de evidencia observacional, conductual, y lingüística; (ii) nos atribuimos dolor a nosotros mismos sin apelar a evidencia alguna, simplemente porque (causal) sentimos dolor. Es decir, nuestros estados mentales nos resultan transparentes, tenemos un acceso directo, privilegiado a ellos; en cambio, los estados de los otros, si bien no nos resultan definitivamente opacos (no estamos reducidos a acceder a ellos vía simulación o metarrepresentación) no se nos dan de la misma manera que los propios; por lo que en este caso nos basamos en elementos observacionales, conductuales, lingüísticos o no, para hacer atribuciones. La cuestión de la atribución será tratada en el próximo apartado.

La propuesta clásica de Trevarthen apunta a la identificación de periodos evolutivos en el desarrollo de las capacidades de relación interpersonal, cognitivas y afectivas, de los niños. Su descripción de las características propias de cada uno de estos momentos evolutivos, cuando es tomada en consideración, es aceptada por todas las perspectivas mencionadas arriba: hay consenso entre los teóricos de las tres perspectivas mencionadas (tanto entre filósofos como psicólogos) respecto de la naturaleza de estas interrelaciones. Repasemos brevemente la descripción clásica de lo que nosotras denominaremos *formas básicas de la intersubjetividad*, expresión que usaremos para englobar la intersubjetividad primaria y la secundaria.

De acuerdo con Trevarthen, la *intersubjetividad primaria* refiere al modo más primitivo de conexión psicológica que caracteriza los primeros meses de vida del bebé. En tanto el bebé realiza conductas que evocan respuestas en el otro y viceversa, puede hablarse de intersubjetividad. Se trata de una conexión diádica, basada en la atención mutua, en la que se comparten estados emocionales. En estos intercambios puede observarse muy tempranamente una alternancia de turnos en las vocalizaciones emitidas por la díada. Por tal motivo se los denominó protoconversaciones. Hasta alrededor de los dos meses, la toma de turnos se establece a la manera de un lento adagio; poco después, el pulso que comparten en sus juegos vocálicos se acelera a un andante y los encuentros más animados semejan cada vez más juegos que protoconversaciones. Estos intercambios incluyen frecuentemente episodios o ciclos de imitación mutua; en ellos puede observarse también una sincronía interactiva en la que toma lugar un ‘reflejar simpático’ (*sympathetic mirroring*), una correspondencia transmodal simpática, por ejemplo, del ritmo de las vocalizaciones de la madre con los movimientos de brazos del bebé (Condon & Sander, 1974, citados por Trevarthen, 2000). Actualmente, estos casos de reflejar simpático se describen como uno de los tantos modos de acoplamiento (transmodales o no) de los ritmos y los gestos sonoros y motores de uno y otro que se engloban bajo el término de musicalidad comunicativa (Malloch & Trevarthen, 2009; Stern, 2010). Por ejemplo, Malloch y Trevarthen analizaron los intercambios vocálicos entre una beba de tres meses y su madre y encontraron una reiterada y particular imitación vocálica de la madre: después de cada emisión de la beba, las tres medidas del timbre de la madre bajaban: la madre hacía su voz más parecida a la del bebé. Con ese solo recurso, le señalaba que la estaba escuchando, que estaba ahí, con ella, atenta y dispuesta al intercambio. Este caso es considerado el ejemplo paradigmático de musicalidad comunicativa. Estos intercambios no incluyen más que las emanaciones de uno a otro y suponen esencialmente estados de simpatía, de «sentir con» el otro; en ellos, los objetos, el mundo externo entero, queda fuera.

El tránsito hacia la *intersubjetividad secundaria* ocurre alrededor de los nueve meses, cuando el bebé comienza a compartir el interés por los objetos y eventos del mundo. Sus primeras manifestaciones son el fenómeno de atención conjunta (en el que niño y el adulto alternan la mirada hacia algún objeto de interés y hacia el compañero de interacción) y las primeras pautas preverbales de comunicación intencionada, especialmente las pautas protodeclarativas de comunicación (mediante las cuales el niño dirige la atención del adulto hacia un objeto con el mero fin de compartir la experiencia con el otro) que emergen alrededor de los doce meses. En particular, el gesto de señalar con función declarativa es considerado la primera referencia motora. El periodo de intersubjetividad secundaria da cuenta de un momento evolutivo particular —desde la atención conjunta hacia las pautas preverbales comunicativas durante el último trimestre del primer año de vida—, en el que el contacto psicológico se establece entre sujetos que dirigen conjuntamente su atención a objetos de referencia compartidos. A diferencia de las relaciones diádicas, sujeto-sujeto, propias de la intersubjetividad primaria, la relación ahora es triangular, se establece una relación sujeto-sujeto-objeto en la que hay dos sujetos con estados psicológicos claramente intencionales (de *aboutness*) y un mundo compartido de referencia y manipulación.

Ahora bien, aunque la distinción entre intersubjetividad primaria y secundaria es la más reconocida y usada en la literatura psicológica y filosófica, algunos autores, próximos a la perspectiva de segunda persona, han descrito e incluido en su tipología otros modos de intersubjetividad. Por ejemplo, Gómez (1998) distingue entre un modo de intersubjetividad de segunda persona, en la que dos subjetividades⁴ entran en contacto y son recíprocamente conscientes de la conciencia ajena (el ejemplo que da implica la atención mutua a objetos, por lo que, en principio, parece restringirse a la intersubjetividad secundaria) y la intersubjetividad de tercera persona, en la que una subjetividad toma otra subjetividad como objeto, no se requiere contacto y se piensa acerca de lo que el otro tiene en mente. Así también Bråten y Trevarthen (2007) distinguen, además de la intersubjetividad primaria y secundaria, un nivel de la terciaria en el que hay relaciones intersubjetivas reflexivas y recursivas, en el sentido de una comprensión comunicativa mediada por metarrepresentaciones que incluye referencia simbólica a mundos reales o ficcionales de la imaginación o simulación conjunta, y en el que hay un habla narrativa, la cual implica predicación y un sentido verbal o narrativo de sí mismo y del otro. Nótese que en tal intersubjetividad terciaria las capacidades metarrepresentacionales no se separan de la comprensión comunicativa, es decir, de los procesos anclados en el aquí y ahora de la interacción.

⁴ Por *subjetividad* hay que entender simplemente la forma en que percibimos, pensamos o sentimos acerca del mundo.

Con el fin de abordar el problema del tránsito hacia la intersubjetividad terciaria, proponemos reorganizar las varias distinciones mencionadas de la siguiente manera: (i) englobar las experiencias de intersubjetividad primaria y secundaria bajo el rótulo de *formas básicas de intersubjetividad*, y (ii) distinguirlas de las *formas no básicas de la intersubjetividad*, que incluyen como modo prototípico el pensar distanciado, meta-representacional, hacia los contenidos opacos de la mente del otro, en particular sus creencias, pero que puede abarcar otros modos intermedios de encuentros mediados por signos complejos, símbolos o representaciones de segundo orden, así como comprensión comunicativa mediada por metarrepresentaciones.

3. LA ATRIBUCIÓN PSICOLÓGICA Y LAS FORMAS NO BÁSICAS DE LA INTERSUBJETIVIDAD

En filosofía hay una larga tradición que se remonta al silogismo práctico aristotélico para dar cuenta de los mecanismos de atribución psicológica. Este esquema clásico de atribución psicológica involucra la explicación de la acción por deseos y creencias, es decir por estados de actitud proposicional. Los conceptos de actitud proposicional se aplican a estados con contenido, esto es con condiciones de verdad, con propiedades semánticas. Estos estados involucran cuatro elementos: 1. Sujeto / organismo/ individuo /sistema, 2. actitud psicológica (de creencia, deseo, etcétera), 3. condiciones de verdad (los estados de cosas del mundo representado), y 4. un modo de presentación del mundo, o contenido proposicional.

La tradición filosófica ha tomado como paradigma de concepto de actitud proposicional el de *creencia*. Quien posee el concepto de creencia comprende que aquel al que se le atribuye una creencia puede representar el mundo verdadera o falsamente, es decir que el mundo puede o no coincidir con nuestras creencias o, en otros términos, que puede haber un desacople entre el mundo y lo creído (3.), que explicaría la acción fallida. Asimismo, quien tiene el concepto de creencia es capaz de darse cuenta de que dos individuos pueden referirse al mismo estado de cosas en el mundo sin saberlo, por desconocer las identidades relevantes, es decir, comprende que las atribuciones de creencias involucran opacidad referencial (4.).

Esta comprensión de las actitudes proposicionales permite explicar la acción intencional a partir de las peculiares conexiones entre el mundo deseado y el mundo representado por las creencias del agente. Es porque un individuo desea que p , y cree que solo se dará p en caso de que haga q , que el individuo está dispuesto a hacer q . Nótese que en este caso paradigmático deben atribuirse estados con *contenido proposicional* para que la explicación de la acción pueda llevarse a cabo, dado que la inferencia práctica depende de inferencias realizadas entre los contenidos de los

estados psicológicos, contenidos que para entrar en genuinas relaciones inferenciales deben ser poseedores de valores de verdad. Así, entendiendo que los intercambios interpersonales están basados en la posibilidad de comprender y anticipar las acciones ajenas, los teóricos clásicos han focalizado sus estudios en la capacidad de atribución de creencias a nuestros congéneres y, en particular, la posibilidad de atribuir creencias falsas para dar cuenta de acciones fallidas. En efecto, la atribución de conocimiento no permite explicar adecuadamente *toda* acción humana: conocer es un verbo de logro y, por tanto, si uno actuara solo en función de lo que conoce, jamás podría realizar acciones ineficaces, porque en todas nuestras acciones el mundo coincidiría con nuestra representación del mundo. Pero somos humanos, imperfectos, muchas veces fallamos, y la única manera de explicar esto es apelando a estados psicológicos como el de creencia, en el que hay posibilidad de un desacoplamiento entre el contenido proposicional creído y el estados de cosas efectivos del mundo. Por ello el «test de falsa creencia» ha sido la piedra de toque para establecer en qué momento los niños han logrado una acabada comprensión de las mentes humanas y, con ello, de las interacciones interpersonales característicamente humanas.

En las primeras formulaciones del test (Wimmer & Perner, 1983; Baron-Cohen, Leslie & Frith, 1985) pareció quedar establecido que los niños eran capaces de atribuir creencias falsas hacia los cuatro años y medio. En trabajos más recientes (Southgate, Senju & Csibra, 2007), se ha mostrado que niños más pequeños (de veinticinco meses) también son capaces de pasar versiones reformuladas del test⁵. No vamos a detenernos a evaluar todos los problemas metodológicos que surgen de las diferentes formulaciones del test y de la interpretación de los datos recogidos ya que no importa —para los propósitos de este trabajo— cuándo se da exactamente el pasaje (si a los dos años, como afirman los estudios más recientes, o los cuatro, como dicen los clásicos); lo importante es que hay un salto entre la atribución psicológica propiamente dicha (es decir, la atribución de estados psicológicos con contenidos proposicionales) y los modos de contacto psicológicos anteriores respecto de los cuales cabría preguntarse si corresponde considerarlos casos en los que hay involucradas atribuciones psicológicas.

En efecto, si por *atribución psicológica* vamos a entender solamente aquellos casos que se ajustan al modelo de las actitudes proposicionales, hay una infinidad de situaciones en las cuales no tratamos a los demás como objetos, pero tampoco los tratamos como sujetos que exhiben estados de actitud proposicional. Consideremos el caso de quien actúa de cierta manera porque está celoso o quien reacciona

⁵ Algunos autores incluso aventuran edades más tempranas (quince meses, Onishi & Baillargeon, 2005). Sin embargo estos experimentos han sido criticados porque puede darse una explicación más simple de los desempeños de los niños sin atribuir creencias falsas (Southgate, Senju & Csibra, 2007).

ante una situación peligrosa escondiéndose o quien grita de dolor. En estos casos no resulta indispensable atribuirle creencias y deseos a la persona para comprender su acción. Cuando poseemos maestría en el uso de conceptos psicológicos la tentación a usarlos es muy grande (tal vez inevitable). Por lo que probablemente lo haríamos sin dudar, si la conducta en cuestión es desplegada por un adulto. Pero dudaríamos en hacerlo si estamos tratando de comprender la conducta de un bebé, de un niño pequeño o de un animal no humano.

Los defensores de la segunda persona, por ejemplo Gomila (2002), han aceptado que hay mecanismos de atribución psicológica que se ponen en juego en las interacciones humanas en las que participan individuos menores al año, es decir, en lo que llamamos formas básicas de intersubjetividad, más específicamente, en el tipo de interacciones recogidas en los triángulos propios de la intersubjetividad secundaria (periodo en el que, de acuerdo con todos los test realizados hasta el momento, los niños no son capaces de pasar ninguna versión del test de falsa creencia, por lo que no resulta adecuado atribuirles la posesión del concepto de creencia). En efecto, se considera que este tipo de interacción desplegada hacia los doce meses está regida por la reactividad emocional ante la expresión (natural) de la emoción ajena, así como la imitación de la acción intencional ajena y la comprensión de una intención inscrita en la acción. Así, parece que sería apropiado considerar que el bebé que reacciona ante la expresión de miedo de la figura de crianza pone en marcha un mecanismo de atribución de estados psicológicos, en este caso de atribución de estados emocionales. De la misma manera, cuando un bebé percibe que alguien intenta agarrar un objeto, o alterna la mirada entre los ojos de un adulto y los de un objeto, señalándolo y emitiendo sonidos, podría pensarse que el bebé está poniendo en juego ciertos mecanismos de atribución psicológica, comprendiendo la acción del adulto como una acción motivada por el deseo de satisfacer las necesidades del niño, o por la creencia de que el niño necesita aquello que está señalando. Y usualmente nosotros, los adultos que acompañamos al bebé, no dudamos en pensar de esta manera.

Sin embargo, no parece indispensable dar cuenta de los mecanismos psicológicos puestos en juego por el bebé apelando a mecanismos de atribución psicológica «maduros». Es más, si —como mencionamos arriba— hay buenas razones para pensar que a los doce meses ningún bebé testeado con ninguno de los paradigmas experimentales disponibles ha pasado exitosamente el «test de falsa creencia», es razonable pensar que el bebé no ha puesto en juego mecanismos de atribución de estados proposicionales complejos como la creencia. Y tampoco parece necesario suponer que sea conveniente atribuir al bebé la comprensión de ningún concepto de actitud proposicional, es decir, de ningún concepto de estado psicológico con contenidos con las características que poseen los contenidos de las creencias (3. y 4.).

Por ello, a nuestro juicio, resulta inadecuado considerar que en estos casos existe atribución psicológica entendida de la manera en que se caracteriza clásicamente. Dicho de otra manera, en el nivel de la intersubjetividad secundaria consideramos que no hay propiamente hablando procesos de atribución psicológica, es decir, de atribución de estados con contenido proposicional al otro ni a uno mismo.

Podríamos preguntarnos si en estos triángulos existe algún tipo de atribución psicológica o, por el contrario, no hay nada que así pueda denominarse. Puesto en otros términos, la pregunta podría plantearse así: ¿cuáles son los conceptos psicológicos disponibles a un bebé que participa en las formas básicas de la intersubjetividad? Pensamos que, si hubiera algo parecido, esto es, si pudiera afirmarse que los niños involucrados están en posesión de algunos conceptos psicológicos, solo se podría considerar que los sujetos poseen conceptos de algunas sensaciones corporales (por ejemplo, dolor), de algunas emociones con contenido objetual pero no proposicional (por ejemplo, miedo al león), del concepto de deseo hacia un objeto (por ejemplo, querer un yogur) y del concepto de atención/intención dirigida a un objeto (por ejemplo, mirar el trompo, agarrar el vaso, etcétera).

Ahora bien, es claro que las interacciones prototípicas de la intersubjetividad terciaria solo son posibles si hay atribución de estados psicológicos con contenido proposicional. En estos casos el sujeto tiene maestría no solo de los conceptos involucrados en el contenido proposicional sino también en el concepto de la actitud psicológica involucrada: es decir debe poseer el concepto de creencia que *p*, deseo que *p*, intención que *p*, miedo que *p*; en otras palabras, debe ser capaz de manejar metarrepresentaciones (tiene que representarse a alguien como un individuo que maneja representaciones, es decir, que desacopla el mundo pensado del mundo real). El contenido de estas actitudes psicológicas, dado su carácter proposicional, tiene a su vez tres características ausentes en las atribuciones de estados psicológicos con contenidos no-proposicionales, como las mencionadas más arriba, que resultan ser las propias de las formas básicas de la intersubjetividad. Nuestra propuesta acerca de las tres características ausentes en las atribuciones de estados psicológicos con contenidos no-proposicionales podría considerarse una generalización a otros estados psicológicos (como deseo e intención) de la distinción trazada por Dretske entre el 'ver simple' (*simple seeing*) y el 'ver epistémico' (*epistemic seeing*) (tomado de Bermúdez, 2003). El ver simple es caracterizado de la siguiente manera:

- (1) toma complementos no-proposicionales,
- (2) crea contextos transparentes,
- (3) es hecha verdadera por relaciones perceptuales directas entre los perceptores y los particulares.

En cambio, el ver epistémico se caracteriza porque:

- (1') toma complementos proposicionales,
- (2') crea contextos opacos,
- (3') es hecho verdadero por relaciones epistémicas basadas perceptualmente entre perceptores y proposiciones.

Extrapolando esta idea podríamos sostener, entonces, que en la atribución psicológica prototípica/clásica hay atribución de estados psicológicos con contenido proposicional, es decir con un contenido que:

- (1) tiene valor de verdad, es decir es V o F,
- (2) tiene una estructura predicativa, y
- (2) está constituidas por conceptos, es decir, elementos constituyentes recombinables.

Por el contrario, en el caso de las atribuciones psicológicas sin contenido proposicional, estas características están ausentes, ya que se trata de estados con contenidos en los que:

- (1') solo puede hablarse de satisfacción o no,
- (2') no hay predicación, y
- (3') puede no haber conceptos involucrados, es decir, puede haber una actitud dirigida directamente hacia un objeto en el mundo, sin mediación de ningún modo de presentación bajo el cual se establece esta direccionalidad y, por tanto, sin opacidad alguna.

El desafío que nos planteamos, entonces, es avanzar en la explicación evolutiva del pasaje de los modos no proposicionales de atribución presentes en las formas básicas de intersubjetividad a los modos canónicos y proposicionales de atribución psicológica que aparecen hacia los cuatro a cinco años de vida.

Como es bien sabido, Davidson en su trabajo *Subjective, Intersubjective, Objective* agradeció no ser un psicólogo del desarrollo:

La dificultad para describir la emergencia de los fenómenos mentales es un problema conceptual: es la dificultad para describir los estadios más tempranos en la maduración de la razón, los estadios que preceden a la situación en la cual conceptos como intención, creencia y deseo tienen una aplicación clara. Tanto en la evolución del pensamiento en la historia de la humanidad, como en la evolución del pensamiento de cada individuo, hay un estadio en el cual no hay

pensamiento seguido de otro en el cual hay pensamiento. Describir la emergencia del pensamiento sería describir el proceso que lleva del primero al segundo de estos estadios. [...] Agradezco no estar en el campo de la psicología del desarrollo (2001, pp. 127-128; la traducción es nuestra).

La dificultad conceptual con la que se topó Davidson (y de la que nunca pudo salir) se deriva, a nuestro juicio, de dos tesis que defendió profusamente: (i) que solo pueden atribuirse creencias a quien posee los conceptos de creencia y de verdad objetiva, y (ii) que todos los verbos psicológicos son —explícita o implícitamente— verbos de actitud *proposicional*. Si nuestra propuesta conceptual desarrollada hasta aquí es correcta, ninguna de estas dos tesis debe ser aceptada. En efecto, poseer una creencia no implica tener el concepto de creencia ni de verdad objetiva, aunque seguramente tener el concepto de creencia va de la mano de tener el concepto de verdad objetiva, tal como sostuvo este autor (2001, pp. 95-105). Los niños en cierto momento —entre los dieciocho meses y los cuatro años— tienen creencias acerca del mundo pero no tienen el concepto de creencia ni de verdad objetiva, ya que solo son capaces de comprender y atribuir estados psicológicos con contenido no-proposicional. La posibilidad de atribuir cierto tipo de estados psicológicos como deseos o intenciones sin contenido proposicional, transparentes, dirigidos directamente hacia objetos del mundo permite identificar verbos psicológicos que no están conceptualmente atados a creencias ni a otros verbos de actitud proposicional, mostrando así que no todas las atribuciones psicológicas involucran contenidos proposicionales.

Nuestra impresión es que, en la medida en que hemos alcanzado el logro cognitivo consistente en poner en marcha mecanismos metarrepresentacionales, se vuelve inevitable ponerlos en funcionamiento cada vez que observamos acciones de seres animados. Pero, sin embargo, en muchos casos resulta perfectamente viable «reescribir» las escenas que presenciamos, por ejemplo, entre bebés prelingüísticos y adultos, sin apelar a metarrepresentaciones para describir las acciones del bebé. Y también podríamos ensayar estas reescrituras cuando consideramos ciertos peculiares intercambios entre adultos: por ejemplo, entre los artistas y el público en el caso de aquellas artes que no implican contenidos proposicionales, como la música y la danza, y también en nuestras experiencias interpersonales íntimas, especialmente en encuentros amorosos.

4. EL PASAJE

En este apartado, una vez establecidos los tipos de atribución psicológica involucrados en los distintos modos de intersubjetividad, trataremos de avanzar en la aún incompleta explicación genética del pasaje de un tipo de atribución a la otra.

La hipótesis que queremos presentar en esta ocasión es que hay dos elementos que aparecen durante el segundo y tercer año de vida, respectivamente, que cumplen un papel importante en la explicación de este pasaje y que permanecen en juego en nuestra vida adulta paradigmáticamente cuando consumimos literatura (y otras formas asociadas de artes narrativas: cine, teatro, etcétera). Se trata del juego de ficción (y funcional), es decir, de la capacidad de simular/pretender, y de la capacidad narrativa, esto es, de la capacidad de organizar el mundo social de forma narrativa, de comprender y producir narraciones). En esta ocasión nos ocuparemos con detalle del primero.

Es llamativo que la actividad lúdica ficcional de los niños haya sido desatendida en los modos habituales de tratar la cuestión del desarrollo de la atribución psicológica cuando, en el ámbito de la psicología del desarrollo, desde mediados de siglo XX, el juego de ficción (i) es considerado una de las claves para el desarrollo de las capacidades simbólicas, (ii) es considerado precursor de la teoría de la mente y (iii) es el hito evolutivo que aparece en el desarrollo inmediatamente después del establecimiento de la intersubjetividad secundaria.

En psicología del desarrollo, en el marco de la psicología cognitiva clásica, fue Leslie (1987 y 1994) quien llamó la atención sobre el juego de ficción como un indicador de la presencia temprana de capacidades metarrepresentacionales. En efecto, fue el primero en establecer un vínculo claro, aunque discutible, entre las habilidades mentalistas y el juego de ficción. Su intuición consistió en suponer que existe un isomorfismo profundo entre las propiedades de la ficción y la intencionalidad propia de los enunciados con verbos de creencia. Según Leslie, el dominio de la «teoría de la mente» es el resultado del funcionamiento de un módulo, de una competencia cognitiva relativamente específica e independiente de otras, que se define por la capacidad de construir y manipular metarrepresentaciones, las cuales son producto de un tipo especial de operación cognitiva —el desacoplamiento— de origen innato. Seguidamente, el mecanismo de desacoplamiento es el resultado de un potencial prefigurado por información genética que está programado para desatarse durante el segundo año de vida: su primera manifestación es el juego de ficción. Leslie considera que el desarrollo del módulo mentalista consiste en un proceso de desacoplamiento que forma un nuevo plano representacional cuya fase crítica se sitúa entre los dieciocho meses y los cinco años. ¿Cómo es posible —pregunta— que un niño piense una banana como siendo un teléfono, o una taza vacía como conteniendo té? Si el sistema representacional está en desarrollo, ¿cómo puede tolerar estas distorsiones en las relaciones semánticas? Para que el juego de ficción sea ficción y no un error categorial es necesario mantener un nivel de representaciones primarias, que representan directamente objetos y situaciones en el mundo; pero, además, ha de incluirse otro orden representacional que no involucra representaciones, sino representaciones

de representaciones. Ellas son metarrepresentaciones o representaciones secundarias en las que han dejado de regir las relaciones ordinarias de referencia y verdad que existen entre las representaciones primarias y las cosas. El rasgo básico de su modelo es la creación de la ficción mediante el copiado de una representación primaria dentro de un contexto metarrepresentacional. Este contexto vuelve opaca la expresión que era primariamente transparente; las relaciones de referencia, verdad y existencia de la representación se suspenden mientras aparecen en este contexto. El contexto metarrepresentacional desacopla la representación primaria de sus relaciones de referencia, verdad y existencia; mientras tanto, la representación primaria original —una copia de la cual fue llevada a un segundo orden— continúa con su referencia literal y sus relaciones de verdad y referencia. Dada las características de su modelo, Leslie necesita marcar las expresiones desacopladas y decide hacerlo utilizando comillas («...»).

Así, este autor sugiere que a cada una de las propiedades semánticas de las expresiones de creencia le corresponde una forma básica de juego: (1) la opacidad referencial recuerda el proceso de sustitución de objeto, (2) el no compromiso de verdad se relaciona con la posibilidad de atribuir propiedades ficcionales, no reales a los objetos, (3) el no compromiso de existencia recuerda los juegos en que se definen objetos puramente imaginarios. La razón de este isomorfismo reside en que ambos operan con metarrepresentaciones.

A continuación, intenta responder otra pregunta: ¿quién desacopla la expresión? Para responderla recurre al lenguaje natural que contiene términos que denotan la relación entre agentes y proposiciones: *pretend* ('simular' o 'ficcional'), es uno de esos términos. Adhiere, entonces, a su modelo formal un elemento que representa la relación informacional. La metarrepresentación de ficción tendría entonces la siguiente forma general: agente-relación informacional-«expresión». En esta forma el agente puede ser «yo», «Juan», «mi madre»; por lo que el mecanismo que permite la comprensión del juego propio es el mismo que explica la comprensión del juego de los demás. La relación informacional es «fingir o simular» y la representación desacoplada es algo del tipo «que esto es un avión». Como la expresión desacoplada ya no refiere directamente, no tiene una relación directa al mundo, decide no usar el término *referencia* y utiliza el término *anclaje*. La idea es que las expresiones desacopladas no tienen referencia a objetos, entonces ellas están ancladas a partes de la representación primaria (en nuestro caso, el avión podría estar anclado a un lápiz, por ejemplo). Y señala, la «relación informacional» es necesaria porque la expresión desacoplada no se relaciona automáticamente con el sistema de representación primaria. Las relaciones informacionales pueden ser vistas como funciones computacionales que relacionan el agente, la expresión desacoplada y la representación primaria.

La propuesta modular e innatista de Leslie está en las antípodas de la teoría de la intersubjetividad. Asimismo, hay muchos trabajos que han apuntado a mostrar que el juego de ficción no presupone capacidades metarrepresentacionales como Leslie sostiene (véanse por ejemplo, Jarrold y otros, 1994; Currie, 1998). Sin embargo, trataremos de argumentar, es posible proponer una aproximación al juego como eslabón del pasaje entre las formas básicas de la intersubjetividad y la atribución psicológica madura adoptando una perspectiva de segunda persona, que contemple el cambio y la novedad en el desarrollo y no se comprometa con la presencia de metarrepresentaciones como elemento fundante de la capacidad ficcional, sino que, a la inversa, considere que es la capacidad de involucrarse en determinados tipos de juego lo que da lugar a la aparición de las capacidades metarrepresentacionales. Así, la falta de juego de ficción (que puede observarse a veces en algunos niños) no sería un síntoma de una falla en las capacidades metarrepresentacionales, sino que por fallar la motivación y/o capacidad para involucrarse en cierto tipo de juegos se retrasa o dificulta la adquisición de las capacidades metarrepresentaciones. No daremos un argumento concluyente a favor de esta tesis (básicamente porque no creemos que sea posible un argumento tal ni para uno ni para el otro lado). Nos enfocaremos en contar una historia plausible del pasaje que nos ocupa adoptando este supuesto, que, esperamos, resulte más iluminadora que la propuesta alternativa.

Pese a su impronta modular y a su hipótesis de un mecanismo desacoplador innato, la teoría de Leslie ha atraído la atención de autores cercanos a la perspectiva de segunda persona como Rivière y Hobson. De acuerdo con Rivière (2003[1999] y 2003[2000]), el juego de ficción señala la aparición de un nuevo orden representacional: el niño mantiene en su conciencia dos niveles de representación, una literal y otra que lo sitúa en un plano nuevo que implica el uso de ciertas formas específicas de representación, representaciones entrecomilladas o suspendidas. Y en esto acuerda con Leslie. Pero este orden representacional es similar, no igual, al que se someten los enunciados con verbos de creencias. No son representaciones de relaciones representacionales, no requieren una representación tacita de un agente, una relación representacional y una expresión entrecomillada. Al jugar, el niño puede hacer que «A es Z» sin necesidad de la cláusula «yo finjo que A es Z». «Ello permite que la mente del niño “se despegue” progresivamente —sin por ello perder suelo de las realidades inmediatas y de la necesidad de acomodación literal de representación de ellas» (Rivière, 2003[1999], p. 228). Entre las representaciones suspendidas y las metarrepresentaciones, en sentido estricto, media una amplia distancia evolutiva. Las segundas solo son asequibles alrededor de los cuatro años y medio, momento en que los niños se muestran capaces de resolver la tarea de falsa creencia. Pero, sin embargo, y este es el punto en el que rescata lo que denomina «la genial intuición

de Leslie», ambas representaciones tienen algo en común: ambas están desligadas de la exigencia de sumisión adaptativa a la realidad a la que se someten las representaciones primarias que el hombre comparte con otros animales; tanto en una como en otra han quedado «en suspenso» las relaciones ordinarias de referencia y verdad que existen entre las representaciones primarias y las cosas o eventos del mundo. En realidad, sugiere Rivière, la cualidad esencial de los enunciados de creencia, ese rasgo de intencionalidad que suspende compromisos de verdad, de existencia y de transparencia referencial, no es más que una consecuencia de la suspensión de mundos representados que se inicia con el juego de ficción. Es en este sentido que el juego de ficción es un precursor del Sistema de Teoría de la Mente (pero no porque comparta el mismo tipo de requisitos representacionales) y su emergencia no es consecuencia de un mecanismo desacoplador de carácter innato. El niño puede fingir realidades alternativas, ver una cosa y actuar independientemente de lo que ve, gracias al desarrollo de los conocimientos que ha ido adquiriendo en sus interacciones con los otros, gracias a su impulso o tendencia hacia la interacción con otros y gracias a un mecanismo semiótico que viene operando desde momentos evolutivos previos. Es el desarrollo de la acción y la interacción lo que lo torna posible el juego de ficción y no un mecanismo innato preprogramado para dispararse a los dos años.

La distinción entre representaciones suspendidas y metarrepresentaciones no es cuestión de poca monta. Indica, aunque Rivière no lo diga explícitamente, que puede existir sustitución de un objeto por otro, pura y genuina ficción, sin lenguaje o, mejor dicho, sin la predicación inherente al lenguaje, y sin conciencia de la actitud proposicional que se mantiene con la situación. Su teoría no se compromete con la presencia de metarrepresentaciones como elemento fundante de la capacidad ficcional, al contrario, considera que es la capacidad de involucrarse en determinados tipos de juego lo que da lugar a la aparición de las capacidades metarrepresentacionales.

Nuestra propuesta consiste en ahondar en este camino abierto por Leslie y Rivière destacando algunos hechos que a nuestro juicio permiten avanzar en la explicación del pasaje desde las formas básicas de la intersubjetividad a la atribución psicológica madura.

Dado que estamos adoptando como punto de partida una perspectiva de segunda persona, queremos enfatizar el hecho de que cada logro cognitivo que alcanza el niño no debe considerarse como algo que ocurre en una interna y aislada mente cartesiana, sino que ocurre en la medida en que el niño en relación con otros es capaz de involucrarse en cierta actividad y *al mismo tiempo* es capaz de reconocer que otro individuo (la figura de crianza, por ejemplo) también está involucrado en exactamente la misma actividad. Veamos esto con un ejemplo concreto: una niña A (quince meses, trece días) recibe de regalo un juego de té. La figura de crianza abre el paquete,

agarra la taza de juguete, hace que revuelve, toma el líquido que no existía, hace ruido y dice «¡Qué rico!». Inmediatamente A empieza a hacer lo mismo. Lo que queremos destacar es que A entendió qué era lo que el adulto estaba haciendo y empezó a hacerlo en lo sucesivo, es decir, A entiende a la vez qué hace ella misma y qué hace el otro (no se da uno antes que el otro). Y si en estas acciones hay involucradas adscripciones de estados psicológicos, esto es, si resulta correcto decir que A comprendió que el adulto estaba jugando a tomar el té, que el adulto quiere jugar, que pretende tomar el líquido, que el adulto sabe que no hay líquido en la taza, etcétera, así como comprende lo que ella misma está haciendo, entonces hay a la vez auto- y hetero-adscripción y, por tanto, posesión de los conceptos psicológicos mencionados⁶.

Los niños solo imitan o se involucran en aquello que comprenden. Alrededor de los doce meses aparecen los primeros juegos funcionales simples (usar un objeto con su uso convencional pero de forma descontextualizada, de manera tal que no tiene los efectos que normalmente tendría: llevarse una cuchara vacía o un vaso vacío a la boca, por ejemplo) y rápidamente empieza a extenderse el ámbito de los posibles receptores de la acción (llevan la cuchara hacia la boca del adulto o del algún muñeco; o el adulto lleva la cuchara vacía a su boca o a la del niño). Entre los quince y los dieciocho meses, aproximadamente, empiezan a incorporarse sonidos simples a los esquemas de acción descontextualizados, como chistidos, y comienza el ordenamiento secuencial, es decir, ponerlos uno detrás del otro (revolver con la cuchara en un vaso vacío y llevarla vacía a la boca). El niño ha ido adquiriendo estas habilidades a través del aprendizaje del uso de instrumentos (cucharas, vasos) posible gracias a la interacción con los adultos que le muestran o enseñan explícitamente su uso, y gracias a que cuando los niños se obsesionan, casi, con el uso descontextualizado de objetos, este es inmediatamente continuado, seguido, iniciado —en cualquier caso, tomado como actividad con sentido— por los adultos que, por ejemplo, abren la boca cuando le acercan una cuchara vacía. Este es el camino típico del desarrollo del juego funcional. Aunque no lo desarrollemos acá, es importante señalar que las imitaciones avanzadas entre niño y adulto, propias del juego en el que uno y otro dan sentido a sus conductas y se imitan, son un logro evolutivo resultado de un largo proceso de andamiaje durante el periodo de despliegue de las formas básicas de intersubjetividad (en los que la imitación mutua y la imitación han cumplido un rol esencial). El bebé nace a un mundo de seres que otorgan sentido a sus conductas: sus movimientos, llantos y acciones iniciales son sometidas desde el principio a un filtro subjetivo de interpretación humana, de forma tal que algunas de sus conductas

⁶ Estamos presuponiendo que la posesión de conceptos psicológicos implica maestría para su uso tanto en primera como en tercera persona. Para una defensa de esta idea véase Diana Pérez, «Conceptos psicológicos» (en elaboración).

no intencionadas, carentes de sentido, son consideradas como relevantes y coherentes en términos humanos. Si los bebés llegan a comunicarse de forma intencionada es precisamente porque sus conductas han sido consideradas como intencionadas y humanamente significativas desde el inicio, cuando aún no eran producto de intenciones. Y es este mismo filtro subjetivo el que andamia, avala, permite, favorece el despliegue del significado compartido del juego funcional (Español, 2004). Con todo este bagaje cuenta A cuando recibe de regalo su juego de té. Así, A comprende de inmediato la propuesta de juego funcional, la imita y hace que el juego continúe.

Nos hemos detenido en este ejemplo porque nos interesa partir de un hecho establecido en algunas observaciones sobre la génesis del juego simbólico en los niños que, creemos, permite explicar con más detalle el pasaje que nos ocupa. Se trata de distinguir claramente entre diferentes tipos de juego. Vamos a enfocarnos en distinguir entre el juego funcional (como el juego de A) y el juego de ficción, que evolutivamente se presenta posteriormente (es decir, el juego funcional es siempre anterior al juego de ficción). La distinción entre juego funcional y de ficción, aunque no siempre se destaca, es reconocida por varios autores (como Hobson, 1995; Rivière, 1984; McCune & Agayoff, 2002; Español, 2004) y coincide con la distinción de Piaget (1977[1946]) entre la ritualización de esquemas y el esquema simbólico propiamente dicho. Además de reconocerse que en el desarrollo normal el juego funcional antecede al juego de ficción, en el desarrollo alterado la distinción permite demarcar niveles de afectación. De acuerdo con Rivière (2003[2000]), en el espectro autista existen diversos niveles de afectación de los trastornos cualitativos de las competencias de ficción e imaginación. En nivel 1, el más afectado, hay una ausencia completa de actividades que sugieren juego funcional o simbólico, así como de cualquier clase de expresiones de competencias de ficción. Pero en el nivel 2 hay presencia de juegos funcionales (que no implican sustitución de objetos o invención de propiedades, consistentes en aplicar a objetos funciones convencionales, de forma simple, por ejemplo, hacer rodar un cochecito o llevar una cucharita vacía a la boca). Y no hay juego simbólico. En los niveles 3 y 4, que suelen corresponder al autismo de alto funcionamiento y al síndrome Asperger, hay juego simbólico pero poco flexible. El juego funcional prácticamente se solapa con el establecimiento de las pautas comunicativas preverbales (el extremo final de lo que suele considerarse intersubjetividad secundaria): aparece al principio del segundo año de vida. El juego de ficción aparece meses después, hacia el final del segundo año de vida. Consideremos los siguientes dos ejemplos como paradigmáticos:

- (1) Ejemplo de juego funcional: agarrar una taza vacía y « tomar».
- (2) Ejemplo de juego de ficción (*pretend play*): agarrar una banana y usarla como si fuera un teléfono.

Nuestra propuesta consiste en mostrar las diferencias en lo relativo a las estructuras cognitivas involucradas en cada uno de estos tipos de juego. Propondremos que este estricto orden evolutivo responde a una diferencia en el grado de complejidad cognitiva involucrado en cada uno de ellos. En ambos casos hay un *como si*, una acción que no es dirigida hacia el mundo tal cual es, no por ignorancia acerca del mundo, sino sabiendo que el mundo es de una forma diferente a la manera en la que el nene se dirige a él. Por tanto en estos casos el niño tiene un estado psicológico que no puede describirse usando un concepto psicológico como saber/conocer, sino que necesariamente hay que usar un concepto diferente. Pero tampoco es posible apelar al concepto de *creencia* porque el nene *no* cree que haya leche en la taza ni que la banana sea un teléfono, ni —obviamente— atribuye estas creencias erróneas al adulto (el adulto no expresa sorpresa al «tomar» de la taza y encontrarla vacía, ni el niño). El nene juega a que / simula que / imagina que hay un líquido en la taza / habla por teléfono, y comprende que es eso mismo lo que hace el adulto. ¿Cuáles serían las capacidades cognitivas mínimas que deberíamos atribuirle a este niño que es capaz de involucrarse en este tipo de juegos?

Consideremos en primer lugar el juego funcional. En este caso parece que no basta con atribuir estados psicológicos con contenido objetivo. Hemos argumentado que en el periodo entre los nueve y doce meses, en el que hay interacciones propias de la intersubjetividad secundaria, el bebé es capaz de atribuir al adulto y a sí mismo estados psicológicos como desear algo, intentar algo o temer algo, pero que no había razones para pensar que sería posible en este momento evolutivo contar con atribuciones psicológicas (ni a sí mismo ni a los demás) con contenido proposicional. La idea sería que *es en el juego funcional en el que por primera vez es necesario atribuir al niño pensamientos con contenido proposicional*: el nene sabe *que no hay leche* pero simula *que la hay* (juega a que la hay). Se da, entonces, por primera vez el contraste entre estados psicológicos en los cuales existe una proposición afirmada y la misma proposición negada (bajo verbos psicológicos diferentes), sin que esto lleve al niño a cambiar ninguno de los dos estados (es decir, no es algo que lo lleve a corregir el conocimiento que tiene del mundo, por ejemplo, ni algo que sorprenda al niño), sino que, al contrario, parece que el placer se da en este caso no exclusivamente por el movimiento corporal —como ocurre con el juego sensorio-motor—, sino por el desajuste entre los diferentes estados cognitivos.

Consideremos ahora el juego de ficción. En primer lugar, es importante recordar que siempre aparece evolutivamente después del juego funcional, o sea que lo supone: podemos hipotetizar que el juego de ficción es más complejo cognitivamente que el juego funcional. Nuestra propuesta es señalar que la complejidad de este nuevo tipo de juego deriva del hecho de que el contenido auto- y hetero-atribuido en este caso

es más complejo: en efecto, no se trata de una misma proposición —«hay líquido»— afirmada como contenido de un estado simulado —«pretendo que...»— y negado en otro —«sé que...»—, sino que *en el caso del juego de ficción hay una negación que no afecta a la proposición como un todo, sino a la predicación*. Veamos el ejemplo: el nene juega a que la banana es un teléfono. ¿Es necesario atribuirle al niño una creencia con el contenido «la banana es un teléfono»? Creemos que no: lo que hay que atribuirle al niño es un estado psicológico con el siguiente contenido proposicional: (el nene sabe que) esto no es un teléfono/es una banana y (el nene juega a/simula/imagina que) esto es un teléfono. En este caso hay por primera vez incompatibilidad a nivel de predicados (no de proposiciones completas). Existe un único objeto físico (el que el nene tiene en la mano)⁷ que soporta predicados incompatibles en las distintas actitudes psicológicas adoptadas simultáneamente por el niño. Un vez más, es este desajuste lo que hace atractivo/motiva la acción.

Recordemos que, según habíamos argumentado en el tercer apartado, en el caso de los niños de un año (en el periodo de la intersubjetividad secundaria) solo se podría considerar que aquellos poseen y son capaces de reconocer en los demás (y en sí mismos) estados psicológicos como sensaciones corporales (dolor), algunas emociones con contenido objetual pero no proposicional (miedo al león) y el concepto de deseo hacia un objeto (querer un yogur) y de atención / intención dirigida a un objeto (mirar el trompo, agarrar el vaso, etcétera), así como saberes procedurales. El pasaje a la intersubjetividad terciaria suponía lograr manejar estados psicológicos con contenido proposicional, en los cuales: (1) hay V o F del contenido, (2) hay predicación y (3) están constituidas por conceptos, es decir, elementos constituyentes recombinables. Si nuestra descripción de lo que ocurre en el juego funcional es correcta, podríamos decir que es allí donde por primera vez aparece (1), en otras palabras, por primera vez allí el niño puede confrontar la verdad o falsedad de los contenidos de sus pensamientos. Seguidamente, con el juego de ficción aparece (2), es decir, la predicación acerca de un objeto físico presente que bajo diferentes actitudes psicológicas resulta ser y no ser un teléfono (o una banana).

El camino por recorrer para dar cuenta en forma exhaustiva del pasaje de las formas básicas de intersubjetividad a la atribución psicológica madura es largo, y falta mucho por recorrer. En particular, falta explicar el surgimiento del tercer rasgo involucrado en la atribución psicológica de contenidos proposicionales, es decir, la posibilidad de recombinar estos contenidos, para lo cual es necesario que esta atribución psicológica no apunte a través de un índice «esto» al sujeto del contenido proposicional,

⁷ Es importante destacar que en una primera etapa evolutiva este juego solo es posible con un objeto físico como disparador. Es necesario avanzar en la complejidad cognitiva un nivel más (o varios más) para incorporar sujetos ficticios.

sino que genuinamente haya una mediación conceptual para el acceso tanto al predicado como al sujeto constituyentes de la proposición atribuida. Probablemente sean dos los procesos centrales implicados en este tránsito: el juego protagonizado y la práctica de hablar y narrar en la vida cotidiana. Dejamos para otra ocasión el desarrollo de estas ideas.

REFERENCIAS BIBLIOGRÁFICAS

- Baron-Cohen, Simon, Alan Leslie & Uta Frith (1985). Does the Autistic Child Have a «Theory of Mind»? *Cognition*, 21, 37-46.
- Bermúdez, José Luis (2003). *Thinking Without Words*. Oxford: Oxford University Press.
- Bråten, Stein & Colwyn Trevarthen (2007). From Infant Intersubjectivity and Participant Movements to Simulation and Conversation in Cultural Common Sense. En Stein Bråten (ed.), *On Being Moved. From Mirror Neurons to Empathy* (pp. 21-34). Ámsterdam-Filadelfia: John Benjamin Publishing Company.
- Currie, Gregory (1998). Pretence, Pretending and Metarepresenting. *Mind and Language*, 13(1), 35-55.
- Davidson, Donald (2001). *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- Español, Silvia (2004). *Cómo hacer cosas sin palabras. Gesto y ficción en la infancia temprana*. Madrid: Machado.
- Gallagher, Shaun (2001). The Practice of Mind: Theory, Simulation or Practice? En Evan Thompson (ed.), *Between Ourselves. Second Person Issues in the Study of Consciousness* (pp. 83-108). Exeter: Imprint Academic.
- Gómez, Juan Carlos (1996). Non-human Primate Theories of (Non-human Primate) Minds: Some Issues Concerning the Origins of Mind-reading. En Peter Carruthers y Peter K. Smith (eds.), *Theories of Theories of Mind* (pp. 330-343). Cambridge: Cambridge University Press.
- Gómez, Juan Carlos (1998). Do Concepts of Intersubjectivity Apply to Non-human Primates? En Stein Bråten (ed.), *Intersubjective Communication and Emotion in Early Ontogeny* (pp. 245-259). Cambridge: Cambridge University Press.
- Gomila, Antoni (2002). La perspectiva de segunda persona de la atribución mental. *Azafea*, 4, 123-138.
- Hobson, R. Peter (1995). *El autismo y el desarrollo de la mente*. Madrid: Alianza.
- Hutto, Daniel (2008). *Folk Psychological Narratives. The Sociocultural Basis of Understanding Reasons*. Cambridge: The MIT Press.

- Jarrold, Chris, Peter Carruthers, Peter Smith & Jill Boucher (1994). Pretend Play: Is it Metarepresentational? *Mind and Language*, 9(4), 445-468.
- Leslie, Alan (1987). Pretence and Representation: The Origins of «Theory of Mind». *Psychological Review*, 94, 412-426.
- Leslie, Alan (1994). Pretending and Believing: Issues in the Theory of ToMM. *Cognition*, 50, 211-238.
- Malloch, Stephen & Colwyn Trevarthen (eds.) (2009). *Communicative Musicality: Exploring the Basis of Human Companionship*. Oxford: Oxford University Press.
- Onishi, Kristine H. & Renée Baillargeon (2005). Do 15-month-old Infants Understand False Beliefs? *Science*, 308, 255-258.
- McCune, Lorraine & Joanne Agayoff (2002). Pretending as Representation: a Developmental and Comparative View. En Robert W. Mitchell (ed.), *Pretending and Imagination in Animals and Children* (pp. 43-55). Cambridge: Cambridge University Press.
- Piaget, Jean (1977[1946]). *La formación del símbolo en el niño. Imitación, juego y sueño. Imagen y representación*. México DF: Fondo de Cultura Económica.
- Rivière, Ángel (1984). Acción e interacción en el origen del símbolo. En Jesús Palacios, Álvaro Marchesi y Mario Carretero (comps.), *Psicología evolutiva*. Volumen II (pp. 145-174). Madrid: Alianza.
- Rivière, Ángel (2002). *Idea: inventario de espectro autista*. Madrid: Fundec.
- Rivière, Ángel (2003[1999]). Educación y modelos del desarrollo. En Mercedes Belinchón y otros. (comps.), *Ángel Rivière. Obras escogidas*. Volumen III (pp. 243-284). Madrid: Médica Panamericana.
- Rivière, Ángel (2003[2000]). Teoría de la mente y metarrepresentación. En Mercedes Belinchón y otros. (comps.), *Ángel Rivière. Obras escogidas*. Volumen I (pp. 191-231). Madrid: Médica Panamericana.
- Southgate, Victoria, Atsushi Senju & Gergely Csibra (2007). Action Anticipation through Attribution of False Belief by Two-year-olds. *Psychological Science*, 18(7), 587-592.
- Stern, Daniel (2010). *Forms of Vitality. Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy and Development*. Nueva York: Oxford University Press.
- Trevarthen, Colwyn (1979). Communication and Cooperation in Early Infancy: A Description of Primary Intersubjectivity. En Margaret Bullowa (ed.), *Before Speech: The Beginning of Interpersonal Communication* (pp. 321-348). Cambridge: Cambridge University Press.
- Trevarthen, Colwyn (1998). The Concept and Foundations of Infant Intersubjectivity. En Stein Bråten (ed.), *Intersubjective Communication and Emotion in Early Ontogeny* (pp. 15-46). Cambridge: Cambridge University Press.

- Trevarthen, Colwyn (2000). Musicality and the Intrinsic Motive Pulse: Evidence from Human Psychobiology and Infant Communications. *Musica Scientiæ*, Special issue, 155-215.
- Trevarthen, Colwyn & Penelope Hubley (1978). Secondary Intersubjectivity: Confidence, Confiding and Acts of Meaning in the First Year. En Andrew Lock (ed.), *Action, gesture and symbol: The emergence of language* (pp. 183-229). Londres: Academic Press.
- Trevarthen, Colwyn & Vasudevi Reddy (2007). Consciousness in Infants. En Max Velmans y Susan Schneider (eds.), *The Blackwell Companion to Consciousness* (pp. 41-57). Cambridge: Blackwell Publishing.
- Wimmer, Heinz & Josef Perner (1983). Beliefs about Beliefs: Representation and the Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. *Cognition*, 13, 103-128.

SECCIÓN III
ENSAYOS SOBRE LA EVOLUCIÓN DEL LENGUAJE

PERSPECTIVAS SOBRE EL ORIGEN Y LA EVOLUCIÓN DEL COMPONENTE SINTÁCTICO

Paola Cépeda

Pontificia Universidad Católica del Perú

Gabriel Martínez Vera

Pontificia Universidad Católica del Perú

1. INTRODUCCIÓN

Visto de manera amplia, el lenguaje puede ser descrito como compuesto de tres grandes componentes: uno que se encarga de la producción de los sonidos del habla (el sistema sensorio-motor o articulatorio-perceptual), otro que involucra los procesos de simbolización (el sistema conceptual-intencional) y otro que combina piezas lingüísticas pequeñas para formar otras más grandes (el sistema sintáctico). En los estudios sobre la evolución del primer componente, se han desarrollado comparaciones entre la fisiología del aparato fonador de la especie *Homo sapiens sapiens* (o humana) con la de los fósiles de especies ancestrales y de primates y otros animales contemporáneos (como en Lieberman, 1975), crucialmente, vinculadas con la laringe descendida de los adultos humanos (véase la argumentación de Lieberman y otros, 1969, y la contrapropuesta de Fitch, 2010). Del mismo modo, se han desarrollado propuestas en relación con el control vocal y respiratorio (Deacon, 1997 y Fitch, 2010), la especialización de la percepción del habla, entre otros asuntos.

Con respecto al segundo componente, los estudios evolucionistas han sido también comparativos y multidisciplinarios. La discusión se ha concentrado en diversos temas, entre ellos, el de si atribuir o no capacidades cognitivas complejas a los animales, como categorización, memoria, inteligencia social, uso de herramientas, comunicación, tradiciones, entre otras (véase Fitch, 2010 para un excelente recuento de los estudios en cognición animal). También se han formulado propuestas en relación con la teoría de la mente y la cohesión social (Dunbar, 1997). Otro elemento central de la discusión académica es la aparición de la capacidad simbólica

en la especie *Homo sapiens sapiens*. Deacon (1997) cree que un cambio en la dieta y la necesidad de representar relaciones sociales sancionables podrían haber sido el elemento de selección para la capacidad simbólica. Por su lado, Bickerton (1990) considera que el lenguaje es un sistema de representación secundario y que fue la adaptación a un nicho ecológico específico (el reclutamiento para aprovechar los cadáveres de la megafauna antes de que llegaran los depredadores carroñeros) la que propició la aparición capacidad simbólica (Bickerton, 2009).

Finalmente, en cuanto a la evolución del tercer componente, la discusión académica se ha incrementado notablemente en los últimos veinte años. Las aproximaciones se han enfocado en dos puntos centrales. El primero es la propuesta de un protolenguaje (Bickerton, 1990) y su caracterización. Por ejemplo, existen debates que enfrentan un protolenguaje composicional y léxico (Bickerton, 1990; Jackendoff, 1999 y 2002) con uno holofrástico (Arbib, 2010 y Wray, 1998). El segundo punto es el de la capacidad combinatoria propiamente dicha y qué elementos son los que la caracterizan. Diversas perspectivas han abordado este asunto: una funcionalista (Pinker & Bloom, 1990; Jackendoff, 1999 y 2002), una formalista (Hauser, Chomsky & Fitch, 2002; Chomsky, 2005), una basada en modelamientos virtuales (Christiansen & Kirby, 2003; Kirby, 2007), entre otras.

En este artículo, nos ocupamos del componente sintáctico. La razón de ello es que consideramos que los mecanismos combinatorios del lenguaje son los que lo hacen especial. A diferencia de los sistemas de comunicación animal, el lenguaje es el único sistema cognitivo recursivo utilizado para la comunicación. Como plantea Bickerton en su caracterización de la *paradoja de la continuidad* (Bickerton, 1990), existe una diferencia de carácter cualitativo entre el lenguaje y los sistemas de comunicación animal, y esta diferencia está basada, crucialmente, en el complejo carácter combinatorio del lenguaje. Para ofrecer un ejemplo, es posible encontrar que, en las formas de comunicación de los primates, hay casos de concatenación de dos elementos que provee un nuevo significado. No obstante, estas reglas son más una excepción que elementos generalizados: «ellas funcionan con tipos de llamadas específicas y estructuradas innatamente en un único contexto, y no representan el conjunto de reglas que se aplican a todas las llamadas (o a clases de llamadas), como sí lo hacen las reglas sintácticas del lenguaje. [...] Por ello, no son verdadera evidencia para un «precursor» de la sintaxis en el UAC (último ancestro común)» (Fitch, 2010, p. 185; la traducción es nuestra)¹.

¹ «[...] they apply to specific, innately structured call types in a single context, rather than representing rules that are applied to all calls (or call classes), as syntactic rules in language are. [...] Thus, these provide little evidence of a 'precursor' of syntax in the LCA (Last Common Ancestor)».

Dentro de un marco teórico que asume la existencia de una capacidad especializada para las habilidades lingüísticas, de carácter exclusivamente humano, nuestro objetivo es, entonces, vincular las propuestas que se han formulado sobre el origen y la evolución del componente sintáctico con las evidencias paleoantropológicas sobre la evolución humana que se han presentado en los últimos años. En este sentido, partimos de la necesidad de plantear hipótesis tentativas que puedan ir evaluándose a través de evidencia teórica y empírica. Esta pretensión nos lleva a establecer límites precisos para nuestra discusión. En primer lugar, establecemos un periodo temporal circunscrito entre hace 200 mil y 50 mil años, tal como puede indicarnos la evidencia paleoantropológica. En segundo lugar, tomamos como puntos de partida únicamente dos de las propuestas más famosas: la de Bickerton (1990 y 2009) y la de Hauser, Chomsky y Fitch (2002). A estas propuestas les añadiremos aquellos elementos que les hacen falta para mejorar su capacidad explicativa.

Esta presentación está organizada de la siguiente manera: en primer lugar, presentamos algunos hechos del componente sintáctico de las lenguas del mundo, a manera de una caracterización del estado actual de la evolución del lenguaje; en segundo lugar, ofrecemos una aproximación cronológica a los procesos involucrados, tratando de vincular las propuestas lingüísticas con algunos datos paleoantropológicos; finalmente, proponemos algunas conclusiones.

2. HECHOS SINTÁCTICOS ACTUALES DE LAS LENGUAS DEL MUNDO

En su estado evolutivo actual, el componente sintáctico del lenguaje tiene características especiales que necesariamente deben tomarse en cuenta dentro de una aproximación a su origen y evolución. Nos gustaría concentrarnos en los cinco hechos que presentamos a continuación².

2.1. Autonomía de la sintaxis

El componente sintáctico puede producir enunciados bien contruidos aun cuando su contenido no sea ni verosímil con nuestro conocimiento del mundo ni conocido por nosotros. En otras palabras, los mecanismos computacionales sintácticos operan de manera independiente, lo que se conoce como *autonomía de la sintaxis*. Observemos los siguientes ejemplos (el primero es de Chomsky, 1965):

- (1) a. Las ideas verdes incoloras duermen furiosamente.
- b. Las trastelas nautisoides protiraban los niqueses topilamente.

² Si bien la mayoría de los ejemplos son del castellano, consideramos que son extrapolables a la capacidad lingüística general.

Los ejemplos de (1) constituyen enunciados bien contruidos, independientemente de su significado. Aun cuando, en (1a), podemos reconocer, por ejemplo, la contradicción entre «verdes» e «incoloras», la construcción de toda la frase sigue perfectamente las reglas de combinación de unidades léxicas. Por su parte, a pesar de que (1b) es un invento, un juego lingüístico nada más, podemos reconocer las piezas léxicas, las relaciones entre los elementos oracionales, las funciones que cumplen esos elementos, incluso sin conocer el significado de cada uno de ellos.

El postulado de la autonomía de la sintaxis no debe entenderse mal: no se trata de postular la separación radical entre el sistema sintáctico y el sistema conceptual-intencional. Más bien, consiste en proponer que los mecanismos combinatorios del lenguaje son formulados sin referencia a su significado, al discurso en el que se enmarcan o al uso del enunciado en un contexto específico. No hay duda de que existen conexiones relevantes entre la estructura sintáctica y el contenido que se expresa, tal como podemos apreciar en los siguientes hechos sintácticos.

2.2. Estructura argumental y jerarquía de constituyentes

Las expresiones lingüísticas no son simples secuencias lineales de elementos, sino que cuentan con una estructura jerárquicamente organizada. La organización de estas jerarquías obedece, primariamente, a la estructura argumental de las piezas que se combinan. Las relaciones entre predicados y argumentos se establecen entre elementos sintácticos que se encuentran dentro del mismo constituyente, es decir, de manera local. Así, una estructura jerárquica no es sino la proyección de un predicado que ha acogido, en posiciones estructurales específicas, los argumentos que requiere para satisfacer sus requisitos léxicos de combinación. Veamos el ejemplo (2):

- (2) a. LEER < [Roberta] , [revistas de Lingüística] >
 b. [[Roberta] [lee [revistas de Lingüística]]].

Una evidencia de que las relaciones en la sintaxis son jerárquicas es el hecho de que a una misma secuencia lineal se le puede atribuir dos estructuras jerárquicas diferentes. Así, la oración «El gerente despidió al empleado sin escrúpulos» puede interpretarse de dos maneras: en (3a), el gerente realiza el despido sin escrúpulos; en (3b), es el empleado quien no tiene escrúpulos:

- (3) a. [El gerente] [despidió [al empleado [sin escrúpulos]]]
 b. [El gerente] [despidió [al empleado] [sin escrúpulos]]

Incluso secuencias que solo difieren en una palabra pueden estar ocultando estructuras completamente diferentes. Así, los enunciados de (4a) y (4b) poseen

estructuras subyacentes diferentes, que nos llevan a interpretar a «José» como el que ama en (4a), pero como el amado en (4b):

- (4) a. José es capaz de amar.
 AMAR < [José] , [(a otros)] >
- b. José es fácil de amar.
 AMAR < [(otros)] , [José] >

2.3. Categorías funcionales

Las categorías funcionales son elementos sintácticos que posibilitan las estructuras jerárquicas a través de la caracterización de rasgos relevantes para la proyección sintáctica. Esto permite explicar dos hechos sintácticos: en primer lugar, se están cumpliendo los requisitos argumentales de las piezas léxicas aun cuando no tengamos una manifestación auditiva de alguno de los argumentos; en segundo lugar, se garantiza una configuración asimétrica de las estructuras sintácticas.

En cuanto al primer punto, la postulación de categorías funcionales obedece a la idea de que las relaciones sintácticas y las condiciones de la buena formación de estructuras se establecen con elementos precisos de carácter obligatorio. Algunas categorías funcionales carecen de rasgos fonéticos, por lo cual se las llama *vacías*. Su comportamiento puede ser predicativo o argumental. Por ejemplo, en (5a), el predicado «querer» satura sus requisitos argumentales con las piezas fonéticamente plenas <Gabriel> y <comprar un libro>. ¿Acaso «comprar» no satisface sus requisitos? En verdad, sí lo hace: con <un libro> y con una categoría vacía <PRO>, cuya interpretación depende (está controlada) del sujeto de la cláusula principal:

- (5) a. Gabriel quiere comprar un libro.
 b. QUERER < [Gabriel_i] , [PRO comprar un libro] >
 c. COMPRAR < [PRO_i] , [un libro] >

Las categorías vacías incluyen, también, los elementos léxicos que pierden sus rasgos fonéticos una vez que se han copiado en una posición diferente de donde se originaron (a esto se le llama «movimiento»; véase 2.5.). Un ejemplo de esto es la cliticización de «have» en inglés: el auxiliar «have» puede adoptar una forma clítica «'ve». Observemos los ejemplos de (6):

- (6) a. They have studied the lesson. → They've studied the lesson.
 b. Will they have studied the lesson? → *Will they've studied the lesson?
 c. Will they have studied the lesson? → Will_i they <will_i> have studied the lesson?

La cliticización de «have» sigue una condición de adyacencia: el clítico debe estar inmediatamente adyacente a su anfitrión acentual. (6b) resulta agramatical, pues, aunque no lo escuchemos, hay un elemento sintáctico que está bloqueando la cliticización (como se ve en [6c]); dado que este elemento no tiene rasgos fonéticos, no puede ser el anfitrión acentual del clítico.

En cuanto al segundo punto, existe mucho más estructura de la que realmente se observa en relación con la expresión sintáctica de las estructuras argumentales de las piezas léxicas. Esto asegura la asimetría de las configuraciones que el componente sintáctico construye, que son requeridas para la interpretación de la derivación sintáctica en los componentes sensorio-motor y conceptual intencional (Kayne, 1994 y Moro, 2000). Por ejemplo, la categoría funcional Tiempo (T) no solo contribuye a la interpretación del enunciado, información que es leída por el sistema conceptual-intencional, sino que también permite el cotejo de rasgos formales que aseguran la buena formación del enunciado. Así, en (7), T está definido con el rasgo semántico [pasado], que es interpretado semánticamente, y, a la vez, está especificado con el rasgo [nominativo], que no es interpretable por los sistemas externos, sino que asegura que «Leonardo» coteje su propio rasgo [nominativo] en la posición de sujeto; estos rasgos, unido a los rasgos no interpretables [3ª persona, singular], permiten que T se realice fonéticamente como «-ió» en esta oración:

(7) a. Leonardo vendió dos globos en la feria.

b. [_{FT} Leonardo]_[3,sing,nom]_i vend_j-ió_[pas,3,sing,nom] <Leonardo>_i <vend->_j dos globos en la feria]

2.4. Infinitud discreta y recursividad

El lenguaje es sorprendentemente productivo: con pocos elementos, podemos producir muchos enunciados. Esto es posible dado que cualquier expresión lingüística recién construida siempre puede combinarse con otro elemento más. Las reglas de combinación permiten que, en esta combinatoria, un tipo de frase pueda estar contenido en otra de su mismo tipo (*embedding*). A esto se le llama *recursividad*. Así, una oración cualquiera siempre puede quedar subordinada a otra oración y a otra y a otra:

(8) a. José adora la música.

b. Juan cree que José adora la música.

c. Mariana piensa que Juan cree que José adora la música.

d. Liliana dice que Mariana piensa que Juan cree que José adora la música.

Etcétera.

2.5. Movimiento

Hemos señalado en 2.2 que las relaciones entre predicados y argumentos se dan de manera local. Sin embargo, en las expresiones lingüísticas, es posible encontrar argumentos en una posición distinta del lugar en el que son interpretados. Sucede, por ejemplo, en las preguntas. En (9), el argumento interno de «comprar» aparece desplazado al comienzo de la oración:

- (9) a. ¿Qué_i compró_j Gabriel <compró>_j <qué>_i?
 b. COMPRAR < [Gabriel] , [qué] >

El movimiento es una operación sintáctica fundamental motivada por el chequeo (o eliminación) de ciertos rasgos no interpretables por los sistemas de actuación externos. Su acción no es irrestricta, sino que existen condiciones, básicamente de localidad, que limitan los movimientos. Los ejemplos de (10) muestran una restricción universal que impide la extracción de un elemento adjunto dentro de la frase nominal:

- (10) a. La muchacha de falda azul compró pantalones de tres diseñadores.
 b. *¿[De qué]_i compró_j la muchacha <de qué>_i <compró>_j pantalones de tres diseñadores?

3. EVIDENCIA PALEOANTROPOLÓGICA

Con el fin de abordar el tema del origen y la evolución de la sintaxis, conviene que nos situemos en una perspectiva que tenga como punto de partida la evidencia paleoantropológica. Así, podemos fijar un periodo comprendido entre los últimos 200 mil y 50 mil años de evolución del ser humano contemporáneo. Tres tipos de evidencia sustentan este fechado: genética, anatómica y arqueológica. Los tres tipos de evidencia servirán para fijar la fecha más lejana; sin embargo, solo emplearemos evidencia genética y arqueológica para el límite más cercano.

3.1. Límite temporal más lejano

Hace 200 mil años aproximadamente, aparecería el *Homo sapiens sapiens* en África como descendiente de un grupo de *Homo erectus* (Lock & Peters, 1999, p. 61); esta sería la fecha más lejana en que una capacidad lingüística común a toda la especie podría surgir. La evidencia genética se concentra en los datos relativos a la Eva mitocondrial. El ADN nuclear de un individuo está compuesto por la herencia del padre y de la madre. El ADN mitocondrial, en cambio, solo proviene de la madre.

Esto es así porque, cuando el espermatozoide fertiliza el óvulo, pierde su propia mitocondria; el cigoto, por tanto, obtiene su mitocondria solo del óvulo. A través del análisis del ADN mitocondrial, se postula que el último ancestro común del ser humano contemporáneo es una única mujer (de ahí el nombre «Eva») que radicaba en África en una población de 10 mil individuos hace alrededor de 200 mil años (Lewin, 1998, p. 414; Lock & Peters, 1999, p. 48; Tattersall, 2008, pp. 90-91). La variación en el ADN mitocondrial entre los humanos constituye solo la décima parte de la de los chimpancés, lo cual sugiere una de dos cosas: (1) que los humanos evolucionaron relativamente recientemente (hace 140-280 mil años), o (2) que los humanos pasaron por un cuello de botella selectivo en algún momento de su historia evolutiva. Estas explicaciones no son excluyentes y están todavía sometidas al debate científico (Lewin, 1998, p. 416).

Los restos fósiles corporales, por su parte, constituyen la evidencia anatómica. Tres fósiles encontrados en Herto, Etiopía, son, claramente, modernos, es decir, pertenecen a la misma especie que el ser humano contemporáneo. Parece que estos tienen 160 mil años de antigüedad. Se trata de cráneos con expansión del hueso parietal (*parietal boss*) (Coolidge & Wynn, 2009, p. 209). Además, también en Etiopía, se han hallado los fósiles Omo I y Omo II en la formación Kibish de Omo. Estos fósiles tienen hasta 195 mil años de antigüedad y presentan también los parietales desarrollados (De la Torre Sáinz, 2008, p. 225). Otros restos fósiles con fechados cercanos revelan los mismos rasgos anatómicos del ser humano contemporáneo, que incluyen decrecimiento de la robustez dental y del esqueleto, modificaciones en la anatomía de la locomoción e incremento en el volumen cerebral. Sin embargo, aunque con rasgos modernos desde un punto de vista anatómico, no parecen modernos en su comportamiento³. Como podemos observar, tanto la evidencia anatómica como la genética sugieren que el origen de los humanos modernos puede fijarse en África. De ahí, salieron y se expandieron a Asia y Europa.

Finalmente, la evidencia arqueológica incluye el examen de los restos materiales. Se calcula que, hace 300 mil años, se inicia la Edad de Piedra Intermedia en África subsahariana y, hace 200 mil años se inicia el Paleolítico Medio en Eurasia (Lewin, 1999, p. 311). En este periodo, no solo la tecnología de las herramientas se refina, sino que también se desarrollan estrategias más eficientes para la organización social y empieza a aparecer la representación indexical. En relación con el primer punto, en Etiopía, la producción de cuchillas tiene hasta 180 mil de antigüedad, mientras que en Kenya, 240 mil; asimismo, en Zaire, se han encontrado puntas de hueso (como cabezas de arpón) que datan de hace 90 a 160 mil años (Lewin, 1999, p. 439).

³ Volveremos sobre este punto en la presentación de la evidencia arqueológica.

En cuanto al segundo punto, hay evidencia de que la búsqueda de alimentos y la caza no eran oportunistas, sino organizadas, en equipos y con conocimiento del territorio. Finalmente, respecto del tercer punto, hay evidencia de uso de pigmentos de diversos colores en Twin Rivers, Zambia, de hace 200 mil años. Se puede descartar que estos pigmentos se usaran para esconderse o como parte de la higiene personal, pues no hay razón para escoger una variedad de colores si con uno solo se puede conseguir el mismo efecto; por tanto, parece que es el color por sí mismo la razón para el empleo de los pigmentos (Coolidge & Wynn, 2009, p. 171). Sin embargo, estos comportamientos no son considerados estrictamente modernos, lo que contrasta significativamente con los rasgos anatómicos modernos descritos para este periodo.

Los tres grupos de evidencia, en conjunto, nos permiten fijar la fecha más lejana para el origen de la sintaxis en relación a la fecha de origen del *Homo sapiens sapiens*, hace 200 mil años.

3.2. Límite temporal más cercano

Por otra parte, hace 50 mil años, se produjo la segunda salida de África de nuestros antepasados directos. Esa es la fecha límite, entonces, en que una facultad lingüística común a toda la especie humana tendría que estar ya presente. El análisis de la evidencia genética para este fechado se centra en los datos relativos al ancestro común más reciente según el cromosoma Y (ACMR-Y), también conocido como «Adán cromosómico». Los hombres poseen un cromosoma X y uno Y, mientras que las mujeres poseen dos cromosomas X. El cromosoma Y, por tanto, se hereda del padre, lo que significa que los hijos varones de un mismo padre tienen el mismo cromosoma Y, y lo pasarán a sus propios hijos varones. Los estudios acerca del Adán cromosómico ubican a ancestro común como un hombre que radicaba en África hace alrededor de 51 mil y 39 mil años (Tattersall, 2008, pp. 92-93) y que poseía el cromosoma Y del que descienden todos los cromosomas Y contemporáneos (de ahí el nombre de «Adán»). El fechado es compatible con las migraciones desde el continente africano hacia Asia y Europa (Fagan, 2002, p. 93).

Junto con el dato genético, la evidencia arqueológica sugiere que, si hace 200 mil años ya existía alguna forma de abstracción, hace 50 mil años aproximadamente, esta estaba desarrollada. El comportamiento moderno está presente en la producción de herramientas, para las que se emplearon diversos materiales (huesos, marfil, cuernos) con fines específicos. Las representaciones se afinan con figurillas de humanos y animales, ornamentos personales hechos de marfil y conchas, entierros rituales, sitios habitacionales abiertos y organizados, contactos de larga distancia y comercio, instrumentos musicales (específicamente, flautas de hueso) (Coolidge & Wynn, 2009, p. 214).

La evidencia en Europa sugiere que hubo un gran cambio en la época, virtualmente instantáneo en tiempo arqueológico, radical y no gradual (Lewin, 1998, p. 435). Una explicación es que el cambio se debió o bien a una mutación genética aditiva, o bien a un evento epigenético que afectó la organización del cerebro (Mithen, 1996; Coolidge & Wynn, 2009). Una vez más, este cambio abrupto sería una sugerencia de que la evolución de la anatomía moderna ocurrió por separado de la evolución del comportamiento moderno.

¿Por qué no tomamos evidencia anatómica para este fechado? La razón es que, como hemos señalado, los rasgos anatómicamente modernos comienzan a aparecer alrededor de hace 200 mil años y no parece haber habido cambios evolutivos significativos en este aspecto que marquen una diferencia con los rasgos anatómicos de hace 50 mil años. Por esta razón, no solo no recurrimos a la evidencia anatómica, sino que esta no existe para fijar el límite que nos interesa.

En conclusión, la evidencia genética y la arqueológica nos permiten fijar los 50 mil años como la fecha más cercana para el origen de la sintaxis.

4. PROPUESTA CRONOLÓGICA EN LA EVOLUCIÓN DEL COMPONENTE SINTÁCTICO

Para establecer los momentos en la evolución del componente sintáctico, partimos de dos de las propuestas más citadas en la bibliografía contemporánea sobre el tema: la de Bickerton (1990 y 2009) y la de Hauser, Chomsky y Fitch (2002). Fijamos estas bases teóricas debido a su capacidad explicativa, a su adecuación a la información que la evidencia paleoantropológica nos brinda y a la necesidad de hacer explícita la coherencia argumentativa de ambas propuestas. Esta selección no ha sido sencilla; de hecho, hemos tenido que poner entre paréntesis algunas de las ideas más difundidas acerca del tema, ya que no siguen los criterios que hemos establecido. Entre ellas, están las siguientes:

- a) Jackendoff (1999 y 2002), porque su propuesta en hasta nueve etapas sucesivas no resulta verosímil con los datos cronológicos de la evidencia paleoantropológica
- b) Dunbar (1997), porque consideramos, junto con Bickerton (2007), que su propuesta establece las relaciones causales incorrectas entre lenguaje y el cotorreo como elementos de cohesión social
- c) Christiansen y Kirby (2003) y Tomasello (2008), porque sus aproximaciones no consideran una especialización cognitiva para el lenguaje (como una Gramática Universal).

Adicionalmente, debemos reconocer que las fuentes tomadas como base carecen de explicaciones para ciertos aspectos consignados entre los hechos sintácticos de las lenguas. Intentaremos completar esos vacíos en nuestra discusión.

Dividimos esta sección en dos partes: la primera corresponde a una revisión de la evolución del componente sintáctico con una antigüedad de 200 mil a 100 mil años, a la que llamaremos «etapa 1»; la segunda, con una antigüedad de 100 mil a 50 mil años, llamada «etapa 2».

4.1. Etapa 1

Como ya hemos observado, combinar unidades lingüísticas no es una simple operación de poner una después de otra. Requiere de la aplicación sistemática de los hechos sintácticos detallados en el apartado 2. Sin embargo, si nos ubicamos hace 200 mil años aproximadamente, nuestro límite más lejano en el origen del componente sintáctico, encontraremos algunas primeras muestras de combinación de unidades que, más adelante en la historia evolutiva humana, manifestarán los hechos sintácticos presentados.

Para Bickerton (1990), el componente sintáctico aparecería de manera abrupta en la historia evolutiva de la especie. Desde su perspectiva, las estructuras cognitivas del cerebro permitían a nuestros antepasados hacer un cálculo social, es decir, distinguir no solo a los individuos del entorno, y los tipos de acción y eventos, sino también crear una representación de los roles de cada individuo en un evento determinado. En algunas oportunidades, un individuo era el agente de una acción; en otras, era el paciente; en otras, en cambio, era el beneficiario. En este sentido, estamos frente a una inteligencia social que le permite al sujeto contar con estructuras relacionales de carácter semántico (esto es, interpretativas) para intervenir y relacionarse de manera exitosa en la vida social del grupo al que pertenece.

Bickerton propone que, como no hay nada completamente nuevo en la evolución y que la semántica existía antes que la sintaxis, debemos observar, en esas relaciones semánticas, la base para las estructuras oracionales. En efecto, desde su perspectiva, la sintaxis aparece hace alrededor de 150 mil años cuando se comienzan a emplear las relaciones semánticas como elementos obligatorios en los enunciados. Esto es lo que Bickerton denomina *cartografía del cálculo social en el protolenguaje*. Los elementos primordiales del componente sintáctico son, así, exaptaciones de la representación de las relaciones sociales.

El traslado de las relaciones semánticas al terreno sintáctico de los enunciados lingüísticos permite la generación de una combinatoria incipiente: las unidades semánticas (léxicas) tienen, ahora, requisitos de combinación que aseguran la buena formación de estructuras lingüísticas. Esto significa que las piezas léxicas que denotan

eventos del mundo tienen ahora la necesidad de satisfacer requisitos combinatorios para poder ser interpretados. Así, por ejemplo, la pieza léxica «buscar» debe satisfacer, por lo menos, dos requisitos combinatorios: un agente (alguien que busca) y un paciente (algo o alguien que es buscado). El surgimiento de un correlato sintáctico de las estructuras relacionales semánticas produce, así, un *saber lingüístico* en sentido estricto.

Con esta propuesta, Bickerton ofrece, desde una perspectiva evolucionista, una explicación bastante plausible de la aparición de los dos primeros hechos sintácticos que hemos descrito: la representación sintáctica de las estructuras argumentales y la autonomía de la sintaxis. En efecto, la presencia de una estructura semántica relacional lleva consigo el hecho de que la combinatoria no dependa ni de las representaciones eventivas ni de lo que sucede en el mundo, sino que lo importante es la satisfacción de los requisitos de una pieza léxica en sí misma, es decir, requisitos combinatorios de orden estrictamente lingüístico. Así, el componente sintáctico se va configurando como autónomo. Más aun, estos requisitos de orden lingüístico se almacenan junto con las piezas léxicas como rasgos de su estructura argumental, la cual debe ser saturada con otras piezas léxicas cuya presencia en la estructura sintáctica dependerá estrictamente de aquellas⁴. De este modo, se van disponiendo las piezas léxicas en grupos de predicados y argumentos.

Este mismo escenario puede permitir una explicación de la aparición de las categorías funcionales o, por lo menos, de una de sus características⁵. En efecto, los requisitos lingüísticos que exigen satisfacer los argumentos de las piezas léxicas habrían permitido que su cumplimiento se considere efectuado gracias a elementos vacíos, puramente funcionales. Sin embargo, dado que su aparición obedece a criterios de buena formación de enunciados lingüísticos, no podrían tratarse de elementos condicionados por el contexto o situación comunicativa, sino por condiciones lingüísticas que también se van almacenando a manera de rasgos en las piezas léxicas. Incluso se pueden asociar algunos rasgos específicos con elementos fonéticamente nulos en ciertas posiciones estructurales. Por ejemplo, dado que, en el enunciado «Daniela ya comió», la pieza léxica «comer» tiene como requisitos un agente (el que come, «Daniela») y un tema (lo que se come, una pieza léxica sin rasgos fonéticos en este caso), se podría pensar que el segundo requisito es «natural» y queda sobreentendido puesto que «siempre se come *algo*». Sin embargo, en el enunciado «Anoche nevó»,

⁴ Ello no implica que la jerarquía hay sido compleja desde su aparición (que haya contado con constituyentes o que haya sido binaria, por ejemplo). Lo relevante en este punto es que hay unidades de las cuales otras dependen.

⁵ Este punto no es tratado por Bickerton, sino que es nuestra extensión de su propuesta para dar cuenta de los hechos sintácticos mencionados.

reconocemos un sujeto fonéticamente nulo que no está condicionado por el contexto: el uso regular del predicado «nevar» exige, en castellano, un sujeto nulo, no porque no haya nada en el mundo que nieve, sino porque es la condición lingüística de esa pieza léxica⁶. En otras lenguas, la condición lingüística de los equivalentes de este predicado es combinarse con un pronombre explícito pero expletivo, es decir, no referencial; esto muestra el paso de elementos meramente referenciales, vinculados con el mundo, a elementos que se desprenden de su carácter referencial y aparecen como piezas necesarias para la construcción de oraciones.

En consecuencia, sobre la base del escenario planteado por Bickerton, se puede dar cuenta de dos de los hechos sintácticos que explican la combinatoria lingüística actual y se puede empezar a emprender una explicación del lugar de las categorías funcionales en la evolución del componente sintáctico.

¿Por qué incluimos el desarrollo de estos hechos de las lenguas contemporáneas en esta etapa? Recordemos que los restos fósiles y arqueológicos permiten relacionar el crecimiento de la capacidad cerebral (la expansión de los parietales) con el incremento de las capacidades cognitivas, aun cuando no con características modernas. Entre estas capacidades cognitivas, destaca la cada vez más sofisticada organización de los grupos sociales y el establecimiento de alianzas internas, lo que indica cognición social (Dunbar, 1997). Nuestros antepasados podían predecir las consecuencias de su propio comportamiento en otros individuos, entender los motivos de los otros y obtener beneficios a través de relaciones de contacto con otros. Esta complejidad genera, sin duda, más complejidad: los competidores ahora están dentro de la misma comunidad. El intercambio de información confiable, efectiva y oportuna es la clave para el altruismo recíproco y su impacto en los mecanismos de comunicación social (Calvin & Bickerton, 2000). Las relaciones de predicados y argumentos en este *protolenguaje*, en el sentido de Bickerton (1990), funcionan como mecanismos que posibilitarían la complejización de la teoría de la mente así como la transmisión cultural (enseñanza). Pudo no haber sido necesaria la instrucción lingüística, pero sí la atención compartida a un índice específico, un modo de representar la realidad (Coolidge & Wynn, 2009, pp. 173-174). Se trata de un comportamiento nuevo sobresaliente en esta etapa. El cerebro cambió y dio origen a nuevas habilidades. De hecho, parece haber signos de un área de Broca en el *Homo rudolfensis*, pero no en los australopitecinos. Algún evento evolutivo debe haber habido, pues, al comienzo del linaje Homo, que potenciaría sus capacidades (Lewin, 1998, pp. 457-460).

⁶ De hecho, resulta bastante contradictorio afirmar que, para «comer», se puede usar un segundo argumento nulo porque «siempre que se come, se debe comer *algo*», mientras que, para «nevar», se usa un argumento nulo porque «no hay *nada* que nieve».

Sin embargo, debemos ser cautos. Las lenguas producen jerarquías de grupos de palabras, mientras que las vías de producción de herramientas producen secuencias de acciones. Hay quienes sugieren que las secuencias de acciones complejas pueden ser evidencia de la capacidad lingüística ya desarrollada en la especie; sin embargo, una vez que conocemos cómo funciona la sintaxis y cuáles son los hechos sintácticos en las lenguas del mundo, esta idea puede quedar descartada. Del mismo modo, mientras un niño expuesto a una lengua determinada puede desarrollar el conocimiento de esa misma lengua sin mayor esfuerzo, no podrá hacer lo mismo con la producción y el uso de herramientas. Así, pues, aun cuando la producción secuencial de una herramienta refinada puede ofrecernos indicios importantes sobre la evolución de la cognición, no puede revelar mucho sobre los mecanismos combinatorios jerárquicos de las lenguas (Coolidge & Wynn, 2009, p. 175). Por esa razón, para esta etapa, queremos resaltar la aparición autónoma de una (proto-)sintaxis, sin duda eficaz para satisfacer las necesidades que surgen frente a la transmisión de tecnologías complejas no solo a través de la observación, sino también a través de la instrucción verbal.

4.2. Etapa 2

Los cambios más notables en el componente sintáctico terminarían configurándose de manera abrupta y reciente, tomando como límite máximo hace alrededor de 100 mil años. Esta etapa permitiría explicar la asimetría generada por las categorías funcionales, la recursividad y el movimiento.

Hauser, Chomsky y Fitch (2002) consideraron pertinente proponer una separación de dos distintas nociones de lenguaje: una facultad del lenguaje en sentido amplio (*Faculty of Language Broad sense* o FLB) y una facultad del lenguaje en sentido estrecho (*Faculty of Language Narrow sense* o FLN). De acuerdo con los autores, lo exclusivamente humano del lenguaje es un mecanismo computacional de carácter combinatorio y recursivo, denominado *unbounded Merge* (Chomsky, 2007), que compondría la FLN.

El mecanismo recursivo consiste en la aplicación de la combinatoria de manera jerárquica y de modo tal que una combinatoria puede incluir, dentro de sí, a otra de su mismo tipo. El rasgo recursivo de este componente computacional se habría exaptado hace 100 mil años de otros sistemas cognitivos, en los cuales habría surgido como una adaptación producto de la selección natural. Tomando en cuenta que la exaptación implica una reorganización cerebral, la recursividad es también una consecuencia del cambio generado por la estructura relacional semántica vinculada con los eventos. Dadas las condiciones adecuadas, cualquier evento puede ser participante de otro evento (Calvin & Bickerton, 2000), de manera recursiva. La representación sintáctica de las relaciones semánticas hereda esta capacidad recursiva, lo que resulta

compatible con la propuesta de la recursividad como exaptación de otras habilidades cognitivas, como representación de las relaciones sociales o teoría de la mente (Dunbar, 1997) o la orientación espacial (Hauser, Chomsky y Fitch, 2002)⁷.

La aparición del mecanismo recursivo permite explicar los últimos hechos presentados para el componente sintáctico. En primer lugar, las categorías funcionales deben aparecer junto con el mecanismo recursivo. Dado que las categorías funcionales son elementos combinatorios, el componente computacional se apoya en ellos y busca satisfacer la buena formación de enunciados deshaciéndose de aquellos rasgos que no sean ni semánticos ni fonéticos, es decir, eliminando las marcas propias de la combinatoria sintáctica y dejando únicamente aquello que sea interpretable por los sistemas externos conceptual-intencional y sensorio-motor (Piattelli-Palmarini & Uriagereka, 2004).

Como analogía, podemos imaginar cómo juntamos dos piezas de Lego. Cada una tiene elementos de cohesión que le permiten ensamblarse con otra pieza: por un lado, son protuberancias visibles y, por el otro, son agujeros de encaje. Una vez que dos piezas se ensamblan, esos elementos de cohesión por los cuales las piezas están unidas quedan invisibles, escondidos en la propia unión de las piezas. Pensemos en las protuberancias como los rasgos formales, puramente constructivos, y pensemos en los agujeros de encaje como las categorías funcionales, que permiten tanto la combinatoria (y la eliminación de los rasgos formales) como la interpretación (gracias a la eliminación de los rasgos formales, permiten ir creando aquello que puede ser interpretado por un sistema externo, tal como el constructor de Lego interpreta la unión de las piezas como la figura de un árbol, un robot, una nave espacial).

La aparición de las categorías funcionales y la del componente computacional recursivo deben ser simultáneas, porque la recursividad necesita su presencia para la sucesiva eliminación de rasgos de construcción oracional, del mismo modo que los rasgos formales necesitan de un sistema que pueda leerlos y realizar operaciones combinatorias con ellos. En otras palabras, sin una computación que los utilice para la combinatoria lingüística, los rasgos formales habrían resultado inútiles, no habrían sido seleccionados y, probablemente, habrían desaparecido en la historia evolutiva de nuestro sistema cognitivo. Del mismo modo, sin elementos que permitan la combinación, un componente computacional que no ejecuta acciones habría resultado superfluo y, tal vez, habría podido desaparecer sin mayores consecuencias.

En segundo lugar, el movimiento es también contemporáneo de la recursividad, ya que este surge como una necesidad de eliminar los rasgos no legibles por los interfaces (como los rasgos de las categorías funcionales). Este hecho sintáctico

⁷ Aun cuando la numeración también fue considerada como candidata a ser la fuente de la recursividad sintáctica, consideramos que puede observarse, más bien, como parasitaria de esta (Chomsky, 2011).

supone la presencia de rasgos por eliminar (que, como ya hemos señalado, son los rasgos formales que ingresan en la combinatoria), así como de espacios sintácticos a los cuales desplazar las piezas para el cotejo de sus rasgos. Estos espacios sintácticos mantienen relaciones estructurales específicas, vinculadas con la jerarquía sintáctica, con las posiciones sintácticas en las que los rasgos formales son incorporados en la estructura. Así, un rasgo no interpretable por los sistemas externos puede ser eliminado únicamente si un elemento compatible con él lo «observa» desde una posición desde donde lo c-comanda y lo puede atraer hacia él⁸. Esta es la motivación para el movimiento.

El movimiento, entonces, repercute positivamente en la buena formación de enunciados. Los elementos sintácticos son interpretados en un lugar (su lugar de origen, donde establecen relaciones de predicados-argumentos), pero pueden aparecer en un lugar diferente («desplazados») como producto del movimiento.

Luego de su origen, el componente sintáctico tendría que haber sido seleccionado debido a las ventajas adaptativas que les otorgaba a los miembros de la especie. En este sentido, tuvo que haber seguido un proceso evolutivo igual al de cualquier otro órgano o sistema cognitivo regido por la selección natural. Al respecto, la propuesta de Chomsky (2007) consiste en que, por ejemplo, la exaptación del mecanismo recursivo ocurrió en un individuo. Debido a las ventajas que le daba en su entorno la manipulación de unidades simbólicas mediante dicho mecanismo, tanto para su capacidad de pensamiento como para la comunicación de significados proposicionales, su *pool* genético sobrevivió a lo largo de las generaciones y se propagó paulatinamente en la especie.

La propuesta chomskiana de que únicamente bastaría una simple mutación para la aparición de la FLN ha sido severamente criticada. Sin embargo, creemos que las críticas son infundadas. El genetista J. B. S. Haldane propuso el siguiente ejemplo: «si un nuevo gen les diera a los individuos que los poseen solo el 1% de ventaja en la reproducción (101 descendientes por cada 100 del resto de individuos), entonces incrementaría su frecuencia de 0,1% a 99,9% en solo cuatro mil generaciones. Para los humanos, 4 mil generaciones es solo 100 mil años, un periodo muy breve en la escala temporal evolutiva» (citado por Coolidge & Wynn, 2009, p. 51; la traducción es nuestra)⁹. Esto es, una pequeña mutación en un solo gen puede ser efectiva. Ahora, pensemos en mutaciones que afectan los genes reguladores. Por ejemplo,

⁸ Recordemos que, en la estructura sintáctica, un nudo A c-comanda un nudo B si y solo si el nudo B está dominado por un nudo hermano de A.

⁹ «if a new gene gave the individuals who had it only a 1 percent advantage in reproduction (101 offspring to 100 for everyone else), it would increase in frequency from .1 percent to 99.9 percent in only 4000 generations. For humans, 4000 generations is only 100000 years - a very brief moment in the evolutionary time scale».

si una mutación modifica el número de ciclos de la división celular de las neuronas, habría un cambio drástico en el número total de neuronas de manera muy vertiginosa. «Aun cuando la mutación en sí misma es pequeña, el cambio en la anatomía sería drástico, y, si esto produjera una ventaja reproductiva (y recordemos que los cerebros son costosos en su metabolismo, así que no hay garantía de que haya una ventaja reproductiva), incrementaría rápidamente su frecuencia» (2009, p. 52; la traducción es nuestra)¹⁰. No debemos descartar, pues, la idea de una simple mutación para el caso de la facultad lingüística de los humanos.

La recursión posee una clara ventaja: permite oración más largas y complejas, lo que, a su vez, posibilita representaciones, interacciones y planificaciones más complejas. Del mismo modo, posibilita la flexibilidad de sentidos, la creatividad lingüística y la capacidad referencial (incluso para referirse al lenguaje mismo). No obstante eso, la recursión exige mayor capacidad de almacenamiento y más memoria de trabajo. Bickerton (Calvin & Bickerton, 2000) y Piattelli-Palmarini y Uriagereka (2004) suponen que diversos sistemas sintácticos entraron en competencia entre ellos con el fin de eliminar la ambigüedad de las oraciones y hacerlas más aptas. La propuesta es que la sintaxis habría operado ella misma como un elemento de presión selectiva favoreciendo cualquier cambio en el sistema nervioso que facilitara la construcción e interpretación de enunciados. Un conjunto de sucesivas adaptaciones habría mejorado la aptitud de los individuos y su pool genético. De manera adicional, la selección orgánica (o efecto Baldwin) habría jugado un papel importante en la evolución de la sintaxis (Calvin & Bickerton, 2000). La selección orgánica es un proceso evolutivo por el cual se van convirtiendo en un instinto aquellos rasgos que, en un principio, constituían conductas aprendidas. La idea es que, generación tras generación, los niños realizaban de forma espontánea lo que inicialmente era un proceso consciente en sus mayores. De este modo, la sintaxis quedaría instalada en la especie, de acuerdo con Calvin y Bickerton (2000), un par de miles de años después de su aparición. Muy recientemente, con la ayuda de simulaciones por computadora, se ha mostrado que, cuando un entorno permanece estable, se da una presión selectiva para que las habilidades aprendidas se vayan haciendo progresivamente innatas. Si una habilidad es innata, podrá emerger antes en el transcurso de la vida de un organismo, lo que evita que los más jóvenes se expongan a una suerte desastrosa y mueran prematuramente sin haber tenido acceso a la experiencia para desarrollar dicha habilidad (Pinker, 1994). Sin embargo, estas consideraciones teóricas deben ser todavía exploradas con más profundidad, de manera principal, el tiempo estimado

¹⁰ «Even though the mutation itself is small, the change in anatomy would be dramatic, and if this yielded a reproductive advantage (and keep in mind that brains are metabolically expensive, so an advantage is by no means guaranteed), it would increase rapidly in frequency».

por Calvin y Bickerton (2000) para la internalización y generación de un comportamiento innato («un par de miles de años» no parece ser plausible).

Frente a la idea de que parece que el comportamiento moderno fue apareciendo gradualmente, nos colocamos en la posición de quienes consideran que, solo con un «gran salto hacia adelante», el comportamiento moderno se hizo posible. El cambio biológico fue expandiéndose a través de generaciones y transformando las manifestaciones culturales de la especie. Solo con un lenguaje complejo se puede expresar las más abstractas reglas sociales, los mitos, el arte con las palabras. Parece razonable suponer que esta complejidad empieza a aparecer con los humanos anatómicamente modernos y llegó a su desarrollo más tardíamente. Solo en las etapas más recientes de la historia de los homínidos, la organización social adquiere un grado de sofisticación que parecería demandar ya habilidades lingüísticas complejas (Lewin, 1999, p. 463). Lo mismo se puede decir de la inteligencia, el aprendizaje general, algunos procesos de memoria, la inhibición de respuesta, las estrategias de procesamiento y pensamiento, y la teoría de la mente.

5. CONCLUSIONES

Como hemos revisado, desde la perspectiva adoptada, el origen y la evolución de la sintaxis pueden fecharse entre hace 200 mil y 50 mil años. Se trata, así, de un fenómeno exclusivamente humano y, más importante, muy reciente. En este artículo, hemos intentado armonizar las propuestas lingüísticas de Bickerton (1990) y de Hauser, Chomsky y Fitch (2002) con los datos que ofrece la paleoantropología a fin de plantear una imagen coherente del contexto en que se habría desarrollado la evolución del componente sintáctico. Como vimos, desde la perspectiva de Bickerton (1990), se puede dar cuenta de la autonomía de la sintaxis, de la estructura argumental y, con ella, de un rasgo de las categorías funcionales, habilidades compatibles con el registro paleoantropológico que va de hace 200 mil a hace 100 mil años. Asimismo, sobre la base de la propuesta de Hauser, Chomsky y Fitch (2002), podemos evaluar un escenario en el que habría surgido el rasgo de recursividad y, junto con él, los rasgos estrictamente gramaticales de las categorías funcionales y el movimiento como mecanismo para su eliminación. Estas habilidades parecen corresponderse conforme con el registro paleoantropológico que va de hace 100 mil a hace 50 mil años. Aun cuando los estudios sobre la evolución del lenguaje todavía deben evaluar la pertinencia de los supuestos teóricos y la validez empírica de las afirmaciones, consideramos que, de manera integrada, contamos con algunos elementos de origen interdisciplinario para ofrecer una aproximación más amplia y coherente del origen y evolución de la sintaxis.

REFERENCIAS BIBLIOGRÁFICAS

- Arbib, Michael (2010). Holophrasis and the Protolanguage Spectrum. En Michael Arbib y Derek Bickerton (eds.), *The Emergence of Protolanguage: Holophrasis vs Compositionality* (pp. 153-166). Ámsterdam: John Benjamins.
- Bickerton, Derek (1990). *Language & Species*. Chicago: The University of Chicago Press.
- Bickerton, Derek (1997). Language Evolution: A Brief Guide for Linguists. *Lingua*, 117(3), 510-526.
- Bickerton, Derek (2009). *Adam's Tongue. How Humans Made Language, How Language Made Humans*. Nueva York: Hill and Wang.
- Calvin, William H. & Derek Bickerton (2000). *Lingua ex Machina. Reconciling Darwin and Chomsky with the Human Brain*. Primera edición. Cambridge: The MIT Press.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge: The MIT Press.
- Chomsky, Noam (2005). Three Factors in Language Design. *Linguistic Inquiry*, 36, 1-22.
- Chomsky, Noam (2007). Of Minds and Language. *Biolinguistics*, 1, 9-27.
- Chomsky, Noam (2011). Language and Other Cognitive Systems. What Is Special About Language? *Language Learning and Development*, 7(4), 263-278.
- Christiansen, Morten H. & Simon Kirby (2003). Language Evolution: Consensus and Controversies. *Trends in Cognitive Sciences*, 7(7), 300-307.
- Coolidge, Frederick L. & Thomas Wynn (2009). *The Rise of Homo sapiens: The Evolution of Modern Thinking*. Malden: Wiley-Blackwell.
- Deacon, Terrence (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. Nueva York: Norton.
- De la Torre Sáinz, Ignacio (2008). *La arqueología de los orígenes humanos en África*. Madrid: Akal.
- Dunbar, Robin (1997). *Grooming, Gossip and the Evolution of Language*. Cambridge: Harvard University Press.
- Fagan, Brian (2002). *World Prehistory: A Brief Introduction*. Quinta edición. Nueva Jersey: Prentice Hall.
- Fitch, W. Tecumseh (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298, 1579-1579.

- Jackendoff, Ray (1999). Possible Stages in the Evolution of the Language Capacity. *Trends in Cognitive Sciences*, 3(7), 272-279.
- Jackendoff, Ray (2002). *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Nueva York: Oxford University Press.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge: The MIT Press.
- Kirby, Simon (2007). The Evolution of Language. En Robin Dunbar y Louise Barrett (eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 669-681). Oxford: Oxford University Press.
- Lewin, Roger (1999). *Human Evolution: An Illustrated Introduction*. Malden: Blackwell.
- Lieberman, Philip (1975). *On the Origins of Language. An Introduction to the Evolution of Human Speech*. Nueva York: Macmillan.
- Lieberman, Philip, Dennis L. Klatt & William A. Wilson (1969). Vocal Tract Limitations on the Vowel Repertoires of Rhesus Monkeys and other Nonhuman Primates. *Science*, 164(3884), 1185-1187.
- Lock, Andrew & Charles R. Peters (eds.) (1999). *Handbook of Human Symbolic Evolution*. Oxford: Blackwell.
- Mithen, Steven (1996). *The Prehistory of the Mind*. Londres: Thames and Hudson.
- Moro, Andrea (2000). *Dynamic Antisymmetry*. Cambridge: The MIT Press.
- Piattelli-Palmarini, Massimo & Juan Uriagereka (2004). The Immune Syntax: The Evolution of the Language Virus. En Lyle Jenkins (ed.), *Variation and Universals in Biolinguistics* (pp. 341-377). Ámsterdam: Elsevier.
- Pinker, Steven (1994). *The Language Instinct*. Nueva York: William Morrow and Company.
- Pinker, Steven & Paul Bloom (1990). Natural Language and Natural Selection. *Behavioral and Brain Sciences*, 13(4), 707-784.
- Tattersall, Ian (2008). *The World from Beginnings to 4000 BCE*. Oxford: Oxford University Press.
- Tomasello, Michael (2008). *Origins of Human Communication*. Cambridge: The MIT Press.
- Wray, Alison (1998). Protolanguage as a Holistic System for Social Interaction. *Language & Communication*, 18, 47-67.

HOMO COMBINANS. EXPLORANDO LA GRAMÁTICA UNIVERSAL MINIMISTA

Paola Cépeda

Pontificia Universidad Católica del Perú

1. INTRODUCCIÓN

Las propuestas de Noam Chomsky sobre el lenguaje aparecen en la década de los cincuenta como una reacción frente a las asunciones conductistas sobre el desarrollo del lenguaje. El modelo de Teoría Estándar (TE) (Chomsky, 1965), en buena cuenta, intenta abordar el denominado Problema de Descartes, la pregunta por el uso creativo del lenguaje. Para poder explicar la creatividad, TE propone una gramática interna que permite la generación de estructuras de frase, así como su transformación en otras a partir de cambios estructurales. Años más tarde, desde fines de la década de los setenta, se va configurando el modelo de Principios y Parámetros (GB, por su etapa Gobierno y Ligamiento) (Chomsky, 1986) como una propuesta para resolver el Problema de Platón aplicado a la competencia lingüística: cómo es que los niños llegan a saber tanto con una limitada experiencia lingüística. El modelo, así, intenta alcanzar un nivel de adecuación teórica explicativa: la teoría lingüística debía decir algo con respecto al proceso de adquisición de una lengua. Para ello, se postula una arquitectura modular del lenguaje, es decir, un conjunto de módulos (o teorías) internos que conferirían a la facultad del lenguaje las cualidades de ser bastante rica, altamente estructurada y, claro, única.

En la década de los noventa, aparece el Programa Minimista (PM) (Chomsky, 1995b). Su proyecto de investigación se alejaba del Problema de Platón, pues buscaba concentrarse en el problema de la inclusión del lenguaje en la arquitectura de la mente. Así, va más allá de un nivel de adecuación explicativo, porque no solo se intenta describir qué propiedades contiene la facultad del lenguaje (los principios de la Gramática Universal [GU]), sino también por qué son de esa manera: si el lenguaje es básicamente un sistema computacional, tiene sentido analizarlo en esos términos, es decir, evaluando su eficiencia computacional. Si en la aproximación de GB

se buscaba dar respuesta a la pregunta de cuánto debemos atribuirle a la GU para dar cuenta de la adquisición de una lengua, PM busca dar respuesta a la pregunta de qué es lo mínimo que debemos atribuirle a la GU si tomamos en cuenta que, como parte de la arquitectura de la mente, el lenguaje se vincula con otros sistemas que le imponen ciertas condiciones para que pueda ser usado (Chomsky, 2007b, p. 4).

Casi al mismo tiempo, empieza a intensificarse la investigación sobre el origen y la evolución del lenguaje. Uno de los elementos centrales en la discusión ha sido la definición misma de lenguaje y qué es aquello que puede hacerlo especial, tan especial que solo una especie sobre el planeta —nosotros, los seres humanos— tiene este sistema lingüístico altamente económico, creativo, simbólico y especializado. En esta comunicación, busco explorar aquellas propiedades que se han visto especiales en la GU y sugerir que lo especial y, quizá, lo exclusivo del lenguaje son los rasgos lingüísticos formales, y no *Merge* o la recursividad. En efecto, los rasgos formales son responsables de la derivación sintáctica (la combinatoria) y desencadenan mecanismos de corrección de imperfecciones; así, la computación lingüística ofrece solo productos óptimos a los sistemas cognitivos con los que interactúa. Estos rasgos combinatorios son piezas claves en las habilidades lingüísticas de nuestra especie, a la que bien podemos denominar *Homo combinans*¹.

Las secciones de esta comunicación están organizadas de la siguiente manera. En primer lugar, presento una breve descripción de la GU minimista. Luego, sucesivamente, exploro *Merge*, la recursividad, los rasgos formales y la externalización, para mostrar que casi todos estos elementos se encuentran o bien en otros sistemas cognitivos humanos, o bien entre las capacidades de otros animales. Digo «casi todos» porque únicamente los rasgos formales aparecen exclusivamente en la capacidad humana del lenguaje. En el último apartado, ofrezco las conclusiones de esta comunicación.

2. LA GRAMÁTICA UNIVERSAL MINIMISTA

En el modelo PM, el lenguaje, como subsistema de la mente/cerebro, es un mecanismo generativo que produce un conjunto de expresiones estructuradas e infinitamente discretas que deben ser interpretadas en dos interfaces: el sistema Conceptual-Intencional (CI), encargado del significado y el pensamiento, y el sistema Sensor-Motor (SM), responsable de la percepción y producción del lenguaje. Para alcanzar una Lengua-I, no basta solo la experiencia lingüística y la dotación genética específica del lenguaje, sino también las condiciones que el funcionamiento de la mente y de los sistemas SM

¹ Tomo prestada esta denominación del nombre de un capítulo del texto *Language in Cognition*, de Cedric Boeckx (2010).

y CI imponen al lenguaje (incluso otras consideraciones extrabiológicas, como una computación mínima o elemental). En este sentido, las condiciones (restricciones y principios) que guían la derivación lingüística ya no son internas a la gramática, sino son más bien externas a ella pues son determinadas por los sistemas con los que interactúa; a estas condiciones se las denomina el «tercer factor» (Chomsky, 2005).

Así, se asume que el lenguaje satisface las condiciones de los sistemas externos de manera óptima gracias a una computación eficiente. Esta idea está expresada en la Tesis Minimista Fuerte (TMF):

La facultad del lenguaje está perfectamente diseñada para cumplir las condiciones que le imponen los sistemas externos (Chomsky, 2007b, p. 4).

La GU minimista, entonces, deja de ser modular (como se había postulado en el modelo previo) y se restringe únicamente a aquello que no puede atribuirse a los sistemas externos al lenguaje. En otras palabras, la GU es lo mínimo que se necesita y que no puede ser explicado por la TMF una vez que se reconocen y se dejan de lado las condiciones y efectos del tercer factor. El carácter biológico de la GU exige, asimismo, una explicación acerca de cómo pudo haber aparecido en el transcurso de la evolución de la especie y, claramente, mientras menos contenga la GU, más sencillo resultará explicar su evolución (Chomsky, 2005).

Chomsky no está en contra de un modelo que apela a la selección natural o a los procesos de adaptación²; su posición es que un modelo adaptacionista no logra explicar el diseño de la GU. No solo la GU es mínima, sino que no parece ser genéticamente variable en la especie. De haber habido muchos cambios genéticos con algún tipo de impacto en el lenguaje, esto habría conllevado diversos niveles de variación visibles en este momento. Sin embargo, actualmente, todas las lenguas son manifestaciones de un mismo sistema cognitivo; no hay evidencia de variación. Frente al argumento de que hubo momentos en que varios rasgos estuvieron en competencia hasta que, después de mucho tiempo, uno de ellos finalmente se asentó en la especie, la evidencia paleoantropológica es contundente: no parece haber habido tanto tiempo para que eso pudiera haber sido el caso (véase Céspedes y Martínez Vera, en este mismo volumen).

² Señala Chomsky: «Las leyes físicas proveen canales estrechos dentro de los cuales los organismos complejos pueden variar, y, sin duda, la selección natural es un factor determinante en la distribución de rasgos y propiedades dentro de esos límites; *un* factor, no *el* factor» (Chomsky, 1995a, p. 56; la traducción es mía). («*Physical law provides narrow channels within which complex organisms may vary, and natural selection is doubtless a factor in determining the distribution of traits and properties within these constraints. A factor, not the factor*»). De aquí en adelante, todas las traducciones que ofrezco son mías.

Entonces, la GU es el producto inesperado y abrupto de la combinación casual de elementos preexistentes en la mente. En este sentido, el lenguaje es como un copo de nieve (citando una metáfora repetida), que toma su forma a partir de leyes naturales, no por presión externa. Esta es, precisamente, la apuesta chomskiana. El mecanismo generativo surge en el individuo tal vez como consecuencia de una pequeña transformación en el cerebro: por ello, debe ser lo más simple posible, el mínimo cambio desde el estado anterior. Así lo dice Chomsky (2000, p. 94)³:

Imaginemos un primate con la arquitectura mental humana, con el aparato sensor-motor ya desarrollado, pero sin órgano del lenguaje. Este primate tiene nuestros modos de organización de la percepción; nuestras actitudes proposicionales (creencias, deseos, esperanzas, temores, etcétera), siempre y cuando no estén mediadas por el lenguaje; quizá un «lenguaje del pensamiento», en el sentido que le da Jerry Fodor. Sin embargo, no tiene ningún modo de expresar sus pensamientos por medio de expresiones lingüísticas, así que esos pensamientos son inaccesibles para él y para otros. Supongamos, además, que algún evento reorganiza el cerebro de tal manera que, como consecuencia, inserta en él la facultad del lenguaje. Para que esta facultad pueda ser usada, el nuevo órgano debe cumplir ciertas «condiciones de legibilidad». Otros sistemas de la mente/cerebro deben poder tener acceso a las expresiones generadas por la facultad del lenguaje (las lenguas-I), «leerlas» y usarlas como «instrucciones» para el pensamiento y la acción. Nosotros podemos intentar formular con claridad (y, si es posible, responder) la pregunta de qué tan buena solución es la facultad del lenguaje para las condiciones de legibilidad y solo ellas. Este es esencialmente el tema del Programa Minimista (Chomsky, 2000, p. 94).

Ahora bien, ¿qué es lo mínimo posible para la GU? ¿Cuál es esa parte de la dotación biológica que queda luego de abstraer las condiciones del tercer factor? ¿Qué es aquello que no puede ser explicado por la TMF? Estas preguntas (que apuntan todas a la misma respuesta) no son nuevas en la teoría lingüística. De hecho, ya se han ofrecido respuestas, concentradas básicamente en dos posibles candidatos para la GU: la computación lingüística llamada *Merge* y la propiedad de la recursividad.

³ «Imagine some primate with the human mental architecture and sensorimotor apparatus in place, but no language organ. It has our modes of perceptual organization, our propositional attitudes (beliefs, desires, hopes, fears, etc.) insofar as these are not mediated by language, perhaps a «language of thought» in Jerry Fodor's sense, but no way to express its thoughts by means of linguistic expressions, so that they remain largely inaccessible to it, and to others. Suppose some event reorganizes the brain in such a way as, in effect, to insert FL. To be usable, the new organ has to meet certain «legibility conditions». Other systems of the mind/brain have to be able to access expressions generated by states of FL (I-languages), to «read» them and use them as «instructions» for thought and action. We can try to formulate clearly—and if possible answer—the question of how good a solution FL is to the legibility conditions, and these alone. That is essentially the topic of the Minimalist Program» (Chomsky, 2000, p. 94).

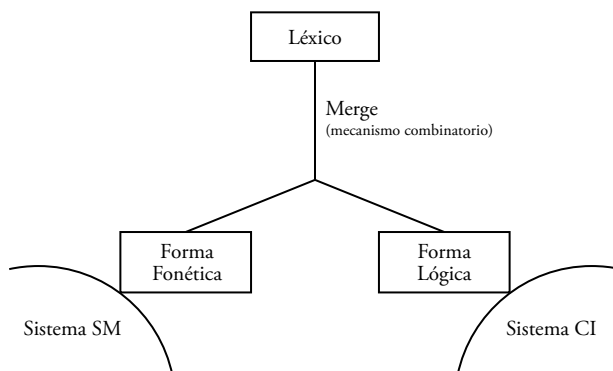
Mi respuesta es diferente. Propongo señalar que aquello que no es explicado por las condiciones de las interfaces es el conjunto de rasgos formales que guían la combinatoria sintáctica. Exploremos cada una de estas tres alternativas, así como los mecanismos para la externalización del lenguaje, desde una perspectiva evolucionista.

3. EXPLORANDO MERGE

En un artículo ampliamente conocido y debatido, Hauser, Chomsky y Fitch (2002) distinguieron una facultad del lenguaje en sentido amplio (FLB, por sus siglas en inglés) y una facultad del lenguaje en sentido estricto (FLN). Para los autores, la primera incluye un conjunto de capacidades cognitivas involucradas con el lenguaje (incluidos los sistemas SM y CI), las cuales, en gran medida, son compartidas con otras especies. La segunda, FLN, incluye un mecanismo lingüístico combinatorio, que es recursivo y que es, de acuerdo con los autores, *exclusivamente humano*. Este mecanismo recibe el nombre de *Merge*.

De acuerdo con Chomsky (2005, 2006 y 2011), la FLN es una solución óptima para vincular los sistemas SM y CI. En efecto, *Merge* construye estructuras lingüísticas de la manera más económica posible y que son plenamente interpretables por estos dos sistemas. *Merge* combina dos objetos lingüísticos (digamos, A y B) para formar uno nuevo (en nuestro caso, {A, B}). La disponibilidad de *Merge* permite un sistema ilimitado de expresiones jerárquicamente estructuradas, lo cual evidencia el carácter infinitamente discreto del lenguaje. *Merge* entrega los productos que construye a dos sistemas externos a él, SM y CI, a través de sus niveles de interface Forma Fonética (FF) y Forma Lógica (FL), respectivamente. Así, un modelo general de la gramática es el que se muestra en (1)⁴:

(1)



⁴ Figura realizada por la autora.

Ahora bien, dado que el lenguaje es parte del mundo natural y, por ello, un sistema físico, asumo (siguiendo el razonamiento de Krivochen, 2012) que comparte las propiedades y principios de los otros sistemas físicos; la diferencia está en los elementos que son manipulados por esos principios. El mismo Chomsky (1965, p. 59) señala que los principios de organización neural pueden estar incluso profundamente enraizados en las leyes físicas. Así, dejando de lado los elementos que *Merge* manipula para el caso del lenguaje, podemos reformular la definición que hemos ofrecido en el párrafo anterior para hacerla más amplia, de modo que pueda aplicarse también a otros sistemas físicos. La noción amplia de *Merge* es que este mecanismo combina dos objetos naturales (digamos, A y B) para formar un nuevo objeto natural (en nuestro caso, {A, B}). Con esta nueva aproximación, resulta claro cómo es que *Merge* puede estar presente en los sistemas naturales y, del mismo modo, ser parte de los sistemas cognitivos distintos del lenguaje (como la composición musical, véase Katz & Pesetsky, 2011).

Si lo anterior es correcto, parece plausible, entonces, suponer que el surgimiento de *Merge* en el lenguaje podría deberse simplemente a la aplicación de una capacidad combinatoria básica a nuevos elementos, y no una capacidad combinatoria radicalmente nueva. Los resultados, sin embargo, de esta aplicación a esos nuevos elementos constituyeron una innovación completa, cuyas consecuencias son la representación y el procesamiento de estructuras jerárquicas recursivas, así como la vinculación de SM y CI a través de esta nueva aplicación⁵.

A manera de una pequeña desviación, valdría la pena preguntarse incluso si *Merge* es un solo mecanismo. Parece que podría desglosarse en los mecanismos del cálculo cognitivo general (Hornstein, 2009 lo llamaría Concatenación) y un añadido, el de la jerarquía, más precisamente, jerarquía endocéntrica (Hornstein, 2009)⁶, que habría mejorado notablemente las capacidades para el pensamiento. Es en este sentido en que debe entenderse la idea de que el diseño del lenguaje aparece como un instrumento del pensamiento:

Tal cambio habría tenido lugar en un individuo, no un grupo. El individuo así dotado habría tenido la habilidad de pensar, planificar, interpretar, de maneras

⁵ A una idea similar llega Krivochen (2012) cuando postula su Tesis Radicalmente Minimista Fuerte (TRMF), que hace eco de la TME, presentada anteriormente. La TRMF señala que las diferencias entre los sistemas físicos son superficiales, pues todos hacen uso de las mismas operaciones y responden a los mismos principios; las diferencias solo se basan en las características de las unidades que se manipulan y en algunos ajustes en la formulación de las operaciones y restricciones (lo que el autor considera que se trata de diferencias en la notación).

⁶ Hornstein (2009) ha propuesto una teoría de la sintaxis a partir de tres primitivos: *Concatenate*, *Copy* y *Label*. Para este autor, el último sería el específicamente lingüístico y permitiría describir la infinitud discreta y las condiciones sintácticas para la derivación de enunciados. Creo que la propuesta que desarrollo en este artículo puede resultar compatible con esa visión en varios aspectos.

novedosas, lo que le produciría ventajas selectivas que transmitiría a su descendencia, hasta expandirse por todo el pequeño grupo de reproducción del que, parece, todos descendemos (Chomsky, 2007b, p. 14)⁷.

Como ya he señalado, entonces, no podemos buscar en *Merge* aquello que es exclusivo del lenguaje, pues bien puede caracterizarse como una combinatoria presente en otros sistemas naturales, incluyendo otros sistemas cognitivos humanos. Lo que sí es especial de *Merge* es que es un mecanismo capaz de manipular unidades que solo el lenguaje puede manipular. Debemos, pues, continuar explorando la GU para averiguar si la otra propiedad considerada especial para el lenguaje, la recursividad, puede ser caracterizada como tal.

4. EXPLORANDO LA RECURSIVIDAD

Una segunda aproximación a aquello que es exclusivamente humano en el lenguaje es apelar a la recursividad. En general, una regla es formalmente recursiva si puede aplicarse a su propio output un número ilimitado de veces. En su respuesta a Hauser, Chomsky y Fitch (2002), Pinker y Jackendoff (2005) llaman «hipótesis de la recursividad» (*recursion-only hypothesis*) a la propuesta de aquellos autores. Como resultado, muchas de las críticas planteadas a Hauser, Chomsky y Fitch (2002) asumen que la recursividad es un producto colateral y toman esta afirmación como elemento central en la discusión. Pinker y Jackendoff, así, expanden el malentendido de que el artículo de 2002 propone que la recursividad es lo estrictamente lingüístico. Esto no es exactamente así. Una lectura atenta de Hauser, Chomsky y Fitch (2002) nos lleva a observar que el mecanismo computacional combinatorio del lenguaje (*Merge*) tiene una propiedad de recursividad, y que esta, de hecho, está presente, de diversas maneras, en distintos sistemas cognitivos, no solo en el lenguaje.

De acuerdo con el malentendido de la «hipótesis de la recursividad», lo único estrictamente lingüístico, en el sentido de FLN, sería la recursividad. En su crítica a Hauser, Chomsky y Fitch, señala Jackendoff⁸:

[...] la recursividad sintáctica es realmente un elemento de la FLB, la cual permite una integración más elegante del lenguaje con el resto del cerebro. Sin duda,

⁷ «Such change takes place in an individual, not a group. The individual so endowed would have the ability to think, plan, interpret, and so on in new ways, yielding selectional advantages transmitted to offspring, taking over the small breeding group from which we are, it seems, all descended».

⁸ «[...] syntactic recursion is actually part of the general-purpose part of the language faculty (FLB), permitting a more graceful integration of language with the rest of the brain. To be sure, language is different, in that it has put recursion to use in mapping between signals and thoughts; that sort of recursion is indeed unique in the animal kingdom. But if recursion occurs in the visual system as well, it does tell us something new about language, pace Chomsky, namely that recursion per se is not what makes language distinctive».

el lenguaje es distinto, porque utiliza la recursividad para relacionar señales y pensamientos; este tipo de recursividad es, de hecho, única en el reino animal. Sin embargo, que la recursividad ocurra también en el sistema visual nos dice algo nuevo acerca del lenguaje, con el debido respeto a Chomsky: que la recursividad per se no es lo que hace al lenguaje distinto (2011, p. 594).

La cita Jackendoff es iluminadora en diversos aspectos. Voy a resaltar dos: en primer lugar, la recursividad sintáctica es un tipo especial de recursividad, única en el reino animal; en segundo lugar, la recursividad está presente en otros sistemas cognitivos.

En cuanto al primer punto, es necesario tener un panorama acerca de aquello a lo que llamamos «recursividad». Específicamente en la sintaxis, una regla es recursiva si se aplica a estructuras de constituyentes que pueden incrustarse en otras de su mismo tipo (*self-embedding*) de manera ilimitada (Fitch, 2010b). Llamemos a este tipo de recursividad *estructural*. Así, la regla de (2a) es estructuralmente recursiva, pues puede aplicarse dentro de sí misma en el elemento A; ello trae como productos (2b), (2c) y otros más:

- (2) a. $A = B A C$
- b. $A = B [B A C] C$
- c. $A = B [B [B A C] C] C$

Pongamos un ejemplo. Tenemos una secuencia *S* del tipo [AABB]. Esta secuencia puede ser construida de diversas maneras. Podría ser una sucesión de elementos (como en [3a]). También podría ser producto de un cálculo: el subíndice *n* indica que la cantidad de veces que aparece A debe ser la misma cantidad de veces que aparece B (como en [3b]):

- (3) a. $S = A A B B$
- b. $S = A_n B_n$

Sin embargo, ni (3a) ni (3b) son estructuralmente recursivos. La misma secuencia *S* sería estructuralmente recursiva únicamente si hay verdadero incrustamiento. Esto significa que los elementos anidados se relacionan de afuera hacia adentro manteniendo la integridad de los constituyentes. Así, *S* es estructuralmente recursiva solo si es producto de cualquier patrón de los que se observa en (4) (adaptado de Corballis, 2007):

- (4) a. $[A [A B] B]$
- b. $A_1 A_j B_j B_1$
- c. 

Los patrones mostrados en (4) muestran que existe incrustamiento de un constituyente [A B] en otro idéntico a él que lo contiene. (4a) muestra la incorporación utilizando corchetes, (4b) emplea subíndices y (4c) una representación gráfica de ramas. Como podemos observar, los patrones de (4), que muestran incrustamiento, son diferentes de los patrones de (3), que no lo muestran. En otras palabras, secuencias linealmente idénticas pueden ser producto de diversas reglas. Únicamente aquellas que presentan verdadero incrustamiento son estructuralmente recursivas, como las expresiones del lenguaje.

¿Pueden construir secuencias estructuralmente recursivas otras especies? No parece ser el caso. Numerosos experimentos han mostrado que los animales tienen una vida cognitiva muy rica, pero parece que carecen de algún componente que constituiría un elemento central del lenguaje. En muchas especies animales, las llamadas siguen reglas o restricciones, y no parece que se dieran al azar (véase un ejemplo significativo en Gentner y otros, 2006). Sin embargo, estas reglas son más simples en comparación con las del lenguaje. Además, la combinatoria del lenguaje tiene que ver estrictamente con estructuras, no con secuencias o sucesiones. No se trata, entonces, solamente de seguir una regla, sino, crucialmente, de reconocer una estructura recursiva, estructuralmente recursiva, como la que he presentado. Aun cuando pueda haber elementos compartidos, de todas maneras, parece existir una gran distancia entre los mecanismos combinatorios de los animales y la sintaxis de las lenguas naturales. Todo esto nos lleva a afirmar, junto con Jackendoff (2011), que, en efecto, este tipo de recursividad es única y especial, y es la que está presente en el lenguaje.

El segundo punto, la presencia de la recursividad en otros sistemas cognitivos, puede quedar más claro con la definición de recursividad estructural que he ofrecido. En efecto, es posible encontrar mecanismos de recursividad en la ubicación en el espacio y direccionamiento, la cognición visual (Pinker & Jackendoff, 2005; Jackendoff, 2011), las estructuras eventivas (Calvin & Bickerton, 2000), la teoría de la mente (Dunbar, 2009), la fabricación de herramientas, entre otros. Sin embargo, no parece que hagan uso de la recursividad estructural. Revisemos dos casos. Primero, la argumentación de Jackendoff (2011) incluía la atribución de la propiedad de recursividad al sistema visual. Aun cuando él mismo afirma que sí se trata de una recursividad estructural, especifica que la bidimensionalidad de los arreglos visuales (*visual arrays*) no puede ser generada por *Merge*, cuyas propiedades gramaticales son distintas (Jackendoff, 2011). Por ello, mi apuesta es que el sistema visual contiene otro tipo de recursividad, diferente de la de las estructuras lingüísticas. Segundo, la recursividad presente en la numeración suele ser vista como una evidencia de que la recursividad no es exclusiva del lenguaje. Sin embargo, nuevamente, no parece que se trate de recursividad estructural para el caso de la numeración. La numeración

infinita puede ser generada por *Merge* y un solo elemento léxico, sin más combinatoria. Además, como señala Chomsky (2007b, p. 7), la infinitud de los números se vincula con el uso del mecanismo, no con la posesión del mismo (y, en ese sentido, es un problema vinculado con la performance, no con la competencia), por lo cual la capacidad numérica sería parasitaria de la facultad del lenguaje, no al revés. El caso de los números, por tanto, no constituiría un ejemplo real.

Sobre la aparición de la recursividad lingüística o estructural, en el proceso evolutivo de la especie, la recursividad habría pasado de ser específica para alguna habilidad cognitiva distinta del lenguaje a ser de dominio general. Las características especiales de la recursividad lingüística podrían obedecer no a un tipo distinto de combinatoria, sino a un tipo distinto de elementos que ingresan en la combinatoria. Así, la propiedad de recursividad del sistema computacional lingüístico es, siguiendo a Hauser, Chomsky y Fitch (2002), un producto colateral de la evolución de alguna otra habilidad cognitiva (como las que he mencionado en el párrafo anterior). La facultad del lenguaje habría aprovechado la propiedad de recursividad para sus propios fines lingüísticos y, en el procesamiento de elementos propios de la combinatoria lingüística, genera un tipo de recursividad particular: la recursividad estructural.

Siguiendo a Jackendoff, podemos afirmar que «la recursividad no es la característica que define al lenguaje; se la puede encontrar en otros sistemas cognitivos» (2011, p. 603)⁹. Y, en cuanto a la recursividad estructural, es el procesamiento de las unidades lingüísticas lo que la genera. En otras palabras, la recursividad estructural no es un primitivo, sino que se deriva de la manipulación de los rasgos formales. Esta idea y la conclusión de que *Merge* es especial en la manipulación de esos rasgos formales (a la que llegamos en el apartado anterior) nos llevan a mirar los rasgos formales como elementos cruciales en la constitución de la GU.

5. EXPLORANDO LOS RASGOS FORMALES

Antes de la aparición de la GU, el sistema conceptual estaba bastante desarrollado. De hecho, en los sistemas cognitivos de los animales, podemos encontrar diversas formas de categorización (representación del mundo, relaciones causales, exteriorización de información reconocimiento de participantes en un evento), teoría de la mente (atribución psicológica), cognición social (pertenecía a un grupo, lazos de parentesco, jerarquías sociales, conflictos), etcétera (Fitch, 2010a). Como apunta Jackendoff¹⁰: «En tanto el pensamiento combinatorio sea posible en un organismo

⁹ «(Recursion) is not the defining characteristic of language; it is found everywhere in higher cognition».

¹⁰ «To the extent that combinatorial thought is possible in an organism without language, we might be correspondingly skeptical of the assumption that syntactic generativity is the source of thought».

sin lenguaje, podríamos igualmente ser escépticos acerca de que el poder generativo de la sintaxis sea la fuente del pensamiento» (2010, p. 72).

Y no lo es. Entre los animales, hay evidencia de mecanismos mentales de cálculo cognitivo (¿protolenguaje?; véase Bickerton, 1990). Lo que falta es un elemento que permita la generación de estructuras sintácticas¹¹ y que, como ya hemos visto, no sea *Merge*. Pues bien, partamos por reconocer que este último no puede operar sin elementos para combinar. Las piezas léxicas son, precisamente esos elementos, los cuales se combinan para formar estructuras sintácticas que sean óptimas para los sistemas SM y CI de FLB.

A propósito de las piezas léxicas, Chomsky señala:

Adicionalmente a *Merge* ilimitado, el lenguaje requiere átomos, o elementos como las palabras, para la computación. Ya sea que pertenezcan estrictamente al lenguaje o a otros sistemas cognitivos, estos átomos plantean serios problemas para el estudio del lenguaje y el pensamiento, y también para el estudio de la evolución de las capacidades cognitivas humanas. El problema básico es que incluso las palabras y conceptos más simples del lenguaje humano y el pensamiento carecen de la relación con entidades independientes de la mente que se han reportado para la comunicación animal (2007a, p. 20)¹².

Pinker y Jackendoff (2005) señalan que cada pieza léxica se almacena en la memoria de largo plazo y está formada por tres tipos de rasgos: fonéticos, semánticos y formales. Los primeros corresponden a los rasgos de pronunciación y son interpretados por el sistema SM. Los segundos corresponden a los rasgos de significado y son interpretados por el sistema CI. Finalmente, los terceros corresponden a los rasgos combinatorios y son leídos por *Merge*, el mecanismo computacional lingüístico.

Los rasgos formales son, entonces, guías para la combinatoria sintáctica. Son ejemplos los rasgos de persona y número en la flexión verbal, así como el rasgo de Caso en las frases nominales. *Merge* lee estos rasgos y sigue sus instrucciones para construir enunciados. La propiedad recursiva de este último está, así, restringida por las instrucciones. Estos rasgos no pueden ser leídos por los sistemas SM y CI,

¹¹ Debe quedar claro que Chomsky no cree que pueda hablarse de predecesores de esta capacidad. Señala él: «Tanto para su evolución como para su desarrollo, no hay muchas razones para postular precursores para *Merge* ilimitado» («*For both evolution and development, there seems little reason to postulate precursors to Unbounded Merge*», Chomsky, 2007a, p. 23).

¹² «*In addition to Unbounded Merge, language requires atoms, or word-like elements, for computation. Whether these belong strictly to language or are appropriated from other cognitive systems, they pose extremely serious problems for the study of language and thought and also for the study of the evolution of human cognitive capacities. The basic problem is that even the simplest words and concepts of human language and thought lack the relation to mind-independent entities that has been reported for animal communication*».

por lo cual se les llama «no interpretables». Por esta razón, *Merge* debe detectarlos y eliminarlos de la derivación sintáctica, de modo que pueda ofrecer a los sistemas de actuación estructuras sintácticas óptimas.

Revisemos un ejemplo:

(5) \emptyset Parece [que los lingüistas están contentos].

En el enunciado anterior, el predicado «parecer» solicita un solo argumento, una cláusula que aparece como su complemento. La posición de sujeto la ocupa un elemento sintáctico sin rasgos fonéticos (esto es, no se escucha) cuya aparición se sigue del principio de que toda oración debe tener sujeto.

(6) Los lingüistas parecen estar contentos.

En el enunciado (6), «los lingüistas» aparece como sujeto de la oración. Su aparición en esta posición obedece a la estrategia de *Merge* de eliminar los rasgos formales de la derivación. En efecto, los rasgos [3.PL] de la flexión del verbo «parecen» son no interpretables para SM o CI. Para deshacerse de esos rasgos, el mecanismo computacional busca rasgos semejantes en la derivación sintáctica que sí sean interpretables, de modo que, a través de la concordancia, los rasgos no interpretables puedan ser cotejados y eliminados. Encuentra, así, a «los lingüistas», cuyos rasgos [3.PL] sí son interpretables. Los rasgos no interpretables son rasgos fuertes, lo que implica que la frase «los lingüistas» debe copiarse desde la cláusula subordinada a la posición de sujeto de la cláusula principal y cotejar, así los rasgos de la flexión. La copia más baja no será leída por el sistema SM:

(7) a. $___$ Parecen_[3.PL] [los lingüistas_[3.PL] estar contentos]
[3.PL] [3.PL]

b. Los lingüistas_[3.PL] parecen_[3.PL] [<los lingüistas> estar contentos]
[3.PL] [3.PL]

Este es solo un ejemplo de que los rasgos formales son los que, efectivamente, están guiando la aplicación de *Merge*.

Los mecanismos combinatorios no son exclusivos de la facultad lingüística. Por esta razón, es necesario aclarar que, cuando se afirma que *Merge* es lo exclusivamente lingüístico (por ejemplo, Chomsky, 2009), no se está hablando de un mecanismo combinatorio cualquiera, sino de un *mecanismo computacional capaz de identificar rasgos meramente formales*. Estos rasgos formales, recordemos, son no interpretables, por lo cual deben ser eliminados en el transcurso de la derivación para que los sistemas de actuación externos solo reciban productos óptimos, con rasgos interpretables (perfectamente legibles). Como producto de la derivación generada por el

mecanismo computacional y dirigida por los rasgos formales aparecen las estructuras recursivas. La recursividad es, así, una propiedad del mecanismo computacional que constituye FLN, *no es FLN*.

No parece haber ninguna otra especie que tenga piezas léxicas como las que combina *Merge*. De hecho, no parece que otro sistema cognitivo humano haga uso de elementos combinatorios formales como lo hace este con los rasgos formales para el lenguaje. Así, pues, dado que *Merge* sigue las instrucciones de los rasgos formales de las piezas léxicas seleccionadas para crear estructuras sintácticas óptimas, el resultado es especial y, quizá, exclusivo del lenguaje.

Piattelli-Palmarini y Uriagereka (2004) desarrollan la Teoría del Virus (*Virus Theory*) para ofrecer una propuesta de cómo aparecieron los rasgos no interpretables en la línea evolutiva del lenguaje. De acuerdo con estos autores, los virus pueden instalarse en un individuo por contacto (transmisión horizontal) e insertarse de manera estable en el genoma de su portador; una vez allí, pueden ser transmitidos por herencia (transmisión vertical) a la siguiente generación. La nueva función del virus, entonces, debe ser adaptativa, de modo que pueda ser preservada por selección natural. En el caso del lenguaje, los rasgos combinatorios parecen comportarse como virus. Piattelli-Palmarini y Uriagereka proponen que, en un momento de la evolución, una forma como el protolenguaje de Bickerton (1990) pudo haberse visto afectada por un cambio en la reorganización de la mente/cerebro. Este cambio generaría que las estructuras sintácticas sean más complejas a partir de un elemento extraño que no es interpretable ni por SM ni por CI. El nuevo sistema, entonces, debe detectar al intruso y generar una «respuesta inmune», es decir, un mecanismo para eliminarlo haciendo uso de los recursos con los que ya cuenta. Estos «antígenos» morfológicos, entonces, son detectados y eliminados por los «anticuerpos» sintácticos atrayendo a los rasgos a posiciones estructurales donde puedan reconocerse como rasgos idénticos a ciertos elementos que sí son interpretables. Así como luego de una única exposición a un virus las células inmunes memorizan al intruso, del mismo modo, los mecanismos de la computación sintáctica pueden operar siguiendo las instrucciones ofrecidas por cada rasgo combinatorio que se comporta como un virus en la derivación. Aun cuando esta propuesta resulta interesante y única para explicar cómo es que proliferaron los rasgos formales en las estructuras sintácticas, ambos autores reconocen que todavía hace falta ir más allá de la metáfora con evidencia lingüística, biológica, genética y neurológica.

El panorama, en general, resulta alentador. La asunción chomskiana es que la aparición de la GU en la historia evolutiva de la especie es consecuencia de la reorganización anatómica del cerebro a partir de su enorme crecimiento en un periodo evolutivo relativamente corto. Este crecimiento del cerebro tendría como

consecuencia la incorporación de *eso que es especial en el lenguaje* en un sistema cognitivo que ya contaba con otras capacidades que resultan cruciales para el lenguaje (como el de los actuales simios o como el de nuestros ancestros en la línea de los homínidos). Revisemos rápidamente estas asunciones.

En cuanto a la reorganización de procesos cognitivos, parece haber habido cambios en las funciones generales de ciertos mecanismos cerebrales que han resultado útiles para el lenguaje. Son ejemplos de estos cambios las estructuras subcorticales que actualmente asumen gran parte de las tareas de procesamiento secuencial que implica el lenguaje, así como las neuronas espejo identificadas en el área de Broca de los primates (Benítez Burraco, 2009, pp. 40-41), entre otras. Desde un punto de vista evolutivo, entonces, una nueva habilidad cognitiva puede aparecer de manera muy rápida gracias a la reutilización de capacidades preexistentes, dedicadas a una tarea en particular, con fines distintos de aquellos para los cuales fueron seleccionadas. *Merge* aplicaría sus capacidades combinatorias al servicio de «inmunizar» la derivación sintáctica de los rasgos no interpretables.

En cuanto a la explicación de la GU como producto derivado del crecimiento cerebral, esta idea podría ser compatible con dos propuestas que provienen de las ciencias cognitivas: una gradualista y otra no. La primera es la Robin Dunbar (2009), que alude a un crecimiento cerebral y una reorganización mental progresivos seleccionados por la capacidad para manejarse en grupos sociales con cada vez más individuos. Para él, existe una relación directa entre el tamaño de los grupos sociales y el tamaño del neocórtex. La segunda propuesta es la de Steve Dorus y otros (2004), que enfatiza el ritmo acelerado del crecimiento del cerebro y su reorganización en la especie *Homo sapiens*. Su hipótesis es que hubo miles de mutaciones repentinas que afectaron a los genes vinculados con el sistema nervioso y que afectaban el crecimiento del cerebro. El resultado ventajoso habría permitido su rápida propagación.

La propuesta de que los rasgos formales podrían ser lo exclusivo del lenguaje encaja naturalmente en el marco generado por la teoría lingüística y la discusión en otras ciencias cognitivas. No serían *Merge* ni la recursividad aquellos elementos exclusivos de la capacidad lingüística, sino los rasgos combinatorios no interpretables por los sistemas con los que el lenguaje interactúa. Aun así, todavía resulta necesario explorar la externalización del lenguaje en el siguiente apartado.

6. EXPLORANDO LA EXTERNALIZACIÓN

Una vez que la GU queda asentada en el sistema cognitivo humano, se externaliza a través del sistema SM. Lo señala Chomsky: «En ese momento, habría una ventaja en la externalización, así que la capacidad se vincularía como un proceso secundario

con el sistema sensoriomotor para la externalización y la interacción, incluyendo la comunicación» (2006)¹³.

¿Por qué sería la externalización un proceso secundario? Chomsky asume que la aparición de la GU confiere a su poseedor una ventaja adaptativa que, en principio, no se refiere a sus mecanismos de externalización, sino a la propia naturaleza interna del lenguaje: como herramienta para el pensamiento. Lo dice así:

En realidad, se puede usar el lenguaje incluso si somos la única persona en el universo con lenguaje, y de hecho tendríamos una ventaja adaptativa. Si de repente una persona desarrollara la facultad del lenguaje, tendría grandes ventajas: podría pensar, podría articular sus pensamientos para sí mismo, podría hacer planes, podría afinar y desarrollar su pensamiento como lo hacemos con el monólogo interior, que ha tenido gran efecto en nuestras vidas (Chomsky, Belletti & Rizzi, 2002, p. 148)¹⁴.

La computación mínima genera, en principio, elementos lingüísticos interpretables por CI; solo posteriormente serían adaptados para ser producidos y percibidos por SM. La aparición de la GU, entonces, habría conferido a su poseedor, en principio y crucialmente, una mejor capacidad para organizar su propio pensamiento. Reformulando la frase extendida de que «el lenguaje es sonido y significado», Chomsky (2011) afirma que el lenguaje es significado con sonido. Dentro de esta visión, el hecho de que utilicemos el lenguaje externamente para interactuar con otros no es algo consustancial al lenguaje. Esta es una de las primeras ideas chomskianas acerca del lenguaje. Sin embargo, se ha querido ver en ella el rechazo de Chomsky a la relación entre lenguaje y comunicación. Este «rechazo» es irreal y no tiene lugar ni en las ideas chomskianas ni en el marco que estamos adoptando.

Una pregunta que surge en este momento es la de por qué la GU no continuó siendo una forma mejorada del «lenguaje del pensamiento», como señala Chomsky, sino que tuvo que externalizarse. Lorenzo (2007) intenta ofrecer una respuesta a esta pregunta señalando que la reorganización anatómica del cerebro a partir de su enorme crecimiento podría haber generado un inesperado contacto entre las zonas cerebrales dedicadas al control vocal y aquellas dedicadas a la conceptualización del entorno (natural o social). Así, se pudo comenzar a exteriorizar las representaciones del sistema CI gracias al mecanismo de la GU ya implementado. Sin embargo,

¹³ «At that stage, there would be an advantage to externalization, so the capacity would be linked as a secondary process to the sensorimotor system for externalization and interaction, including communication».

¹⁴ «Actually you can use language even if you are the only person in the universe with language, and in fact it would even have adaptive advantage. If one person suddenly got the language faculty, that person would have great advantages; the person could think, could articulate to itself its thoughts, could plan, could sharpen, and develop thinking as we do in inner speech, which has a big effect on our lives».

a mí me deja el sabor de que se trata de una salida intrincada para un problema que puede explicarse de otra manera, en términos más sencillos, más económicos. En efecto, sin salir del paradigma de PM, la vinculación GU-SM resulta altamente adaptativa puesto que, en efecto, la externalización del pensamiento no es costosa, sino que se hace a través de un mecanismo de infinitud discreta. Una sugerencia de que esto podría ser así es la repetida afirmación de Chomsky del conflicto entre mecanismos computacionales y facilidad comunicativa:

Las condiciones de una computación eficiente entran en conflicto con la facilidad de comunicación. Universalmente, las lenguas prefieren una computación eficiente. Esto parece ser generalizadamente verdadero. Si es así, el lenguaje pudo haber evolucionado, y tener un diseño, primariamente como un instrumento del pensamiento, con la externalización como un proceso secundario. La comunicación, un caso especial de externalización, también sería un proceso secundario, otra diferencia crucial entre el lenguaje y los sistemas simbólicos de otros animales (2006)¹⁵.

Otro punto que Chomsky considera en su argumentación es el de la disociación entre lenguaje y SM. Para ello, ofrece, por lo menos, dos argumentos. En primer lugar, en cuanto a su aparición en la historia de la especie, los elementos que constituyen el sistema SM estaban presentes cientos de miles años antes de la emergencia del lenguaje. Fitch (2010a) propone que la laringe de las cabras y algunos monos desciende durante las vocalizaciones hasta una posición más o menos semejante a la de los humanos modernos. Lo interesante es que la revisión de sus esqueletos no ofrece esa información: solo es posible observar el dinamismo en sus articulaciones vocálicas en el momento mismo en que las vocalizaciones se están produciendo. Por extensión, se podría especular que una situación semejante podría haber estado presente en los neandertales. La diferencia crucial con los mecanismos humanos está en la posición permanentemente descendida de la laringe en los humanos adultos. Por otro lado, se sabe que los huesos del oído medio funcionan como un sistema amplificador de sonido, que puede capturar distintas frecuencias en los sonidos del habla; sin embargo, estos mismos huesos han evolucionado de las mandíbulas de los reptiles como un efecto del crecimiento del neocórtex en los mamíferos hace 160 millones de años (Chomsky, 2003, p. 274).

¹⁵ «Conditions of efficient computation conflict with facilitation of communication. Universally, languages prefer efficient computation. That appears to be true quite generally. If so, it appears that language evolved, and is designed, primarily as an instrument of thought, with externalization a secondary process. Communication, a special case of externalization, would also be a secondary process, another crucial difference between human language and the symbolic system of other animals».

En segundo lugar, respecto de la dinámica social de las lenguas, la adquisición y la variación están enfocadas, exclusivamente, en los elementos de la externalización. En efecto, Chomsky plantea que el proceso de adquisición de una lengua consiste, básicamente, en el aprendizaje de los elementos que se externalizan, como la fonética, la fonología, las señas, etcétera, y lo que se interpreta a partir de ellos, pero nunca consiste en aprendizaje de los principios de la computación lingüística. Del mismo modo, el cambio lingüístico muestra que la variación y la diversidad en las lenguas encuentran su fuente en los mecanismos de la externalización, que son los que están permanentemente sujetos al cambio. El mecanismo computacional lingüístico es el mismo en todas las lenguas.

No cabe duda de que es una ventaja adaptativa externalizar el pensamiento y compartir información dentro del grupo, sobre todo con los parientes y, en particular, con la descendencia, de modo que se pueda perpetuar la supervivencia. Y hacerlo con la GU, un mecanismo de infinitud discreta, resulta simple y poco costoso. Sin embargo, los mecanismos computacionales y los elementos combinatorios descritos para la facultad del lenguaje no han sido diseñados primariamente para su externalización, sino que esta constituye un elemento circunstancial y secundario en la historia evolutiva de la capacidad lingüística.

7. CONCLUSIONES

He explorado algunos elementos que se han observado previamente como constituyentes esenciales de la GU. En primer lugar, *Merge* no puede ser exclusivo de la capacidad lingüística, pues luce más como un mecanismo combinatorio presente en otros sistemas físicos; su presencia en el lenguaje es solo una evidencia más de que el lenguaje es parte del mundo natural. En segundo lugar, el tipo de recursividad que es relevante en el lenguaje es la recursividad estructural, que incorpora el anidamiento o incrustación de elementos del mismo tipo en otros. Sin embargo, como propiedad general, no puede constituir la FLN, sino que es un producto derivado de la combinatoria de los elementos lingüísticos. En tercer lugar, los rasgos formales no interpretables parecen ser lo «especial» del lenguaje: no solo son piezas únicas en el mundo natural, sino que dirigen la combinatoria sintáctica dándole a *Merge* las propiedades que ostenta y generando recursividad estructural. Finalmente, aun cuando el proceso de externalización es altamente ventajoso y permite que el lenguaje se usado, es evolutivamente secundario al diseño de la GU.

Yendo más allá de lo presentado, algunos autores consideran que, en su estado actual, el lenguaje es especial, pues solo con él es posible romper las barreras de la modularidad (Boeckx, 2010), dado que este puede recoger y combinar información

de distintos sistemas centrales de conocimiento que, de otra manera, permanecerían aislados. En este sentido, el lenguaje sería una nueva forma de representación (Bickerton, 1990) que potenciaría otras capacidades cognitivas como la representación espacial, la organización de los números y la teoría de la mente.

¿Cuál es el siguiente paso en este programa? No estoy segura de que debamos apuntar hacia asuntos novedosos. Creo que lo prudente será regresar a aquellas preguntas que se respondieron previamente o que aún no han terminado de ser respondidas, como el uso creativo del lenguaje, la adquisición de un sistema lingüístico, la relación entre lenguaje y los otros sistemas cognitivos con los que interactúa, y también la evolución y el desarrollo de la capacidad lingüística en la especie. Con toda la experiencia ganada en los más de cincuenta años de este programa científico y con una visión interdisciplinaria en el estudio del lenguaje, revisar lo que se ha hecho y ofrecer respuestas a las viejas preguntas desde una nueva perspectiva resultará, sin duda, «un gran salto hacia adelante».

REFERENCIAS BIBLIOGRÁFICAS

- Benítez Burraco, Antonio (2009). *Genes y lenguaje: aspectos ontogenéticos, filogenéticos y cognitivos*. Barcelona: Reverté.
- Bickerton, Derek (1990). *Language & Species*. Chicago: The University of Chicago Press.
- Boeckx, Cedric (2010). *Language in Cognition*. Malden: Wiley-Blackwell.
- Calvin, William H. & Derek Bickerton (2000). *Lingua Ex Machina. Reconciling Darwin and Chomsky with the Human Brain*. Cambridge: The MIT Press.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.
- Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Nueva York: Praeger.
- Chomsky, Noam (1995a). Language and Nature. *Mind*, 104(413), 1-61.
- Chomsky, Noam (1995b). *The Minimalist Program*. Cambridge: The MIT Press.
- Chomsky, Noam (2000). Minimalist Inquires: The Framework. En Roger Martin, David Michaels y Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* (pp. 89-155). Cambridge: The MIT Press.
- Chomsky, Noam (2005). Three Factors in Language Design. *Linguistic Inquiry*, 36, 1-22.
- Chomsky, Noam (2006). *Bilingualistic Perspectives. Design, Development, Evolution* [manuscrito]. Conferencia magistral. Lima, Pontificia Universidad Católica del Perú.
- Chomsky, Noam (2007a). Of Minds and Language. *Biolinguistics*, 1, 9-27.

- Chomsky, Noam (2007b). Approaching UG from Below. En Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and The View from Syntax-Semantics* (pp. 1-29). Berlín: Mouton de Gruyter.
- Chomsky, Noam (2009). Opening Remarks. En Massimo Piattelli-Palmarini, Juan Uriagereka & Pello Salaburu (eds.), *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country* (pp. 13-43). Oxford: Oxford University Press.
- Chomsky, Noam (2011). Language and Other Cognitive Systems. What Is Special About Language? *Language Learning and Development*, 7, 263-278.
- Chomsky, Noam, Adriana Belletti & Luigi Rizzi (2002). *On Nature and Language*. Cambridge: Cambridge University Press.
- Corballis, Michael C. (2007). Recursion, Language, and Starlings. *Cognitive Science*, 31(4), 697-704.
- Dorus, Steve y otros (1994). Accelerated Evolution of Nervous System Genes in the Origin of Homo sapiens. *Cell*, 119, 1027-1040.
- Dunbar, Robin (2009). Why Only Humans Have Language. En Rudolph Botha & Chris Knight (eds.), *The Prehistory of Language* (pp. 12-35). Oxford: Oxford University Press.
- Fitch, W. Tecumseh (2010a). *The Evolution of Language*. Cambridge: Cambridge University Press.
- Fitch, W. Tecumseh (2010b). Three Meanings of «Recursion»: Key Distinctions for Biolinguistics. En Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives* (pp. 73-90). Cambridge: Cambridge University Press.
- Gentner, Timothy Q. y otros (2006). Recursive Syntactic Pattern Learning by Songbirds. *Nature*, 440, 1204-1207.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298, 1579-1579.
- Hornstein, Norbert (2009). *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.
- Jackendoff, Ray (2010). Your Theory of Language Evolution Depends on Your Theory of Language. En Richard K. Larson, Viviane Déprez y Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives* (pp. 63-72). Cambridge: Cambridge University Press.
- Jackendoff, Ray (2011). What is the Human Language Faculty? Two Views. *Language*, 87(3), 586-624.

- Katz, Jonah & David Pesetsky (2011). *The Identity Thesis for Language and Music*. *LingBuzz*. <http://ling.auf.net/lingbuzz/000959>
- Krivochen, Diego G. (2012). *An Introduction to Radical Minimalism II: Internal Merge beyond Explanatory Adequacy*. *LingBuzz*. <http://ling.auf.net/lingbuzz/001256>.
- Lorenzo, Guillermo (2007). ¿Es el lenguaje una adaptación? En Juan R. Coca (coord.). *Varia biológica. Filosofía, ciencia y tecnología* (pp. 189-219). León: Centro de Estudios Metodológicos e Interdisciplinarios de la Universidad de León.
- Piattelli-Palmarini, Massimo & Juan Uriagereka (2004). The Inmune Syntax. The Evolution of the Language Virus. En Lyle Jenkins (ed.), *Variations and Universals in Biolinguistics* (pp. 341-377). Ámsterdam: Elsevier North-Holland.
- Pinker, Steven & Ray Jackendoff (2005). The Faculty of Language: What's Special About It? *Cognition*, 95, 201-236.
- Uriagereka, Juan (1998). *Rhyme and Reason. An Introduction to Minimalist Syntax*. Cambridge: The MIT Press.

EL «ARGUMENTO DEL DISEÑO» Y LA CRÍTICA CHOMSKIANA A LA EXPLICACIÓN ADAPTACIONISTA DE LA EVOLUCIÓN DEL LENGUAJE

Marcos Herrera Burstein

Pontificia Universidad Católica del Perú

1. INTRODUCCIÓN: LA EVOLUCIÓN DEL LENGUAJE Y EL ARGUMENTO DEL DISEÑO

Se puede afirmar que una de las contribuciones más importantes de Chomsky al estudio de la mente ha sido su crítica del paradigma conductista, que favoreció el surgimiento del paradigma cognitivo a finales de los años cincuenta. Su reseña de *Verbal Behavior* de Skinner (1957) y en general la manera en que se propone dar cuenta de la facultad humana del lenguaje (Chomsky, 1977[1959]) representan un radical cuestionamiento del «mandamiento» fundamental del conductismo de no postular conceptos mentalistas —es decir, conceptos que presuponen estructuras, procesos o representaciones mentales— para explicar el comportamiento o las capacidades de los seres humanos. En el enfoque de Chomsky juega un papel central la noción de la *creatividad lingüística*, entendida como la capacidad de los hablantes de producir y comprender infinitas oraciones. Chomsky (1966, p. 53) apela a una clásica noción de Wilhelm von Humboldt, que es que el lenguaje hace «un uso infinito de medios finitos» (1988[1836], p. 108; la traducción es mía)¹. Este es el *gran truco del lenguaje*. Dichos medios finitos, con los que el hablante puede producir infinitas expresiones, constituyen una gramática, entendida como una estructura cognoscitiva representada en la mente del hablante. Chomsky muestra de este modo que una explicación del lenguaje que incorpora conceptos mentalistas es más satisfactoria que una que deliberadamente renuncia a ellos, como la de Skinner (1957).

Pero Chomsky es especialmente conocido por su tesis de que la facultad humana del lenguaje es innata. Hay que aclarar que lo que es innato, para Chomsky, no es la capacidad del hablante adulto de hablar una lengua materna particular, como el castellano, el quechua o el inglés, sino la capacidad del niño de desarrollar la gramática

¹ En el texto original: «*von endlichen Mitteln einen unendlichen Gebrauch machen*».

de cualquier lengua particular, a la que llama la Gramática Universal. Para defender esta tesis Chomsky se basa en el llamado *argumento de la pobreza del estímulo* (1965, 1975 y 1979), según el cual el tiempo de adquisición de dicha gramática particular es demasiado corto, los estímulos son demasiado fragmentarios, y las estructuras cognitivas resultantes (es decir, la gramática particular) son demasiado complejas y homogéneas en diferentes hablantes, como para que parte importante de lo que se encuentra en el estadio final de dicho desarrollo no haya estado presente ya en el estadio inicial del mismo. De este modo, el conocimiento resultante al final del proceso contiene más de lo que dichos estímulos pueden enseñar. Es lo que Chomsky (1986) denominó el problema de Platón.

A pesar de su carácter controversial, esta idea de que el lenguaje no sería una creación cultural sino más bien parte de la dotación biológica de nuestra especie ha ido ganando cada vez mayor difusión, en parte gracias a obras de divulgación, en particular Pinker (1994). En nuestro idioma, los psicolingüistas españoles De Vega y Cuetos escriben (tomando una imagen de Pinker y Bloom, 1990) que «estos hechos [...] sugieren que el lenguaje es un rasgo biológico específico de nuestra especie. Es más un «instinto» (como la ecolocalización del delfín, la habilidad tejedora de la araña o la visión estereoscópica de los primates) que una invención cultural (como la rueda o la escritura)» (1999, p. 14).

Sin embargo, sostener que una determinada capacidad cognitiva es innata, es decir, que es parte de la dotación biológica de una especie nos obliga a plantear una pregunta fundamental. Desde Darwin la biología nos enseña que aquellas capacidades con las que se compara aquí al lenguaje, esto es, la capacidad del delfín para la ecolocalización, la capacidad de la araña para tejer su tela o la capacidad de los primates para ver de manera estereoscópica son el producto de la evolución, mediante el mecanismo de la selección natural (Dawkins, 1995). Según el enfoque neodarwinista de la evolución, un determinado rasgo que puede surgir aleatoriamente como producto de mutaciones genéticas en algunos individuos de una especie puede ser «seleccionado» si, ante determinados cambios del entorno, confiere a los individuos que lo poseen una mayor capacidad para adaptarse a dichos cambios, sobrevivir y reproducirse, pasando dichos rasgos a sus descendientes, los que se impondrían a los otros individuos que carecen de tales rasgos. Por tanto, si se postula que la capacidad cognitiva del lenguaje es innata y parte de la dotación biológica de la especie humana, habría que enfrentar el desafío de explicar cómo podría haber evolucionado como producto de la selección natural.

Científicos cognitivos como Steven Pinker y Paul Bloom (1990) por un lado, y Ray Jackendoff (1999) por el otro, han acometido dicho desafío, planteando hipótesis explicativas acerca de cómo podría haber ocurrido este proceso, basándose

en el llamado «argumento del diseño» propuesto por el biólogo Richard Dawkins (1986). El argumento del diseño había sido utilizado tradicionalmente por pensadores creacionistas como el teólogo William Paley, quien a inicios del siglo XIX sostenía que si algo tiene un diseño, como un reloj, entonces debería existir un diseñador de este diseño, un relojero, concluyendo así que los organismos del mundo natural, que tienen un diseño muy complejo, también deberían tener un diseñador, por lo que tendría que existir un creador. Dawkins muestra que dicho argumento debe ser interpretado de modo fundamentalmente diferente desde Darwin, pues el diseño de los organismos, por complejo que sea, puede explicarse siempre como un resultado del moldeamiento gradual por parte de la selección natural, la que opera paso a paso a lo largo de enormes periodos de tiempo. Un ejemplo paradigmático es la evolución del complejo diseño del ojo de los vertebrados. Así, según Dawkins, la presencia de un diseño complejo en el mundo natural no debe ser tomada como una evidencia de un dios creador, sino más bien como una evidencia de la acción de la selección natural. Este argumento asume además que dicho diseño sirve siempre para alguna función que incrementa las posibilidades de supervivencia del organismo, y constituye por tanto una adaptación para la misma. El ojo, por ejemplo, sería una adaptación para la función de la visión, y cada paso en su evolución implica una mayor eficiencia en el cumplimiento de dicha función. Apoyándose en el argumento del diseño, Pinker y Bloom (1990) y Jackendoff (1999) han señalado que la facultad del lenguaje muestra también, como el ojo, signos de un diseño complejo, lo que para ellos sería una evidencia incontestable de que se trata de una adaptación moldeada gradualmente por la selección natural. Y la función a la que respondería dicha adaptación sería la comunicación de proposiciones (pensamientos) por un canal serial.

Uno podría esperar que Chomsky hubiese recibido con beneplácito estas propuestas de armonizar su visión innatista y biologista del lenguaje con la teoría neodarwiniana estándar de la evolución mediante el mecanismo de la selección natural. Pero como han señalado Pinker y Bloom (1990), sorprendentemente Chomsky ha rechazado desde hace mucho tiempo y de modo tajante la tesis de que el lenguaje sería una adaptación moldeada por la selección natural, y en particular la idea de que sería una adaptación para la función de la comunicación. Esto ha generado un importante debate, que describiré a continuación².

² Ver también la presentación de la evolución del lenguaje en la primera parte de este volumen. Para otras explicaciones de la evolución del lenguaje, confróntese Bickerton (1990) y Dunbar (2009).

2. LA BATALLA POR LA EVOLUCIÓN DEL LENGUAJE: ADAPTACIÓN O EXAPTACIÓN

La posición de Chomsky ha sido desarrollada y fundamentada extensamente en un clásico artículo de Massimo Piattelli-Palmarini (1989), quien cuestiona la llamada concepción neodarwiniana de la evolución que hemos descrito más arriba. Piattelli-Palmarini anuncia así que en el estudio de la evolución este enfoque neodarwiniano estaría siendo desplazado por un nuevo enfoque neo-neodarwiniano (Gould & Lewontin, 1979; Gould & Vrba, 1982, citados por Piattelli-Palmarini, 1989). Para estos autores la selección natural sería solo uno entre varios mecanismos disponibles para la evolución, en la que otros mecanismos no adaptativos, como los *spandrels*, los productos colaterales (*by-products*) o la exaptación, juegan también un rol fundamental³. Piattelli-Palmarini critica las teorías adaptacionistas o neodarwinianas de la evolución del lenguaje, señalando que es difícil imaginar cómo la selección natural podría haber moldeado gradualmente todos los mecanismos que forman parte de la facultad del lenguaje (entendida por él en términos de la teoría de principios y parámetros desarrollada por Chomsky en los años ochenta), mecanismos que en su opinión son gratuitos, o cómo en dicho proceso cada nuevo estadio podría haber conferido ventajas adaptativas adicionales a los individuos. Como alternativa propone una explicación neo-neodarwiniana de acuerdo a la cual la facultad del lenguaje sería el resultado de mecanismos no adaptativos, en otras palabras, sería un *spandrel*, un producto colateral o una exaptación (Piattelli-Palmarini, 1989, p. 19).

En su respuesta a este artículo Pinker y Bloom (1990), defensores de la postura adaptacionista, insisten en que el lenguaje sí habría sido moldeado por la selección natural, tal como lo plantea la teoría neodarwiniana estándar. Una de las principales razones que invocan en defensa de esta tesis es precisamente el argumento del diseño de Dawkins (1986), pues para estos autores, el punto débil de la crítica de Gould y Lewontin al adaptacionismo es que la selección natural sería la única explicación

³ La noción de *spandrel* es tomada por Gould de la arquitectura, en la que designa un espacio triangular que resulta al erigir una cúpula sobre arcos que se unen en ángulos rectos, y que ha sido usado para representaciones pictóricas, pero sin haber sido diseñado específicamente con esa finalidad. Análogamente, Gould llama *spandrel* a una propiedad biológica que es un producto secundario de otras características arquitectónicas de los organismos y que no ha sido seleccionada en sí misma por proporcionar ventajas adaptativas. Un *by-product* o *producto colateral* es un rasgo que ha sido seleccionado no porque proporcione ventajas adaptativas, sino porque es el producto de un gen que está asociado a otro gen responsable de un rasgo que sí proporciona ventajas adaptativas (por ejemplo, el color blanco de los huesos, que es un producto colateral del rasgo de contener calcio, seleccionado por su fortaleza). La *exaptación*, finalmente, es un mecanismo mediante el cual un rasgo que tenía una función determinada, o que no tenía ninguna función, asume una función que no tenía anteriormente (por ejemplo, las alas de los insectos que originalmente servían a la función de regulación térmica, pero luego fueron exaptadas para la función del vuelo) (ver Piattelli-Palmarini, 1989; Buss y otros, 1998; Pinker & Bloom, 1990).

científica disponible para dar cuenta de la complejidad adaptativa (1990, p. 709). Mecanismos no adaptativos como *spandrels*, *by-products* o *exaptaciones* pueden ser responsables de la evolución de rasgos biológicos simples, pero no serían capaces de producir una estructura con un diseño complejo, como el ojo de los vertebrados. De aquí Pinker y Bloom concluyen que el lenguaje humano, que también presenta un diseño complejo, debe haber evolucionado mediante la selección natural (p. 726), señalando además que una ventaja adaptativa fundamental que el lenguaje podría haber proporcionado a nuestros antepasados es la complejización de las interacciones sociales: «*The social value of complex language probably played a profound role in human evolution*» (p. 725). De hecho, estos autores otorgan un peso muy importante a la cognición social en el desarrollo del cerebro humano en general: «*After all, it doesn't take all that much brain power to master the ins and outs of a rock or to get the better of a berry. But interacting with an organism of approximately equal mental abilities whose motives are at times outright malevolent makes formidable and ever-escalating demands on cognition*» (p. 725).

Es en el contexto de esta controversia que Chomsky, conjuntamente con los biólogos Marc D. Hauser y W. Tecumseh Fitch, publicó en 2002, en la revista *Science*, un influyente artículo en que formula una versión nueva y más fuerte de la explicación no adaptacionista de la evolución de la facultad del lenguaje. Dicha publicación dio lugar a un artículo de respuesta de Steven Pinker y Ray Jackendoff en la revista *Cognition*, en 2005, defendiendo la postura adaptacionista, a la que siguieron, ese mismo año y en la misma revista, una réplica de Fitch, Hauser y Chomsky y una contrarréplica de Jackendoff y Pinker. A continuación, presentaré brevemente dicho debate para discutir luego en mayor detalle algunos aspectos importantes de la argumentación antiadaptacionista de Hauser, Chomsky y Fitch (2002) y Fitch, Hauser y Chomsky (2005); en dicha discusión tomaré en cuenta también otros dos trabajos de Chomsky (2000 y 2010).

Hauser, Chomsky y Fitch defienden, en la misma línea que Piattelli-Palmarini, la tesis de que la facultad del lenguaje, a pesar de ser un producto de la evolución, no sería —al menos en sus aspectos más nucleares— una adaptación moldeada por la selección natural para la comunicación. Pero la argumentación de Hauser, Chomsky y Fitch se distingue en dos puntos centrales de la de Piattelli-Palmarini. El primero es la introducción (o explicitación) de una distinción dentro de la facultad del lenguaje entre la *Facultad del lenguaje en sentido amplio* (*Faculty of language - broad sense*) o FLB, por sus siglas en inglés, y la *Facultad del lenguaje en sentido estrecho* (*Faculty of language - narrow sense*), o FLN, por sus siglas en inglés. FLN sería el sistema computacional lingüístico abstracto y constituiría un componente de FLB que interactuaría con sus otros dos componentes, el sistema sensorio-motor y el sistema

conceptual-intencional (2002, p. 1570). FLN generaría representaciones internas y las proyectaría en la interface sensorio-motora mediante el sistema fonológico y en la interface conceptual-intencional mediante el sistema semántico (formal). El segundo punto en que el artículo de Hauser, Chomsky y Fitch se distingue de la propuesta de Piattelli-Palmarini es la hipótesis de que FLN consistiría fundamentalmente en el mecanismo de la recursión⁴. Así, el mecanismo computacional de la recursión tomaría un conjunto finito de elementos y produciría a partir de él un conjunto potencialmente infinito de expresiones discretas.

En este marco conceptual, y apoyándose en el método comparativo de la biología, Hauser, Chomsky y Fitch proponen la siguiente hipótesis acerca de la evolución del lenguaje: la mayor parte de FLB, sino toda, estaría basada en mecanismos que compartimos con otras especies (esto es, los sistemas sensoriomotor y conceptual-intencional); en cambio, FLN, es decir el mecanismo computacional de la recursión, habría evolucionado recientemente y sería exclusivo de la especie humana. Hay que subrayar que los autores afirman que esta hipótesis anula el argumento del diseño que presentamos anteriormente y cuestiona, de ese modo, que FLN sea una adaptación (2002, p. 1573). Igual que Piattelli-Palmarini, consideran que por esta razón los aspectos nucleares de FLN no serían una adaptación producto de la selección natural, sino, más bien, *spandrels* o productos colaterales (*by-products*) de restricciones preexistentes (p. 1574). Advierten, sin embargo, que la cuestión no es si FLN tomada como un todo es adaptativa, pues al permitirnos comunicar una infinita variedad de pensamientos la recursión claramente lo es. La cuestión es si componentes particulares del funcionamiento de FLN son adaptaciones para el lenguaje, moldeadas específicamente por la selección natural o, si, por el contrario, FLN habría evolucionado por razones distintas que la comunicación (p. 1574). En ese sentido en las conclusiones de su trabajo señalan que FLN podría haber evolucionado por otros motivos que el lenguaje: «*to solve other computational problems such as navigation, number quantification, or social relationships*» (p. 1578)⁵.

En su respuesta a Hauser, Chomsky y Fitch (2002), Pinker y Jackendoff (2005) sostienen —retomando una de las ideas centrales de Pinker y Bloom (1990)— que la facultad del lenguaje, al igual que otros sistemas biológicos que muestran señales de un diseño adaptativo complejo, es un sistema de rasgos coadaptados que evolucionó

⁴ Definida como la capacidad de insertar iterativamente un elemento dentro de un elemento de la misma categoría, como cuando subordinamos una oración dentro de otra, por ejemplo: Jorge come arroz / Marcos cree que Jorge come arroz / Pablo piensa que Marcos cree que Jorge come arroz, etcétera (agradezco a Paola Céspedes por este ejemplo).

⁵ Nótese que en este último caso la recursión no sería ni un *by-product* ni un *spandrel*, sino una *exaptación* de un mecanismo que evolucionó en otros dominios al dominio del lenguaje.

mediante la selección natural para la comunicación de proposiciones complejas: «*The language faculty, like other biological systems showing signs of complex adaptive design (Dawkins 1996, Williams 1966), is a system of co-adapted traits that evolved by natural selection (Jackendoff, 1992, 1994, 2002; Pinker, 1994b, 2003; Pinker & Bloom, 1990). Specifically, the language faculty evolved in the human lineage for the communication of complex propositions*» (p. 204). Pinker y Jackendoff rechazan en ese sentido la «hipótesis de solo-recursión» (*recursion-only hypothesis*), que según Hauser, Chomsky y Fitch nulificaría el argumento del diseño y cuestionaría que FLN sea una adaptación. Con el propósito de cuestionar esta hipótesis, Pinker y Jackendoff llevan a cabo en la mayor parte de su artículo (2005, pp. 205-218) una extensa revisión de evidencia empírica, que en su opinión muestra que existen otros aspectos del lenguaje que a pesar de ser distintos de la recursión también deben formar parte de FLN, por ser inherentemente lingüísticos e inherentemente humanos, a partir de lo cual concluyen que «*the empirical case for the recursion-only hypothesis is extremely weak*» (p. 217). En su opinión, dicha hipótesis de Hauser, Chomsky y Fitch estaría motivada por la presunción de que en última instancia el programa minimalista de Chomsky será validado. Pinker y Jackendoff concluyen finalmente que el argumento de estos autores a favor de que el lenguaje no es una adaptación para la comunicación no es convincente (p. 231).

En su réplica a Pinker y Jackendoff (2005), Fitch, Hauser y Chomsky les reprochan no haber comprendido adecuadamente su propuesta y haber hecho borrosa la distinción central entre FLB y FLN (2005, p. 180). Seguidamente, conceden que FLB puede ser una adaptación que evolucionó como producto de la selección natural para diferentes fines, entre ellos la comunicación, y que en su artículo de 2002 han sostenido claramente que sí consideran que el lenguaje es el resultado de la evolución y que muestra signos de diseño adaptativo (2005, p. 185). Pero insisten en que a diferencia de FLB, FLN no es una adaptación (en todo caso para la comunicación) (p. 189). Observan igualmente que el hecho de que FLN esté compuesta por un único mecanismo simple (la recursión) nulifica el argumento del diseño en que se basan los adaptacionistas (p. 190). Presentan, además, evidencia empírica contraria a la tesis central de Pinker y Jackendoff de que hay otros mecanismos en FLN distintos de la recursión, y consideran que dichos aspectos del lenguaje forman parte de FLB, al no ser exclusivamente lingüísticos y exclusivamente humanos. Finalmente en su contrarréplica a Fitch, Hauser y Chomsky (2005), Jackendoff y Pinker sostienen que la caracterización de FLN, hecha por estos autores en sus artículos de 2002 y 2005, es problemática, al presuponer una dicotomización de las capacidades cognitivas entre las que son absolutamente únicas al lenguaje y a los seres humanos y aquellas que son idénticas a capacidades cognitivas no lingüísticas

o no humanas, dejando de lado aquellas capacidades cognitivas que han sido modificadas durante la evolución humana (2005). Señalan que Fitch, Hauser y Chomsky suelen aplicar la distinción entre FLN/FLB de manera absoluta, y de ese modo utilizan cualquier similitud entre una función lingüística y cualquier otra capacidad cognitiva para desterrar a dicha función de FLN y situarla en FLB (Jackendoff & Pinker, 2005, p. 214).

He presentado en líneas generales los principales argumentos intercambiados por Chomsky, Hauser y Fitch por un lado y Pinker y Jackendoff por el otro, en contra o a favor de una explicación adaptacionista de la evolución de la facultad del lenguaje. Como hemos podido ver, se trata de un debate muy complejo, y sin duda apasionante, para el cual ambos bandos aportan evidencia empírica de fuentes diversas y cuyo desenlace es difícil de predecir. Ambos lados reconocen además que ambas posturas constituyen en este momento tan solo hipótesis de investigación y que actualmente no disponemos de datos empíricos suficientes para decidirnos por alguna (y es posible que nunca dispongamos de ellos). En ese sentido mi propósito en este trabajo será solamente examinar críticamente ciertos aspectos de la posición de Chomsky frente a la explicación adaptacionista de la evolución del lenguaje. No creo que dicho examen crítico esté en condiciones de invalidar la posición de Chomsky ni mucho menos, pero tal vez sí de relativizarla.

3. EL ARGUMENTO CONTRA EL «ARGUMENTO DEL DISEÑO»

Un primer punto acerca del que quiero llamar la atención tiene que ver con una diferencia fundamental entre la argumentación esgrimida por Hauser, Chomsky y Fitch (2002) en contra de la explicación adaptacionista de la evolución del lenguaje, respecto a la planteada trece años antes por Piattelli-Palmarini (1989). Si bien Hauser, Chomsky y Fitch concuerdan con Piattelli-Palmarini en que la facultad del lenguaje no sería una adaptación que habría evolucionado mediante el moldeamiento gradual por parte de la selección natural, y explican su origen filogenético también recurriendo a mecanismos no adaptacionistas como *spandrels*, productos colaterales o exaptaciones, su propuesta difiere de la de Piattelli-Palmarini en relación a la naturaleza de lo que hay que explicar como producto de dicha evolución, es decir, al diseño de la facultad del lenguaje. Así, en la versión de Piattelli-Palmarini, lo que hay que explicar como producto de la evolución es una facultad del lenguaje concebida desde la perspectiva de la teoría de principios y parámetros desarrollada por Chomsky en los años ochenta, y a la que por tanto se le atribuye un diseño interno muy complejo y rico. En cambio, para Hauser, Chomsky y Fitch lo que hay que explicar como producto de la evolución es ahora la facultad del lenguaje en sentido estricto, FLN,

a la que se le atribuye un diseño muy simple o casi ningún diseño, ya que incluye únicamente al mecanismo de recursión y a sus interfaces⁶.

En este contexto cabe plantear la pregunta por los motivos que llevaron a un cambio tan profundo en la concepción de la facultad del lenguaje. Una respuesta plausible, que seguramente suscribirían Chomsky y los seguidores de sus ideas, es que dicho cambio ha sido exclusivamente el producto de años de investigación empírica, el que llevó al paso del paradigma de principios y parámetros al paradigma minimalista. Pero si tomamos en cuenta que entre ambas propuestas, la de Piattelli-Palmarini (1989) y la de Hauser, Chomsky y Fitch (2002), se encuentra la réplica de Pinker y Bloom (1990) a la primera, no es fácil apartar de la mente la impresión de que el cambio planteado en Hauser, Chomsky y Fitch (2002) podría también haber sido motivado en parte por la necesidad de responder a la defensa de la tesis adaptacionista propuesta por Pinker y Bloom en dicho trabajo, y en particular a la objeción que estos plantean sobre la base del «argumento del diseño». Como hemos visto, para estos autores el diseño rico y complejo de la facultad del lenguaje, que el propio Piattelli-Palmarini asumía en su artículo de 1989, constituye una evidencia de que dicha facultad es una adaptación producto de la selección natural. Esta conclusión se apoya, como indicamos, en el argumento del diseño planteado por Dawkins (1986): mecanismos no adaptativos como la exaptación pueden producir rasgos biológicos más simples, pero la selección natural es la única que puede producir estructuras con un diseño complejo, que son adaptaciones para alguna función. Por tanto si la facultad del lenguaje muestra un diseño complejo, lo más plausible sería asumir que es una adaptación producto de la selección natural antes que una exaptación. Notamos, pues, que la hipótesis de Hauser, Chomsky y Fitch de que FLN está compuesta por un único mecanismo simple, la recursión, permite, como lo señalan los mismos, nulificar este argumento del diseño: «*If FLN is indeed this restricted, this hypothesis has the interesting effect of nullifying the argument from design, and thus rendering the status of FLN as an adaptation open to question*» (2002, p. 1573). Vemos claramente por qué: un mecanismo simple como la recursión sí podría ser el producto de un mecanismo no adaptativo como la exaptación (ver también Pinker & Jackendoff, 2005, p. 219).

Sería sin duda excesivo atribuir a Hauser, Chomsky y Fitch la intención de plantear, de manera casi estipulativa, que FLN debe tener un diseño simple, esto es, estar formada por un mecanismo cognitivo simple (la recursión), para poder de este modo

⁶ Podemos conceder que en algún sentido la teoría de principios y parámetros también tenía por objeto de estudio a la facultad del lenguaje en sentido estricto o FLN. Pero en cualquier caso para dicha teoría, asumida por Piattelli-Palmarini, el diseño de FLN era rico y complejo, en tanto que para HCF es más bien muy simple.

explicar a dicho mecanismo cognitivo como el producto de un proceso biológico no adaptativo como la exaptación, el que no puede producir diseños complejos, eludiendo así la objeción planteada por el argumento del diseño esgrimido desde el bando adaptacionista. Pues en un tal escenario la hipótesis de estos autores se parecería más a una astuta estrategia jurídica que a una teoría científica surgida únicamente a partir del análisis neutral de los datos empíricos. Pero, aun si descartamos dicho escenario, tampoco se puede dejar de lado por completo la posibilidad de que el permitir eludir dicha objeción podría haber constituido una motivación adicional para favorecer el enfoque de un diseño simple de FLN. No hay que olvidar, por otro lado, que los propios Pinker y Jackendoff han sugerido —lo que sin duda ha contribuido a la aspereza de este debate— que no todas las propuestas de Hauser, Chomsky y Fitch deberían ser consideradas estrictamente como el resultado de una investigación científica neutral. Así, luego de reprochar al giro minimalista que con él «*most of the technical accomplishments of the preceding 25 years of research in the Chomskyan paradigm must be torn-down*» (2005, p. 220), Pinker y Jackendoff afirman en relación al trabajo empírico hecho en el marco de dicho programa que «*such work should not be taken as empirically vindicating Minimalist hypotheses about the empirical nature of language, but rather as carrying out a mandate to implement this vision of Chomsky's*» (2005, p. 222). En la misma línea encontramos el siguiente cuestionamiento de Lappin y otros —citado también por Pinker y Jackendoff (2005, 222)—:

What is altogether mysterious from a purely scientific point of view is the rapidity with which a substantial number of investigators, who had significant research commitments in the Government-Binding framework, have abandoned that framework and much of its conceptual inventory, virtually overnight. In its place they have adopted an approach which, as far as we can tell, is in no way superior with respect to either predictive capabilities or explanatory power (2000, p. 667).

4. ¿PUEDE LA EVOLUCIÓN PRODUCIR UN ÓRGANO PERFECTO?

Una segunda cuestión que deseo examinar, que está vinculada con la anterior, es que a pesar de reconocer que no contamos con evidencias empíricas suficientes para sostener la hipótesis de que FLN tiene un diseño simple y que está compuesta únicamente por el mecanismo de recursión, Hauser, Chomsky y Fitch optan decididamente por una explicación no adaptacionista de la evolución del lenguaje. Conceden así que eventualmente podría ocurrir que la investigación muestre que FLN sí tiene un diseño interno complejo, con lo que la explicación adaptacionista se vería confirmada: «*This hypothesis and the alternative selectionist account are both viable and can eventually be tested with comparative data*» (2002, p. 1574). Por ello me parece

legítimo plantear la siguiente pregunta: si tanto la explicación adaptacionista como la explicación no adaptacionista de la evolución del lenguaje son plausibles, ¿por qué Chomsky «apuesta» de modo tan enfático por la explicación no adaptacionista?

Contestar pregunta requiere que tomemos en cuenta que el proceso cognoscitivo que lleva al conocimiento científico, del mismo modo que la cognición humana en general, no es únicamente el resultado de un proceso *de abajo hacia arriba* (*bottom up*), que parte de los datos para llegar a ciertas hipótesis explicativas, sino que también se encuentra influido, en un proceso *de arriba hacia abajo* (*top down*), por ciertas expectativas, explícitas o implícitas, que los investigadores pueden tener acerca de los posibles resultados de su labor, las que también revelan la influencia de factores personales, sociales, culturales, institucionales, etcétera⁷. En ese sentido podemos afirmar que hay en Chomsky una clara preferencia por la explicación no adaptacionista y un claro rechazo por la explicación adaptacionista, que no es el resultado de la investigación científica misma, ya que, como él mismo reconoce, esta todavía no nos ha proporcionado los datos necesarios para decidirnos por una o por otra. ¿Qué motiva, entonces, a Chomsky a comprometerse en una crítica tan enfática de la explicación adaptacionista de la evolución del lenguaje? La motivación para su rechazo de la explicación adaptacionista se deriva, en mi opinión, de un supuesto básico de la aproximación chomskiana al lenguaje, que es que las estructuras lingüísticas no están diseñadas para cumplir ninguna función. Esta idea es expresada claramente por Piattelli-Palmarini cuando subraya la ‘gratuidad’ de las estructuras cognitivas que subyacen al lenguaje: «*The study of language has, in fact, disclosed many instances of specificity and gratuity in the design of all natural human languages*» (1989, p. 22).

Desde mi punto de vista, Chomsky está tratando de escapar a la paradoja que surge de postular por un lado estructuras cognitivas lingüísticas gratuitas, cuyo diseño no cumple ninguna función, y postular al mismo tiempo que una parte fundamental de tales estructuras cognitivas es innata y por tanto parte de la dotación biológica de la especie. Lo que origina la paradoja es la teoría de la selección natural de Darwin: si un rasgo es parte de la dotación biológica de la especie, entonces debe ser un producto de la evolución mediante la selección natural, por lo cual debe proporcionar ventajas adaptativas a los organismos que lo poseen. Pero si dicho rasgo

⁷ Esta influencia de expectativas surgidas de factores emocionales en procesos cognoscitivos racionales ha sido elaborada en el marco de la teorización psicoanalítica de la mente, en particular en base a la clásica noción freudiana de la superación del principio del placer (que nos hace representarnos solamente las cosas que nos gustan) por el principio de la realidad (que nos exige representarnos las cosas como son, aunque no nos gusten), la que según Freud (1900, pp. 570 y ss.; 1911, pp. 18 y ss.) nunca llega a ocurrir completamente.

proporciona ventajas adaptativas, entonces su diseño no puede ser gratuito, sino que debe responder al cumplimiento de alguna función adaptativa. Sostengo en ese sentido que los argumentos planteados por Piattelli-Palmarini y Hauser, Chomsky y Fitch constituyen estrategias para resolver dicha paradoja, mediante el recurso de invocar mecanismos biológicos no adaptativos que pueden explicar la evolución de estructuras cognitivas gratuitas. Esta idea de la gratuidad de las estructuras lingüísticas no es, definitivamente, el resultado del trabajo empírico llevado a cabo en el marco de la gramática generativa: es más bien claramente un postulado que subyace a dicho programa de investigación, que se expresa en otro antiguo supuesto del enfoque chomskiano sobre el lenguaje. Me refiero al supuesto de la autonomía de la sintaxis, de acuerdo al cual el estudio de la sintaxis debe ser exclusivamente formal y no debe incorporar consideraciones semánticas o pragmáticas. Dicho postulado tampoco es el resultado de la investigación empírica, pues se encuentra presente al inicio mismo de la teoría generativista, en *Syntactic Structures*, en el que Chomsky afirma ya categóricamente que la gramática es autónoma e independiente del significado, y que la base de la teoría gramatical debe ser puramente formal. En esta obra temprana leemos ya, por ejemplo, que «*grammar is autonomous and independent of meaning*» (1957, p. 17) y que «*only a purely formal basis can provide a firm and productive foundation for the construction of grammatical theory*» (p. 102).

Otra de las propuestas centrales de mi trabajo es, por tanto, que una motivación muy importante que lleva a Chomsky a luchar con todas sus fuerzas contra la explicación adaptacionista neodarwiniana de la evolución del lenguaje es que aceptar dicha explicación lo obligaría a renunciar a una de sus convicciones más enraizadas, que es la de la gratuidad de las estructuras lingüísticas, es decir, la idea de que las estructuras lingüísticas son el producto de un mecanismo computacional puramente formal, indiferente a los contenidos conceptuales que tales estructuras pueden transmitir o a las funciones comunicativas que pueden cumplir. Pues si la explicación adaptacionista fuese correcta, entonces las estructuras lingüísticas presentarían el diseño que tienen porque este sería el más óptimo para codificar tales contenidos conceptuales o para cumplir tales funciones comunicativas. En otras palabras: dicho diseño ya no sería gratuito y no tendría sentido pretender explicarlo en términos puramente formales.

Es interesante en este contexto hacer referencia a otro trabajo de Chomsky, publicado solo dos años antes de Hauser, Chomsky y Fitch, en el cuál plantea una pregunta curiosa, a saber: «*How close does language come to what some super-engineer would construct, given the conditions that the language faculty must satisfy*» (Chomsky, 2000, p. 9). En dicho artículo el autor hace una afirmación central para entender su concepción acerca del lenguaje: que los lenguajes artificiales, es decir, los lenguajes

creados por el hombre con propósitos especiales, como los usados por la inteligencia artificial, deliberadamente evitan incorporar propiedades como los sistemas inflexionales o la propiedad de desplazamiento, las que serían propiedades especiales del lenguaje humano, producto de las condiciones de legibilidad impuestas por los sistemas externos con los que interactúa, que son el sistema conceptual y el sistema motor (2000, p. 12). Al leer este pasaje, nos damos cuenta de que para Chomsky estas dos «imperfecciones» serían externas a la facultad del lenguaje, impuestas a ella por su necesidad de interactuar con otros sistemas, de modo que en su núcleo más íntimo, por decirlo así, el diseño del lenguaje sería para Chomsky similar al de los lenguajes artificiales de la inteligencia artificial, que son lenguajes concebidos con un propósito, en otras palabras, que dicho diseño, que estaría libre de imperfecciones, se acercaría a lo que algún «superingeniero» diseñaría. En otras palabras, que dicho diseño sería perfecto. Encontramos una idea similar en un trabajo reciente de Chomsky, en el que plantea lo que denomina la *tesis minimalista fuerte*, según la cual «*language keeps to the simplest recursive operation, Merge, and is perfectly designed to satisfy interface conditions*» (2010, p. 52).

El invocar una noción como la de «perfección» cuando se está tratando de dar cuenta de una estructura que supuestamente pertenece al ámbito de lo biológico, y por tanto a la naturaleza, no puede dejar de llamar la atención. Vale la pena en ese sentido recordar que en su recordada serie de televisión *Cosmos* Carl Sagan nos cuenta cómo el astrónomo alemán del siglo XVII Johannes Kepler imaginaba que el movimiento de los planetas debía regirse por las leyes pitagóricas de la armonía. Asumiendo el modelo copernicano, pero incorporándolo dentro de una cosmología a la vez platónica y religiosa, Kepler había intentado infructuosamente durante años demostrar que las distancias de los planetas en relación al sol estaban determinadas por esferas en el interior de poliedros perfectos, las que se encontraban anidadas una dentro de la otra de modo sucesivo. Pero su modelo no coincidía con las observaciones de Tycho Brahe, el más importante astrónomo de la época, y Kepler tuvo que aceptar que el movimiento de los planetas no podía ser explicado por su modelo de poliedros perfectos y armonía de esferas y pudo reconocer finalmente que las órbitas de los planetas no eran, como él imaginaba, círculos, que eran figuras geométricas simples y perfectas, realizando así el descubrimiento empírico de que estas son en realidad elípticas. Teniendo esto en mente llama la atención que Chomsky, como un Kepler contemporáneo todavía deslumbrado por la ilusión de una cosmovisión platónica, postule la existencia de una entidad perfecta en el mundo natural, y se plantee incluso la pregunta de cuánto se acercaría a lo que un «superingeniero» podría diseñar. ¿Hay lugar para semejante órgano perfecto dentro de la teoría darwiniana de la evolución? Como señalan Pinker y Bloom,

«*The idea that natural selection aspires toward perfection has long been discredited in evolutionary theory (Williams, 1966)*» (1990, p. 717). De este modo, la facultad del lenguaje de Chomsky sería un objeto sui generis del mundo natural. No sería aprendida, sino innata, formaría parte de la dotación biológica de una especie, el homo sapiens. Pero a diferencia de la mayoría de los rasgos biológicos, no sería el producto de la selección natural. No tendría ninguna función, no proporcionaría a esta especie ninguna ventaja adaptativa, sería solo un producto colateral de la evolución. Su diseño no sería el resultado de la selección natural, pero sería perfecto y comparable al de un lenguaje artificial diseñado por un superingeniero. En ese sentido la explicación adaptacionista que Chomsky rechaza, de que el lenguaje habría evolucionado para la comunicación de saber y de intenciones, tiene la ventaja de que no requiere postular un objeto natural único y, como escriben Pinker y Jackendoff (2005, p. 231) no requiere comprometerse con supuestos como que el lenguaje es no-redundante, perfecto, inadecuado para la comunicación, o diseñado antes para la belleza que para el uso.

5. ¿LA CREATIVIDAD LINGÜÍSTICA ES SOLO PRODUCTO DE LA RECURSIVIDAD?

La tercera cuestión que quiero tratar aquí tiene que ver con la llamada *creatividad lingüística*. Desde mi punto de vista una de las mayores contribuciones de Chomsky a la comprensión del lenguaje y de la cognición humana en general ha sido el poner en relieve el fenómeno de la *creatividad lingüística*, como aquel dato central que debe ser explicado por los investigadores del lenguaje. Tomemos la oración «El atolon-drado clarinetista encontró la arrugada partitura debajo del refrigerador». Es muy poco probable que el lector de este trabajo haya tenido anteriormente contacto con ella. Sin embargo al mismo tiempo es muy probable, si es un hablante competente del español, que no haya tenido ninguna dificultad en comprenderla. Es esto precisamente lo que se entiende por *creatividad lingüística*: la posibilidad de comprender y producir infinitas oraciones en nuestra lengua materna, aunque no las hayamos encontrado antes. ¿Cómo puede explicarse esta capacidad? Chomsky nos hizo ver que el conocimiento del lenguaje no puede consistir en almacenar todas las oraciones de la lengua en nuestra memoria, pues esto es imposible, ya que son potencialmente infinitas. Como señalamos anteriormente, Chomsky explica el *gran truco del lenguaje*, que hace posible la *creatividad lingüística*, asumiendo que tales expresiones infinitas son el producto de medios finitos, los que consisten, en rasgos generales, en un repertorio de elementos léxicos por un lado y la posibilidad de combinarlos de acuerdo con ciertas restricciones o reglas por el otro. Tales medios finitos constituyen el conocimiento que tiene el hablante de su lengua, el que estaría representado en su mente.

No obstante, hay que llamar la atención acerca del hecho de que en las propuestas más recientes de Chomsky se tiende a identificar la creatividad lingüística con el mecanismo de la recursividad, el que da lugar a la llamada propiedad de la *infinitud discreta* (*discrete infinity*). Así, Chomsky sostiene que el lenguaje humano está basado en una propiedad elemental biológicamente aislada, que es la propiedad de la infinitud discreta, que también muestran los números naturales, en los que siempre se puede sumar uno (2000, p. 3). En la misma línea Hauser, Chomsky y Fitch sostienen que el núcleo del lenguaje es su capacidad de poder expresivo ilimitado, capturada por la noción de la infinitud discreta (2002, p. 1576), identificando a continuación esta noción a la de la creatividad lingüística, pues la definen como la capacidad, exclusivamente humana, de recombinar unidades significativas en una variedad ilimitada de estructuras más grandes, que difieren sistemáticamente en su significado (p. 1576).

Desde mi punto de vista esta identificación entre creatividad lingüística, infinitud discreta y recursividad es problemática. Pues la infinitud discreta (como puede apreciarse en los propios experimentos descritos en Hauser, Chomsky y Fitch) consiste fundamentalmente en la posibilidad de agregar una unidad más a una secuencia escalar de unidades, como los números naturales. En ese sentido, estos autores presentan un experimento que muestra la diferencia fundamental en la que los niños adquieren el número, a diferencia de los chimpancés: mientras estos últimos son capaces únicamente de correlacionar dígitos del uno al cinco con las cantidades correspondientes, los niños pueden hacer algo que los chimpancés no pueden, que es comprender la secuencia de tales dígitos y el principio que las gobierna, conocido como función de sucesión, que consiste en poder agregar una unidad más (p. 1577). Pero si bien todo esto es muy interesante, pienso que tiene poco que hacer con la creatividad lingüística, pues cuando un hablante produce una oración que nunca ha escuchado, está haciendo algo muy distinto que «agregar uno más»: está haciendo una combinación novedosa de símbolos, esto es, de signos lingüísticos (que además no tienen entre sí una relación escalar, como los números). Considero por ello que la noción de *infinitud discreta* a la que apela Chomsky distorsiona finalmente el concepto de creatividad lingüística que él mismo propuso en sus obras tempranas. Pues la creatividad lingüística no consiste, en mi opinión, en poder comprender y producir *infinitas* oraciones, sino más bien en poder comprender y producir oraciones *novedosas* de manera *ilimitada*, lo que no es lo mismo. O dicho de otro modo, en poder comprender y producir ilimitadamente novedosas combinaciones de signos lingüísticos (frases y oraciones). Y la noción de recursión tampoco es capaz de dar cuenta de esta capacidad: no siempre que creamos una combinación novedosa de signos lingüísticos estamos introduciendo un elemento dentro de otro elemento

de la misma categoría. De hecho, la oración citada anteriormente es un ejemplo de creatividad lingüística, pero su producción o comprensión no involucran en ningún momento la recursividad.

Por tanto, la creatividad lingüística es fundamentalmente el resultado del mecanismo combinatorio, que puede producir secuencias novedosas a partir de elementos léxicos, en base a restricciones a las secuencias posibles de tales elementos (restricciones que se pueden concebir como reglas, como principios o como propiedades directamente codificadas en los elementos léxicos). Que tales secuencias sean infinitas es un producto secundario del carácter ilimitado de las combinaciones que el mecanismo puede producir. No considero además que la infinitud discreta, que sí es claramente una propiedad del sistema numérico, sea una propiedad central del lenguaje, pues a diferencia de, por ejemplo, los números naturales, las expresiones lingüísticas no constituyen una escala en que la posición relativa de cada una en la secuencia dependa de ser una unidad mayor que la anterior. La recursividad, finalmente, juega solo un papel secundario en la generación de secuencias novedosas de símbolos (frases y oraciones), pues la mayoría de ellas no involucran operaciones recursivas. La aplicación reiterada de una operación recursiva (como por ejemplo la inclusión de una oración relativa dentro de otra) juega un papel muy marginal en el uso cotidiano del lenguaje (salvo, tal vez, en alguna canción infantil).

Chomsky sí propone, sin embargo, un mecanismo que es capaz de explicar la generación de combinaciones novedosas de elementos léxicos: es la operación *Merge*. En una oportunidad la define como la operación irreductible que toma dos objetos ya previamente formados y anexiona uno al otro, formando un objeto mayor que tiene precisamente las propiedades del blanco de la anexión (2000, p. 13). Considero por ello que sería más adecuado, si se quiere adherir a la noción de FLN propuesta por Hauser, Chomsky y Fitch, asumir que esta consiste no tanto en la operación de la recursividad, sino en la operación *Merge*, que es capaz de producir secundariamente recursividad e infinitud discreta. Por ello es curioso que en Hauser, Chomsky y Fitch (2002) y en Fitch, Hauser y Chomsky (2005) la noción de *Merge* prácticamente no sea mencionada y quede opacada por la de recursividad, que como acabamos de ver, no puede realmente explicar el mecanismo combinatorio del lenguaje que subyace a la creatividad lingüística. En cambio en un trabajo más reciente, Chomsky señala explícitamente que FLN está constituida por *Merge*, a la que describe como una operación recursiva (2010). Así, en dicho texto Chomsky afirma que cualquier sistema generativo incorpora una operación *Merge* que toma estructuras preexistentes X e Y y las combina en una nueva estructura Z. Aplicada de manera ilimitada a un lexicón de átomos léxico-conceptuales, produce una infinitud discreta de expresiones estructuradas. Observa, además, que en el caso más simple, *Merge* es la única operación

recursiva dentro de la Gramática Universal (GU), parte del componente genético de la facultad del lenguaje y por tanto un producto de la evolución de este órgano cognitivo (2010, p. 52).

No obstante, a pesar de que las presentaciones de Chomsky (2000) y Chomsky (2010) me parecen en este aspecto más adecuadas que las de Hauser, Chomsky y Fitch (2002) y Fitch, Hauser y Chomsky (2005), al poner a *Merge* en primer plano, considero que esa operación no es capaz por sí sola de dar cuenta por completo de la creatividad lingüística. La razón es que en mi opinión un aspecto fundamental de la creatividad lingüística es lo que lingüistas estructuralistas como Martinet (1980) llamaron la *doble articulación* del lenguaje. Se puede distinguir así dos niveles en la combinación de las unidades lingüísticas. En un nivel inicial se combina unidades sin significado (fonemas) para formar unidades con significado (morfemas). En un siguiente nivel, tales unidades con significado son combinadas para formar unidades que también tienen significado, y que son cada vez mayores⁸. Por ejemplo, en el nivel inicial las seis unidades fonológicas *c, o, r, t, é, s*, que carecen de significado, pueden combinarse para formar el morfema *cortés*; en el siguiente nivel este morfema puede combinarse con el morfema *des-* para formar la unidad léxica *descortés*, la que puede a su vez combinarse con las unidades léxicas *el* y *gato* para formar la frase nominal *el descortés gato*, la que puede finalmente combinarse con la frase verbal *durante tres horas* para formar la oración *El descortés gato maulló durante tres horas*⁹. Una propiedad muy importante de las combinaciones de este nivel es que son interpretables semánticamente. Por un lado, los morfemas reciben directamente una interpretación semántica mediante reglas que los vinculan a determinados contenidos conceptuales, y por otro lado sus combinaciones en unidades mayores (palabras, frases, oraciones) son interpretadas *composicionalmente*: el significado de tales unidades es función del significado individual de sus partes y de la forma en que están combinadas (Chierchia & McConnell-Ginet, 2000).

La cuestión es que si bien *Merge* es capaz de dar cuenta de la combinación de elementos con significado para formar unidades mayores, que pueden luego interpretarse composicionalmente, tal operación no puede dar cuenta de la combinación de las unidades sin significado o fonemas para formar unidades con significado o morfemas. Pues si *Merge*, como vimos, toma dos objetos anteriormente formados y anexiona uno al otro, formando un objeto mayor que tiene las propiedades del blanco

⁸ Martinet llama al nivel de los fonemas «segundo nivel de articulación» y al nivel de las unidades con significado «primer nivel de articulación». Pero para facilitar la exposición he preferido llamar nivel *inicial* al de la combinación de los fonemas y *siguiente* nivel al de la combinación de unidades con significado.

⁹ Agradezco a Jorge Pérez por este ejemplo.

de la anexión, la combinación de fonemas para formar un morfema no puede ser su producto. En consecuencia, si *Merge* es lo único exclusivamente lingüístico, la caracterización del lenguaje de Chomsky deja de lado al nivel de articulación correspondiente a la combinación de los fonemas, relegándolo a la interface de FLN con el sistema sensorio-motor. Es curioso, sin embargo, que en otro trabajo de este mismo periodo Chomsky sí hace referencia a un mecanismo computacional que evoca la combinación de fonemas o unidades sin significado. Aquí sostiene que el sistema computacional que genera expresiones lingüísticas tiene dos operaciones básicas, una que ensambla rasgos para formar ítems léxicos y otra que forma objetos sintácticos más amplios a partir de aquellos que ya han sido construidos, comenzando por los ítems léxicos (2000, p. 10). Si bien la segunda claramente se refiere a *Merge*, la primera debe incluir necesariamente rasgos fonológicos. De hecho en ese mismo pasaje señala que la primera operación puede concebirse como un lexicón o una lista de ítems léxicos, que son asociaciones arbitrarias de sonido y significado, así como elecciones particulares de propiedades flexionales.

Pero frente a ello Hauser, Chomsky y Fitch sostienen enfáticamente que la fonología no forma parte de FLN. Cuestionando la difundida idea de Lieberman (1984) de que «el habla es especial» (*«speech is special»*), de acuerdo a la cual esta se basaría en mecanismos exclusivamente humanos, adaptados para la percepción y la producción del habla, Hauser, Chomsky y Fitch sostienen que habría más bien una continuidad entre humanos y animales respecto al habla, que sería mucho mayor de lo que se había asumido hasta ahora (2002, p. 1574; Fitch, Hauser & Chomsky, 2005, p. 192). Pinker y Jackendoff, por el contrario, subrayan en su respuesta a Hauser, Chomsky y Fitch que las características globales y mayores de la fonología son específicas al lenguaje, únicamente humanas, infinitamente discretas, y no recursivas, por lo que esta representaría un contraejemplo muy significativo para la hipótesis de *solo-recursión* (2005, p. 212). En su réplica a Pinker y Jackendoff, Fitch, Hauser y Chomsky insisten en que la fonología no solo no representa un contraejemplo significativo para la hipótesis de solo-recursión, sino que es irrelevante para dicha hipótesis, y atribuyen la mayor parte de la fonología a FLB: «*Given our present knowledge, much of phonology is likely part of FLB, not FLN, either because phonological mechanisms are shared with other cognitive domains (notably music and dance), or because the relevant phenomena appear in other species, particularly bird and whale “song”*» (2005, p. 200).

En su propuesta acerca de los posibles estadios de la evolución del lenguaje humano Jackendoff (1999) ya había postulado un estadio en el que tendría lugar el desarrollo de un sistema fonológico combinatorio, el que permitiría la creación de una clase abierta e ilimitada de símbolos, anterior y distinto al desarrollo de la sintaxis,

en particular de la estructura jerárquica de frase, la que correspondería a un estadio posterior. Conuerdo con esta posición. Sin las combinaciones de fonemas el resto del edificio lingüístico sería imposible. ¿Qué se estaría combinando en los otros niveles? No podrían ser conceptos, pues las secuencias de fonemas a las que llamamos morfemas y palabras constituyen precisamente lo que nos permite lexicalizar conceptos, y por tanto tener acceso a los mismos, activándolos en el procesamiento cognitivo y combinándolos composicionalmente en unidades conceptuales mayores. Son también aquellas secuencias de fonemas que conocemos como morfemas gramaticales las que nos permiten expresar lingüísticamente las relaciones sintácticas. El nivel de articulación constituido por la combinación de los fonemas para dar lugar a morfemas o unidades lingüísticas con significado es, por tanto, un aspecto fundamental de la creatividad lingüística. La doble articulación es, por tanto, un aspecto clave del *gran truco* del lenguaje.

6. ¿COMUNICACIÓN O PENSAMIENTO?

Una controversia vinculada al debate en torno al adaptacionismo tiene que ver con la función que se atribuye el lenguaje. Hemos visto en este sentido que mientras los autores que defienden la postura adaptacionista, como Pinker, Bloom o Jackendoff, consideran que el lenguaje evolucionó como una adaptación para la comunicación, y en particular para la comunicación de proposiciones, los críticos del adaptacionismo, como Chomsky, Hauser y Fitch, no creen que el lenguaje (en particular la facultad del lenguaje en sentido estrecho o FLN) haya evolucionado como una adaptación para ninguna función en particular, y mucho menos para la comunicación.

Así, en su reseña crítica de Hauser, Chomsky y Fitch (2002), Pinker y Jackendoff (2005) atribuyen a Chomsky la tesis de que el lenguaje estaría pobremente diseñado para la comunicación. Señalan que para Chomsky el lenguaje no es un sistema para la comunicación, sino un sistema para expresar el pensamiento, que puede ser usado para la comunicación (que no sería la función del lenguaje) (Chomsky, 2000, citado por Pinker & Jackendoff, 2005, p. 223). En contra de este punto de vista Pinker y Jackendoff afirman que el diseño del lenguaje (una correspondencia entre significado y sonido) es precisamente lo que se esperaría de un sistema que ha evolucionado para la comunicación de proposiciones (2005, p. 224). Consideran además que la única manera de dar sentido al hecho de que los seres humanos estén equipados con un dispositivo para hacer corresponder significado y sonido producido vocalmente es que este dispositivo permite a una persona lograr que el significado llegue a la cabeza de una segunda persona al hacer un sonido con su tracto vocal (p. 225). En su réplica, Fitch, Hauser y Chomsky objetan a Pinker y Jackendoff el haber atribuido

a Chomsky la tesis señalada, esto es, que el lenguaje no es «para» la comunicación porque es usado como habla interiorizada. Consideran que esta atribución de Pinker y Jackendoff a Chomsky es incorrecta y producto de una mala interpretación. En este contexto señalan, sin embargo, que hay biólogos que sí han planteado que las necesidades comunicativas no habrían constituido una presión selectiva significativa para producir un sistema como el lenguaje (Luria, 1974 o Jacob, 1982), el que estaría íntimamente relacionado con el desarrollo del pensamiento abstracto, al permitir infinitas combinaciones de símbolos. Fitch, Hauser y Chomsky conceden por otro lado que FLB pueda servir, entre otras funciones, para la comunicación y el pensamiento privado (2005, p. 5), pero agregan luego que si bien aceptan que FLB es una adaptación, insisten en su hipótesis de que FLN no es una adaptación para la comunicación (p. 8).

Más recientemente, Chomsky ha planteado una asimetría en la relación entre el lenguaje-I (lenguaje interno, es decir FLN) y los dos sistemas de interface, que son el sistema conceptual-intencional y el sistema sensorio-motriz, sosteniendo que su relación primaria es con la interface semántica, es decir, con los sistemas del pensamiento (2010, p. 55). Hace referencia así a la idea del biólogo François Jacob (1982) de que la cualidad que hace único al lenguaje no es su rol para la comunicación, sino para la simbolización y evocación de imágenes cognitivas. Chomsky añade que la generación de expresiones para satisfacer a la interface conceptual-intencional generaría un «lenguaje del pensamiento», disponible para uso interno, y que este sería el estadio más temprano del lenguaje (2010, p. 55). Más adelante en este mismo artículo Chomsky nos ofrece la siguiente especulación acerca de la evolución del lenguaje: «*Within some small group from which we are all descended, a rewiring of the brain took place in some individual, call him Prometheus, yielding the operation of unbounded Merge, applying to concepts with intricate (and little understood) properties*» (p. 59).

Siguiendo con este relato, Chomsky señala que el lenguaje habría dotado a *Prometheus* con un rango infinito de expresiones estructuradas, con lo cual poseía muchas ventajas frente a sus congéneres, como la capacidad para el pensamiento complejo, el planeamiento, la interpretación, etcétera. Esta capacidad sería transmitida a su descendencia, la que predominaría. En este punto, observa el autor, «*there would be an advantage to externalization, so the capacity might come to be linked as a secondary process to the SM system for externalization and interaction, including communication*» (p. 59). Vemos pues que si bien en textos anteriores como Fitch, Hauser y Chomsky no hay una toma de posición tan clara respecto a esta controversia, en este artículo Chomsky sí plantea de modo más claro que la primera función del lenguaje habría sido el pensamiento (si bien tampoco habría evolucionado para ella), y que recién ulteriormente su externalización mediante el sistema sensorio-motriz, esto es,

mediante el habla, es decir la comunicación, podría haber sido ventajosa. Por tanto la comunicación sería una función secundaria en relación al pensamiento.

Luego de presentar las dos posiciones respecto a si la función primaria del lenguaje es el pensamiento (o como a veces se dice, la representación del mundo) o la comunicación, quisiera señalar que en mi opinión tal controversia carece de sentido, por el hecho de que pensamiento (o representación del mundo) y comunicación se encuentran vinculados de manera indisoluble. Cada vez que en la intimidad de nuestro mundo subjetivo formulamos un pensamiento lingüísticamente, estamos empleando palabras y categorías prestadas, que otros han usado y usarán en innumerables ocasiones para formular sus propios pensamientos. Más aún, cualquier pensamiento que podamos formular en palabras es de por sí un pensamiento que puede ser potencialmente compartido con otros. Vimos anteriormente que Chomsky (1968) se esforzó por vincular su concepción del lenguaje, entre otros, con la de Wilhelm von Humboldt. Pues bien, este autor escribe, en el mismo texto al que Chomsky hace referencia en aquella obra, que «el pensamiento [idea] convertido en lenguaje ya no pertenece exclusivamente a un sujeto» (1988[1836], p. 54; la traducción es mía)¹⁰. Al pensar en palabras ya estamos pensando con otros, de hecho formulamos verbalmente nuestros pensamientos dirigiéndonos a nosotros mismos como si fuéramos un otro. En el mismo texto de Humboldt leemos en esta misma línea que el lenguaje solo se desarrolla socialmente y que el ser humano solo puede comprenderse a sí mismo en la medida en que ha probado en otros que sus palabras son comprensibles¹¹. Además, como señalamos más arriba, accedemos a los conceptos, que constituyen la materia prima de nuestros pensamientos, mediante secuencias de fonemas que constituyen por decirlo así las «direcciones» de tales conceptos en nuestra memoria. ¿Cómo podríamos poseer conceptos en nuestra memoria o manipularlos mentalmente sin poseer las secuencias de fonemas que nos permiten codificarlos? Por tanto para el pensamiento el sistema sensorio-motriz (para usar la terminología de Chomsky) tendría que estar presente desde un inicio, al lado del sistema-conceptual intencional.

Por otro lado nuestros conceptos provienen de la comunidad a la que pertenecemos y en la que nos socializamos. Al crecer, el niño aprende que su comunidad distingue «tipos» que puede encontrar en su experiencia, tanto en el mundo físico como en el social (e incluso en su mundo mental), tales como *silla*, *piedra*, *amigo*,

¹⁰ En el texto original: «*die in Sprache verwandelte Vorstellung nicht mehr ausschließend Einem Subjekt angehört*».

¹¹ «*In der Erscheinung entwickelt sich jedoch die Sprache nur gesellschaftlich, und der Mensch versteht sich selbst nur, indem er die Verstehbarkeit seiner Worte an Andren versuchend geprüft hat*» (Humboldt, 1988 [1836], p. 54).

juego, dormir, saltar, triste, verde, etcétera. Estos «tipos» corresponden a lo que la ciencia cognitiva denomina *categorías*. Propongo así definir un *concepto* como aquella información, almacenada en la memoria semántica, que empleamos para para clasificar a un ítem particular en una categoría. Los significados de las palabras o significados léxicos son, en ese sentido, conceptos (confróntese Herrera, 2002). La adquisición del léxico significa para el niño la apropiación de tales conceptos o de las categorías disponibles en su comunidad, es decir, en su cultura. Si entendemos por pensamiento la manipulación mental de tales conceptos, ¿cómo podría entonces existir un pensamiento antes de la comunicación? Podría argumentarse que las relaciones que establecemos entre los conceptos cuando pensamos no provienen de la cultura en la que crecemos, sino que se derivan de una estructura conceptual innata, que evolucionó en nuestra especie. Eso puede ser cierto, pero no vale para los conceptos mismos que conectamos mediante tales relaciones. Así, si me represento mentalmente, mediante el lenguaje, la proposición o el pensamiento de que «Mi amigo está en la biblioteca», se puede sostener eventualmente que la relación *estar en* pueda no ser cultural, o que las nociones de evento y participantes implícitas en dicha proposición puedan ser innatas, pero sin duda los conceptos de *amigo* o de *biblioteca*, que son conectados mediante dicha relación, no pueden serlo. No hay forma de que yo posea dichos conceptos, y por tanto de que pueda formular dicho pensamiento para mí mismo, antes de interactuar comunicativamente con otros en mi cultura. A menos, claro, que se asuma una posición platonista extrema que defienda la idea de que todos los conceptos que pudiésemos llegar a tener en cualquier lengua, en cualquier cultura y en cualquier época (aun las que nunca han llegado a existir) fuesen innatos y estuviesen guardados en algún lado esperando ser «descubiertos». Extrañamente Chomsky parece haber sostenido algo así cuando en sus *Conferencias de Managua* describe el proceso de aprendizaje de las palabras como el encontrar qué etiquetas son usadas para conceptos preexistentes, una conclusión que en su opinión parecería ser esencialmente correcta (1988, p. 134).

Si por otro lado se quiere sostener que la función primaria del lenguaje es la representación cognitiva del mundo (que es otra manera de definir el término *pensamiento*) y no la comunicación, habría que señalar que la representación del mundo no es una actividad individual y subjetiva, sino que es común e intersubjetiva, al emplear categorías compartidas intersubjetivamente y al construirse en la interacción dialógica. Así, como nos muestra el llamado *paradigma interpretativo* en el estudio de las interacciones sociales, que se alimenta de fuentes diversas como la sociología fenomenológica de Alfred Schütz, el interaccionismo simbólico de Herbert Mead o la etnometodología de Harold Garfinkel (Treibel, 1995), aquello que llamamos el mundo o la realidad, en particular la realidad social en la que vivimos cotidianamente,

no preexiste a las interacciones sociales, sino que es construida por los participantes en dichas interacciones mismas, gracias a las interpretaciones subjetivas que hacen cada uno de los participantes acerca de la situación en la que están involucrados, pero en particular a partir de la ratificación que ambos hacen de tales interpretaciones en la interacción dialógica, mediante la cual tales interpretaciones adquieren el estatus de una realidad intersubjetivamente sancionada. Las representaciones del mundo que el lenguaje hace posibles no son pues nunca individuales, son compartidas e intersubjetivas. En conclusión, representación (pensamiento) y comunicación son procesos que se implican mutuamente.

Para cerrar esta discusión quiero llamar la atención acerca de una similitud entre la especulación chomskiana de *Prometheus* y el origen del lenguaje por un lado, y la posición de las teorías políticas liberales de la modernidad acerca de la relación entre los individuos y la sociedad por el otro. De acuerdo con estas últimas, representadas por ejemplo por Hobbes y Kant, en un primer momento existirían individuos atomizados y aislados, los que recién en un segundo momento establecerían relaciones sociales entre sí. Tales teorías fueron cuestionadas, por ejemplo, en los escritos del joven Hegel, quien desde una perspectiva intersubjetivista rechazó el supuesto de la existencia previa de individuos atomizados y aislados y sostuvo por el contrario que los seres humanos se encuentran *desde siempre* inmersos en relaciones sociales que los definen¹². En la especulación de Chomsky encontramos, igual que en las teorías políticas liberales modernas, la noción de individuos que primero usan el lenguaje para pensar de manera aislada, y que recién posteriormente encuentran ventajoso externalizar dicho pensamiento, comunicándose con los demás. Si nos situamos, por el contrario, en una perspectiva intersubjetivista, tendremos que rechazar la especulación acerca de *Prometheus*, pues consideraremos que desde un inicio estamos en relación con otros, aun cuando cada uno esté pensando individualmente «dentro de su cabeza», ya que al pensar usando el lenguaje siempre estamos pensando «con otros».

7. COMENTARIOS FINALES

Como indiqué más arriba, mi propósito en este artículo ha sido examinar críticamente algunos aspectos de la posición de Chomsky frente a la explicación adaptacionista de la evolución del lenguaje. Señalé igualmente que el objetivo de dicho examen no era el de invalidar la posición de Chomsky, pero sí de relativizarla. Los cuestionamientos planteados aquí no me impiden reconocer que la hipótesis planteada por Hauser, Chomsky y Fitch (2002) puede, en última instancia, ser correcta. Como hemos visto,

¹² En alemán «*immer schon*». Esa idea del joven Hegel ha sido retomada en la filosofía social contemporánea por Axel Honneth (ver Deranthy, 2009 y Escalante, 2012).

los propios autores aceptan que tiene el estatus de una hipótesis empírica que debe ser contrastada con los datos de la investigación. Sin duda lo mismo vale para la hipótesis adaptacionista de Pinker y Jackendoff (2005). En ese sentido me parece importante conceder que hay una crítica muy importante que Hauser, Chomsky y Fitch han hecho a los intentos adaptacionistas de imaginar escenarios acerca de cómo podría haber evolucionado la facultad del lenguaje mediante la selección natural, cuando indican que ven pocas razones para creer que FLN puede ser descompuesta en rasgos independientes interactuantes, cada uno con su propia historia natural independiente, o que cada uno de estos rasgos podría haber sido moldeado por la selección natural, teniendo en cuenta su débil conexión con la eficacia comunicativa (2002, p. 1574). Es posible, entonces, que las hipótesis adaptacionistas acerca de la evolución del lenguaje resulten siendo al final especulaciones para las que se carece de toda base empírica y que a diferencia del caso del ojo de los vertebrados finalmente no se pueda mostrar que la facultad del lenguaje sea una adaptación que ha evolucionado mediante la selección natural.

Pero sería un error pensar que la única lectura que tendría dicho fracaso es la que proponen Hauser, Chomsky y Fitch, es decir, que la facultad del lenguaje, al menos en sus aspectos más nucleares, habría evolucionado por mecanismos extra-adaptativos como la exaptación. Existe también otra lectura. Hemos visto que el razonamiento que lleva a la explicación adaptacionista, basado en el argumento del diseño, es el siguiente: la facultad del lenguaje sería parte de la dotación biológica de la especie y muestra signos de un diseño complejo, y como la selección natural es la única fuerza capaz de producir estructuras biológicas con un diseño complejo, la facultad del lenguaje debería ser una adaptación moldeada por la selección natural para alguna función (probablemente la comunicación). Ahora bien, si el intento de mostrar que la facultad del lenguaje es una adaptación que habría evolucionado mediante la selección natural finalmente fracasa, entonces hay una explicación alternativa a la de Hauser, Chomsky y Fitch, y es simplemente que la facultad del lenguaje no sería un producto de la evolución. Y no sería un producto de la evolución, porque no sería parte de la dotación biológica de la especie sino que, como ha venido sosteniendo Tomasello (1999 y 2008), el lenguaje formaría parte de la cognición cultural y por tanto no sería un producto de la evolución biológica, sino de la evolución cultural¹³. O como han sostenido por otro lado emergentistas y conexionistas, mecanismos de aprendizaje simples, que operan a lo largo de los sistemas humanos para

¹³ En ese sentido el propio Jackendoff, que es un claro defensor de la tesis de que el lenguaje es parte de la dotación biológica de la especie humana, concede que una conclusión posible de la investigación empírica en el futuro podría ser que el lenguaje sea, finalmente, un fenómeno puramente cultural, como sostiene Tomasello (Jackendoff, 2010, p. 66).

la percepción, la acción motora y la cognición, en la medida en que son expuestos a los datos lingüísticos como parte de un entorno social humano comunicativamente rico, por un organismo deseoso de explotar la funcionalidad del lenguaje, serían suficientes para desencadenar la emergencia de representaciones lingüísticas complejas (Ellis, 1998, p. 657), de modo que el lenguaje no sería el producto de un mecanismo innato, sino que emergería a partir de la interacción de múltiples restricciones, aunque ninguna de estas sería específica para el lenguaje (Elman y otros, 1996).

Algo que debe quedar claro al final de esta discusión es que en el estado actual de los conocimientos no se puede afirmar la validez de ninguna de estas hipótesis. Esto vale tanto para la hipótesis biologista-adaptacionista de Pinker, Bloom y Jackendoff como para la hipótesis biologista-antiadaptacionista de Piattelli-Palmarini, Hauser, Chomsky y Fitch, así como para la hipótesis cognitivo-culturalista de Tomasello o para la hipótesis emergentista-conexionista de Ellis y Elman. Y quien sabe tampoco se pueda hacerlo nunca. Por tanto, la actitud que corresponde tomar en esta controversia debe ser una que integre, por un lado, la admisión de la relatividad de la propia postura así como, por el otro, la apertura y la tolerancia frente a puntos de vista alternativos. En este contexto, pretender que alguna de estas posturas constituye una verdad consensualmente aceptada por la comunidad de científicos cognitivos nos revelaría que estamos saliendo del terreno de la discusión académica para entrar al de la defensa fundamentalista y fanática de una teoría que ha dejado de ser tal, para convertirse más bien en una ideología al servicio de la consolidación de una identidad grupal.

REFERENCIAS BIBLIOGRÁFICAS

- Bickerton, Derek (1990). *Language & Species*. Chicago: Chicago University Press.
- Buss, David M. y otros (1998). Adaptations, Exaptations, and Spandrels. *American Psychologist*, 53(5), 533-548.
- Chierchia, Gennaro & Sally McConnell-Ginet (2000). *Meaning and Grammar. An Introduction to Semantics*. Segunda edición. Cambridge: The MIT Press.
- Chomsky, Noam (1957). *Syntactic Structures*. La Haya: Mouton.
- Chomsky, Noam (1977[1959]). Crítica de «Verbal Behavior» de B. F. Skinner. En Ramón Bayés (ed.), *¿Chomsky o Skinner? La génesis del lenguaje* (pp. 21-86). Barcelona: Fontanella [*Language*, 35(1), 26-58].
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.
- Chomsky, Noam (1966). *Lingüística cartesiana. Un capítulo de la historia del pensamiento racionalista*. Madrid: Gredos.

- Chomsky, Noam (1975). *Reflexiones sobre el lenguaje*. Barcelona: Ariel.
- Chomsky, Noam (1979). À propos des structures cognitives et de leur développement: une réponse à Piaget. En Massimo Piattelli-Palmarini (ed.), *Théories du langage. Théories de l'apprentissage. Le débat entre Jean Piaget et Noam Chomsky* (pp. 65-87). París: Seuil.
- Chomsky, Noam (1980). *Rules and Representations*. Oxford: Blackwell.
- Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin and Use*. Nueva York: Praeger Publishers.
- Chomsky, Noam (1988). *Language and Problems of Knowledge. The Managua Lectures*. Cambridge: The MIT Press.
- Chomsky, Noam (2000). New Horizons in the Study of Language. En *New Horizons in the Study of Language and Mind* (pp. 3-18). Cambridge: Cambridge University Press.
- Chomsky, Noam (2010). Some Simple Evo Devo Theses: How True Might They Be for Language? En Richard K. Larson, Viviane Déprez y Hiroko Yamakido (eds.), *The Evolution of Human Language. Biolinguistic Perspectives* (pp. 45-62). Cambridge: Cambridge University Press.
- Dawkins, Richard (1986). *The Blind Watchmaker: What the Evidence of Evolution Reveals a Universe without Design*. Nueva York: Norton.
- Dawkins, Richard (1995). *River Out of Eden. A Darwinian View of Life*. Londres: Phoenix.
- De Vega, Manuel & Fernando Cuetos (1999). Los desafíos de la psicolingüística. En Manuel de Vega y Fernando Cuetos (eds.), *Psicolingüística del español* (pp. 13-52). Madrid: Trotta.
- Deranty, Jean-Philippe (2009). Beyond Communication. A Critical Study of Axel Honneth's Social Philosophy. *Social and Critical Theory*, 7. Leiden-Boston: Brill.
- Dunbar, Robin (2009). Why Only Humans Have Language. En Rudolf Botha y Chris Knight (eds.), *The Prehistory of Language* (pp. 12-35) Oxford: Oxford University Press.
- Ellis, Nick C. (1998). Emergentism, Connectionism and Language Learning. *Language Learning*, 48(4), 631-664.
- Elman, Jeffrey L. y otros (1996). *Rethinking Innateness. A Connectionist Perspective on Development*. Cambridge: The MIT Press.
- Escalante, Soledad (2012). *Intersubjektivität und Anerkennung. Hegels Ansatz (1802-07) und seine kritische Auslegung bei Honneth (1992)* [disertación]. Fráncfort del Meno: Peter Lang.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky (2005). The Evolution of the Language Faculty: Clarifications and Implications. *Cognition*, 97, 179-210.
- Freud, Sigmund (1900). *Die Traumdeutung* [La interpretación de los sueños]. Tomo II. Fráncfort del Meno: Fischer.

- Freud, Sigmund (1911). *Formulierungen über die zwei Prinzipien des psychischen Geschehens* [*Formulaciones acerca de los dos principios del suceder psíquico*]. Tomo III (pp. 13-24). Fráncfort del Meno: Fischer.
- Gould, Steven J. & Richard C. Lewontin (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society B*, 205(1161) 581-598.
- Gould, Steven J. & Elisabeth Vrba (1982). Exaptation—A Missing Term in the Science of Form. *Paleobiology*, 8(1), 4-15.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch (2002). *The Faculty of Language: What It Is, Who Has It, and How Did It Evolve?* *Science*, 298, 1569-1579.
- Herrera, Marcos (2002). ¿Al pan, «pan» y al vino, «vino»? Un enfoque representacionista no esencialista del significado de las palabras. *Lexis*, XXVI(2), 345-393.
- Humboldt, Wilhelm von (1988[1836]). *Über die Verschiedenheiten des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts* *Gesammelte werke* (pp. 1-425). Tomo 6. Berlín: De Gruyter.
- Jackendoff, Ray (1999). Possible Stages in the Evolution of the Language Capacity. *Trends in Cognitive Sciences*, 3(7), 272-279.
- Jackendoff, Ray (2010). Your Theory of Language Evolution Depends on Your Theory of Language. En Richard K. Larson, Viviane Déprez y Hiroko Yamakido: *The Evolution of Human Language. Bilingual Perspectives* (pp. 63-72). Cambridge: Cambridge University Press.
- Jackendoff, Ray & Steven Pinker (2005). The Nature of the Language Faculty and its Implications for Evolution of Language (Reply to Fitch, Hauser, and Chomsky). *Cognition*, 97, 211-225.
- Martinet, André (1980). *Éléments de linguistique générale*. Novena edición. París: Armand Colin.
- Piattelli-Palmarini, Massimo (1989). Evolution, selection and cognition: From «learning» to parameter setting in biology and in the study of language. *Cognition*, 31, 1-44.
- Pinker, Steven (1994). *The Language Instinct. How the Mind creates Language*. Nueva York: Harper.
- Pinker, Steven & Paul Bloom (1990). Natural Language and Natural Selection. *Behavioral and Brain Sciences*, 13, 707-784.
- Pinker, Steven & Ray Jackendoff (2005). The Faculty of Language: What's Special about It? *Cognition*, 95, 201-236.
- Skinner, Burrhus F. (1957). *Verbal Behavior*. Nueva Jersey: Prentice Hall.

- Tomasello, Michael (1999). *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.
- Tomasello, Michael (2008). *Origins of Human Communication*. Cambridge: Harvard University Press.
- Treibel, Annette (1995). Einführung in soziologische Theorien der Gegenwart. En Hermann Korte y Bernhard Schäfers (eds.), *Einführungskurs Soziologie*. Tomo III. Tercera edición. Opladen: Leske + Budrich.

SECCIÓN IV
ENSAYOS SOBRE DESARROLLO Y ESTRUCTURA DEL LENGUAJE

¿CÓMO SE COMUNICAN LOS INFANTES EN EL SEGUNDO AÑO DE VIDA? EL CASO DE LAS FUNCIONES DECLARATIVA E IMPERATIVA

María de los Ángeles Fernández Flecha

Pontificia Universidad Católica del Perú

1. INTRODUCCIÓN

La psicolingüística evolutiva aborda el proceso por el cual los infantes devienen, al cabo de algunos años, usuarios competentes de la lengua de su entorno. Esta disciplina considera posible aproximarse científica y sistemáticamente a los procesos de comprensión y producción lingüísticas de infantes y niños en edad preescolar. No solo es posible, pues, analizar los patrones tempranos de producción para entender cómo funciona el sistema lingüístico en sí mismo sino que puede resultar ilustrativo de realidades más amplias como el desarrollo cognitivo general. En este sentido, el estudio científico del lenguaje infantil arrojaría luces sobre el fenómeno del lenguaje y la mente humana.

Tomasello y Bates (2001, p. 10) han propuesto que la psicolingüística evolutiva analiza los procesos por los cuales los niños se valen de su herencia biológica y cultural para convertirse, tras pocos años, en usuarios competentes de su lengua. Se trata de un proceso en el que confluyen habilidades perceptivas, cognitivas, comunicativas, sociales y de aprendizaje. En la explicación del proceso de aprendizaje lingüístico, los tradicionales conceptos de *nature* (naturaleza o biología) y *nurture* (crianza o cultura) tendrían ambos un papel importante e interactuarían para dar lugar al estadio final de conocimiento y de uso maduro de una lengua.

Al ser el lenguaje infantil un producto aún en proceso de formación, exhibe un carácter inestable, por lo que observamos mayor variabilidad (respecto del modelo adulto) en la relación entre forma y contenido de los ítems prelingüísticos analizados. En este sentido, más allá de los aspectos formales de la conducta comunicativa infantil, tanto los parámetros situacionales de su emisión como las reacciones por parte del entorno familiar serán claves en la determinación de la intención específica expresada.

A pesar de su carácter transicional, el sistema comunicativo infantil —que devendrá, eventualmente, dadas ciertas condiciones mínimas necesarias como la existencia de un entorno lingüístico, en el adulto— encierra también un orden y una regularidad cuyo estudio enriquece la comprensión del lenguaje humano.

El presente texto se interesa, pues, por el desarrollo ontogenético del lenguaje, específicamente por su uso temprano para la comunicación en el contexto de la interacción espontánea en el hogar. Nos enfocamos en el papel de las lenguas como herramientas para la comunicación, que florecen y se nutren en el encuentro intersubjetivo con los usuarios adultos de dicho sistema, a partir de quienes se interiorizan las pautas de expresión y de uso que permitirán a los infantes aproximar progresivamente su producción al modelo adulto.

2. ¿QUÉ SE APRENDE CUANDO SE APRENDE UNA LENGUA?

El aprendizaje de una lengua es el proceso por el que ciertas predisposiciones biológicas humanas entran en contacto e interactúan con ciertos factores sociales y culturales para dar lugar a un conocimiento lingüístico específico, convencional en una comunidad determinada. Aunque ciertas evidencias llevan a pensar que el lenguaje humano tiene una base biológica específica de la especie humana y de ninguna otra (predisposición morfológica para la producción de sonidos articulados, ciertas características comunes a una mayoría de lenguas del mundo, la existencia de un patrón común de desarrollo lingüístico en diversas culturas), resulta innegable el papel que la cultura y el entorno ejercen en el aprendizaje de una lengua por parte del infante.

En este sentido, si bien es la propia constitución biológica del ser humano la que lo pone en condiciones de adquirir una lengua cualquiera, conocer una lengua implica mucho más que ser capaz de construir enunciados gramaticales: supone, también, saber usarla adecuadamente. Así, pues, no basta con aprender palabras y reglas de construcción, sino que, además, el infante debe aprender cómo expresarse según dónde se encuentre, a quién se dirija, qué pretenda obtener de su interlocutor, etcétera. Este conocimiento pragmático es una parte crucial de lo que supone ser un usuario competente de la lengua y no puede existir en el infante de forma determinada, biológica, pues varía cultural y socialmente. Más bien, la interacción con sus cuidadores, en el entorno, le dará al niño las pautas de realización comunicativa de los distintos actos de habla que puede producir.

A partir de la combinación de los elementos biológicos y los factores culturales, el infante atraviesa el proceso que desemboca en la adquisición completa, madura, de su lengua materna. El proceso de aprendizaje de una lengua comprende, pues, una

secuencia de momentos evolutivos, cada uno de los cuales se caracteriza por ciertos aspectos que resultan más predominantes entonces que en otros previos. No ocurre que, con el paso de una etapa a la siguiente, la conducta del niño se modifique de forma absoluta (lo inmaduro desaparece para dar lugar a lo maduro y no reaparecer más); más bien, se acentúa o generaliza un poco más cierta característica que puede ser considerada más madura respecto de la característica predominante en la etapa anterior. Así, va variando la frecuencia de ocurrencia de la conducta más madura, que se impone sobre la menos madura de forma progresiva y no repentina o absoluta.

Además, en el curso del desarrollo de cierta habilidad, pueden observarse patrones de tipo U; en estos casos, antes de alcanzar un nivel de desempeño más cercano al modelo adulto, el desempeño del niño deviene de pronto menos bueno de lo que venía siendo —un aparente retroceso en la maestría alcanzada hasta ese momento— para volverse luego aun mejor. De ahí el nombre de curva U de desarrollo, debido a la aparente caída que podemos observar en el nivel de desempeño del niño.

Los aparentes retrocesos observados en este proceso de desarrollo pueden explicarse a partir del funcionamiento del mismo sistema-que-aprende-la-lengua (el niño). Este administraría los pesos asignados a los diversos tipos de estímulo que recibe (input lingüístico) hasta lograr algo muy parecido a un aprendizaje exitoso. De este modo, el sistema construye progresivamente su conocimiento a partir de los estímulos recibidos y, también, de su propio nivel de conocimiento en cada momento. Así, en dicho proceso, «el sistema que aprende la lengua vuelve sobre cada adquisición formal parcial a distintos niveles de complejidad de las representaciones y del procesamiento, vuelve sobre un mismo problema en distintos momentos del desarrollo (fases)» (López Ornat en López Ornat y otros, 1994, p. 103). De este modo, replantea sus aprendizajes parciales y a menudo tentativos de forma progresiva hasta aproximarse al estadio final de adquisición de un determinado patrón o uso lingüístico.

De acuerdo con el emergentismo, pues, mecanismos simples de aprendizaje pueden, a partir de su exposición al input lingüístico en el marco de un entorno social, ser suficientes para explicar la emergencia de representaciones lingüísticas complejas (Ellis, 1998, p. 631). Estos mecanismos generales de aprendizaje operarían en los sistemas humanos de percepción, acción motriz y cognición a la vez, de forma interrelacionada. Del encuentro entre dichos procesos simples de aprendizaje, y un problema de gran riqueza y complejidad como es el lenguaje, durante largos periodos de práctica en la vida del sujeto, se derivaría un resultado final altamente complejo: el conocimiento de la lengua por parte del individuo. La gramática y las representaciones lingüísticas en general serían resultados emergentes en el contexto de este proceso. El lenguaje puede ser concebido, pues, como un sistema emergente

y, a la vez, autoformante en el que los avances en un nivel contribuyen a los avances en el nivel siguiente.

En suma, la emergencia del lenguaje resultaría de la interacción entre los mecanismos del sistema cognitivo (su estructura biológica) y el ambiente (la información lingüística implícita en el input). El sistema-que-aprende-la-lengua cambiaría para adaptarse a las continuas variaciones de complejidad en el input, las que, a su vez, se originan en el propio sistema: cada cambio en el estado de conocimiento del sistema motiva un cambio en la manera en que el input es percibido por el mismo (López Ornat en López Ornat y otros, 1994; Karousou, 2003).

A continuación, presentamos algunos de los elementos que el niño aprende en el proceso de desarrollo comunicativo y lingüístico. La lista no pretende ser exhaustiva (deja de lado, por ejemplo, el aprendizaje fonológico) sino recoger aquellas habilidades más estrechamente relacionadas con la presente investigación.

2.1. Las intenciones comunicativas

La aparición de la conducta intencional constituye uno de los principales hitos en el desarrollo infantil. Bruner (1973) observa que, ya en el primer año de vida, resultan comunes los episodios en que el niño despliega una intención mientras varía los medios para alcanzar su meta. En el caso específico de las intenciones comunicativas, existe consenso respecto de su aparición hacia los diez meses de edad. Esta intencionalidad sería, además, modelada por el adulto cuidador en el marco de dinámicas o formatos de interacción (Bruner, 1973 y Tomasello, 2001), elementos centrales en el planteamiento del LASS de Bruner (*Language Acquisition Support System* o Sistema de Apoyo para la Adquisición del Lenguaje). El aprendizaje lingüístico y comunicativo se desarrollaría en el contexto de actividades coordinadas con usuarios maduros del lenguaje, a la manera de actos de habla primitivos que facilitan la entrada del niño en el lenguaje. En estos marcos de interacción, el niño aprendería a comunicarse sobre la base de las características predecibles de la acción que se está llevando a cabo de forma conjunta con el adulto. Estas funciones comunicativas serían logradas, tempranamente, a través de medios prelingüísticos (Bruner, 1983, p. 31).

Hacia la misma edad, el niño cuenta con los prerrequisitos sociocognitivos para participar en rutinas de acción conjunta de manera un poco más sofisticada: logra comprender las metas e intenciones de los demás —no solo su objetivo final sino, también, en cierta medida, sus planes de acción, es decir, los medios empleados para alcanzar dicho fin—; comprende, aunque de forma básica, el conocimiento compartido —lo que el interlocutor sabe, porque el infante lo ha experimentado con él, que va más allá de la atención conjunta en tanto trasciende el presente perceptivo—;

y cuenta con la habilidad y la motivación de ayudar a los demás a alcanzar alguna meta. A partir de estas habilidades presentes en infantes de un año de edad, es posible inferir una teoría de la mente mucho más robusta de lo anteriormente sospechado (Carpenter, 2009).

Nos basamos en la concepción de Tomasello (2008) de la comunicación humana como una empresa fundamentalmente cooperativa; en esa medida, la capacidad de establecer y mantener intenciones compartidas con otros resultaría clave para adquirir una lengua. La lectura de intenciones, que emergería entre los nueve y los doce meses y haría posible la aparición de intenciones comunicativas en el infante, incluiría habilidades como las siguientes (Tomasello, 2003): (1) compartir atención con otras personas respecto de objetos o eventos de mutuo interés (atención triádica, también conocida como triangulación en Davidson, 2001); (2) seguir la atención y gesticulación de otras personas respecto de objetos o eventos distintos, más allá de la interacción inmediata; (3) dirigir activamente la atención de otros a objetos distantes por medio de gestos como el de señalar, mostrar o gestos no lingüísticos; y (4) aprender culturalmente, imitativamente, las acciones intencionales de otros, incluyendo sus actos comunicativos.

Hacia su primer cumpleaños, el niño empieza a producir las primeras convenciones léxicas. La adaptación humana para la comunicación simbólica se hace evidente conforme se van desarrollando nuevas habilidades sociocognitivas: atención conjunta, lectura de intenciones y el tipo de aprendizaje cultural caracterizado por la imitación con inversión de roles. En primer lugar, la existencia de marcos de atención conjunta hace posible que tanto adulto como infante comprendan sus acciones o contribuciones al intercambio en curso a la luz de las actividades que están llevando a cabo y de la meta en común que saben que comparten.

En segundo lugar, en estos marcos de atención conjunta, el niño comprende las intenciones comunicativas de su interlocutor. Las intenciones comunicativas constituyen un tipo particular de intenciones en tanto se dirigen no solo hacia un objeto (sobre el que se ejecuta o realiza cierta acción) sino hacia los estados intencionales de otro sujeto (el interlocutor). El terreno intersubjetivo creado por los marcos de atención conjunta permite al niño y al adulto comprender sus intentos comunicativos.

En tercer lugar, una vez que el niño ha comprendido que los demás, al igual que él, establecen relaciones intencionales con el mundo, puede concentrarse en los medios conductuales de los que se valen para lograr sus objetivos. A partir de esto, aprenderá los símbolos comunicativos por medio de una imitación de rol revertido: no solo debe sustituir al adulto como actor, sino, también, sustituirse por el adulto como objetivo del acto comunicativo intencional, es decir, reemplazar su propio estado atencional como meta por el del adulto.

En este proceso de desarrollo, las intenciones comunicativas infantiles que inicialmente son expresadas a través de medios tradicionalmente conocidos como «prelingüísticos» (por ejemplo, el gesto de señalar o ciertas vocalizaciones más o menos cercanas al balbuceo) serán luego expresadas de forma cada vez más lingüística, conforme el niño avance en el proceso de aprendizaje de su lengua.

2.2. La pragmática: actos de habla

Bruner (1974) considera que, en el caso del estudio del lenguaje infantil, conviene concentrarse en el uso, no en la forma, de su producción. El despliegue de funciones lingüísticas que el niño produce resultaría crucial para comprender cómo es finalmente adquirida la lengua del entorno. De acuerdo con dicho autor, el lenguaje no sería adquirido tanto como sistema formal de unidades y reglas de combinación, sino, fundamentalmente, como un instrumento para regular la actividad y la atención conjunta en el contexto de patrones de interacción ritualizada entre infante y cuidador. En este sentido, el lenguaje sería una extensión especializada y convencional de acción cooperativa.

Las teorías que buscan explicar el desarrollo lingüístico y comunicativo infantil sobre la base de la competencia y las habilidades pragmáticas del niño hallan en el acto de habla una noción muy valiosa y útil. El acto de habla constituye un tipo de acto social. Se trata, específicamente, de acciones llevadas a cabo a través de enunciados (Yule 1996, 2008).

De acuerdo con Austin (1962), podemos distinguir tres tipos de actos que tienen lugar cuando se producen enunciados lingüísticos: locutivos, ilocutivos y perlocutivos. El acto locutivo constituye el acto básico de enunciación y supone la producción de una expresión lingüística significativa. Sin embargo, no producimos locuciones lingüísticas sin un propósito o intención en mente. El acto ilocutivo encierra la fuerza comunicativa del enunciado y es conocido, también, como la fuerza ilocucionaria del acto de habla. Searle (1965) ha afirmado que, si bien el acto de habla ilocucionario (o acto de habla) no tiene necesariamente contenido proposicional, sí encierra siempre una fuerza ilocucionaria. De ello se deduce que, ya sea por medios lingüísticos o no lingüísticos (gestuales, por ejemplo), los individuos pueden comunicar la fuerza o intención (la ilocución) correspondiente a sus locuciones —piénsese en una orden, una demanda, una promesa, etcétera. Finalmente, no puede decirse que creamos un enunciado con una función en mente sin pretender que repercuta de una forma particular en el oyente; este efecto de la locución constituye la perlocución o efecto perlocutivo.

A partir de los componentes del acto de habla, Bates y otros proponen una progresión de 3 etapas de desarrollo lingüístico (1975, p. 207). De acuerdo con este planteamiento, ciertas estructuras performativas (tipos de actos de habla entendidos como ilocuciones) tienen una historia evolutiva que precede al habla en sí misma. En primer lugar, el infante atravesaría una etapa perlocucionaria, en la que, si bien su conducta tiene un efecto sistemático en su oyente, carece de un control intencional o consciente de dicho efecto. Piénsese, por ejemplo, en el efecto del llanto de un niño pequeño; sus padres se apresurarán a consolarlo sin que este, en un inicio al menos, haya llorado con esa intención. En segundo lugar, en la etapa ilocucionaria, el niño empieza a hacer un uso intencional de señales no verbales con el objetivo de transmitir demandas y, también, de dirigir la atención del adulto hacia determinados objetos y eventos que hayan llamado su atención. Piénsese en las ocasiones en que el niño señala cierto evento inesperado en su entorno, como la caída de un objeto o el efecto divertido del funcionamiento de algún juguete, mirando al adulto para asegurarse de que este preste atención. Finalmente, en el caso de la etapa locucionaria, el niño es ya capaz de construir proposiciones y de emitir sonidos de habla en el marco de las mismas secuencias performativas que antes producía de forma no verbal. Piénsese, por ejemplo, en el paso de señalar un objeto para demandar al adulto que se lo entregue a producir la forma «dame» con la misma intención.

Dore introduce la noción de «acto de habla primitivo» para referirse a aquellos enunciados que consisten formalmente de una sola palabra o contorno prosódico y que transmiten las intenciones de un niño que no puede aún producir oraciones. El autor sostiene que «[...] antes de que el niño adquiera estructuras oracionales, posee un conocimiento sistemático de la pragmática de su lengua que puede ser mejor descrito en términos de “actos de habla primitivos”» (1974, p. 344).

Esta unidad de análisis permite, pues, captar la productividad comunicativa del infante que no es aún capaz de producir un lenguaje más convencional en términos más estrictamente lingüísticos, aunque sí exhibe un desempeño comunicativo eficaz al lograr transmitir sus intenciones en el contexto de su entorno más inmediato. La noción de acto de habla primitivo rescata la habilidad temprana del infante para la comunicación en momentos en que su conocimiento de la lengua resulta aún bastante inmaduro —en este sentido, los medios empleados para la expresión involucran aspectos como la prosodia, de desarrollo más temprano que la fonología—.

Como habíamos adelantado líneas arriba, a partir de la teoría de los actos de habla, Bates y otros proponen que ciertas estructuras performativas se hallan ya bastante desarrolladas para el inicio del segundo año de vida, como los protoimperativos y los protodeclarativos. Los protoimperativos recogen el uso intencional que hace el niño del oyente como agente o herramienta para lograr una meta determinada.

En el caso del protodeclarativo, se trataría de un «esfuerzo preverbal para dirigir la atención del adulto hacia cierto evento u objeto en el mundo» (1975, p. 208).

Además, a diferencia de Dore, las autoras defienden una relación de continuidad en el desarrollo lingüístico y comunicativo temprano; así, los primeros declarativos de una palabra emergerían a partir de conductas declarativas que involucran vocalizaciones prelingüísticas y gestos como señalar, entregar y mostrar. De acuerdo con la hipótesis de la continuidad, el paso del desarrollo prelingüístico al lingüístico supondría una especie de superposición por la cual, solo gradual y progresivamente, ciertas habilidades devendrían propiamente lingüísticas. No sería un cambio que pueda identificarse con absoluta precisión. La división entre balbuceo y lenguaje parecería reflejar, pues, las características de la percepción e interpretación adulta, más que los procesos evolutivos que tienen lugar en el niño (Vihman y otros, 1985). De este modo, se superponen temporalmente la adquisición gradual de palabras y el uso continuado del balbuceo. El repertorio fonético del balbuceo de un niño se reflejará, además, más o menos con los mismos niveles de frecuencia, en la clase de palabras de tipo adulto que dirá y en cómo las dirá, es decir, en su producción fonética (Vihman, 1996). Las primeras palabras seguirían, pues, ciertos patrones exhibidos ya en el balbuceo.

En suma, se propone que los logros de la etapa comúnmente referida como prelingüística evidencian ya un orden y una sistematicidad que puede rastrearse y explicar, a su vez, las características de la etapa más propiamente lingüística. Finalmente, debe añadirse que el emparejamiento productivo de forma y sentido, de sonido y significado, que se considera el hito representado por la aparición de las palabras y, en ese sentido, el inicio de la etapa comúnmente considerada lingüística, precede, en realidad, a la aparición de las primeras palabras de tipo adulto. Así, aunque aún inmaduras segmental y referencialmente, observamos en las vocalizaciones (típicamente consideradas prelingüísticas) las bases de una relación más o menos estable entre forma (suprasegmental o prosódica, por ejemplo) y sentido (función, por ejemplo).

2.3. Las vocalizaciones y las (proto)palabras

En los momentos iniciales del desarrollo lingüístico, observamos que conviven —o, tal vez, se siguen muy de cerca, superponiéndose en ciertos momentos de tránsito— tres formas referenciales de producción vocal. En primer lugar, entendemos por vocalización aquella emisión vocal que precede a desarrollos más complejos como la protopalabra o la palabra. López Ornat y Karousou definen las vocalizaciones como «producciones tempranas interpretables por los padres, con o sin contenido

segmental, en las que el niño produce grupos prosódicos de duración variable [...] ensayos lingüísticos de distinta naturaleza: articulatorios, prosódicos, comunicativos» (2005, p. 401). Así, la vocalización se caracteriza por no presentar una forma fonológica o segmental estable y clara que vaya asociada a un sentido forma consistente, sino, más bien, por ir acompañada de un contorno entonativo que le brinda unidad. Aunque el balbuceo es el caso más conocido, existen otros tipos de vocalizaciones algo posteriores que involucran una mayor variedad de sonidos sin alcanzar aún el nivel de protopalabras.

En cuanto a las protopalabras, Vihman (1996) las concibe como formas infantiles relativamente estables, de uso relativamente consistente, aunque pueden carecer de vínculo claro con la unidad convencional correspondiente del modelo adulto. En este sentido, se definen no por su parecido formal con el modelo adulto, sino por su uso consistentemente asociado con cierto contexto. En estas formas reconocemos, pues, una asociación más estable entre forma y sentido o uso. Aunque en sus aspectos formales las protopalabras pueden alejarse bastante del modelo adulto correspondiente, debido a la asociación más o menos consistente que establece el niño entre una forma sonora y un contexto de uso las reconocemos como «palabras infantiles» o protopalabras, como símbolos vocales tempranos.

Finalmente, las palabras son las formas léxicas características del lenguaje adulto, en las que se combinan una consistencia de uso aún más fina, y una forma fonológica ya estable y convencional. Tradicionalmente, las palabras son tomadas como la evidencia más clara de que el niño se halla, de manera incontestable, en posesión de una lengua y en proceso de desarrollarla aún más. Para que una emisión vocal sea considerada una palabra debe, por una parte, exhibir un uso consistente en uno o más contextos (lo que evidencia un significado asociado de forma invariable a la forma usada) y, por otra parte, presentar una forma sonora (segmental) que remita, de modo claro, a la forma adulta correspondiente.

2.4. La prosodia

Vihman (1996) propone que los patrones prosódicos de la lengua se adquieren de forma temprana, probablemente debido a que los rasgos prosódicos son especialmente salientes para el infante desde muy temprano y a que el número de patrones prosódicos posibles es menor que el número de segmentos, lo que redundaría en un manejo más temprano y mejor de los aspectos prosódicos del lenguaje, frente a la dificultad que implicarían los aspectos segmentales o fonológicos.

De acuerdo con autores como Snow (2006), la entonación seguiría un curso de desarrollo no lineal, caracterizado por un patrón de regresión y reorganización

de base lingüística. Frente a la teoría de los grupos respiratorios (Lieberman, 1967 y 1984), que predice que los contornos descendentes ocurrirán más frecuentemente y con un rango de acento mayor que los ascendentes, y otras teorías que postulan que la entonación se desarrolla recién después de la adquisición de las primeras palabras (con lo cual estaría vinculada con la experiencia lingüística y no habría, en realidad, diferencias evolutivas entre contornos ascendentes y descendentes), Snow propone un patrón no lineal de desarrollo: la entonación sufriría una regresión hacia los diez meses de edad, que se mantendría por casi todo el periodo de una palabra y terminaría antes del segundo cumpleaños. Este periodo de regresión reflejaría un proceso de reorganización y cambio lingüístico, el paso de un sistema de entonación prelingüístico a uno lingüístico.

De acuerdo con dicha propuesta, en la etapa más temprana, los tonos descendentes se encontrarían bajo control fisiológico. Más adelante, sin embargo, una vez ocurrida la reorganización de base lingüística, pasarían a estar bajo control lingüístico. La regresión en la producción entonativa podría explicarse a partir de la aparición de la comunicación intencional hacia los diez meses de edad. En este punto, la entonación deviene especialmente relevante en tanto los contornos entonativos codifican las intenciones pragmáticas del infante. Más adelante, después de la reorganización, tendría lugar una segunda discontinuidad, que marcaría la verdadera adquisición de la entonación. Hacia el final del periodo de una palabra (entre los dieciocho y los veinticuatro meses de edad dependiendo del niño), coincidiendo con la adquisición de la sintaxis expresiva y la aparición de las combinaciones de dos palabras, tendría lugar un dramático aumento en la producción entonativa. Esto se explicaría a partir del hecho de que los contornos entonativos marcan fonéticamente los límites de las unidades sintácticas mayores, como cláusulas y oraciones, que los niños empiezan a producir en ese momento.

Así, en momentos tempranos del desarrollo comunicativo, la modulación de la prosodia del enunciado deviene para los infantes una herramienta sumamente útil para comunicar sus intenciones; en este proceso, se establecen asociaciones entre forma, y contenido o función comunicativa. La prosodia permitiría al infante descubrir que el proceso comunicativo supone variaciones en el nivel de la forma que corresponden a precisiones en el nivel del sentido o la función pragmática.

Finalmente, diversos estudios de lenguas como el inglés han hallado que los infantes son capaces de asociar en su producción ciertas características prosódicas con ciertas funciones o intenciones comunicativas. Así, Halliday (1975) halló una asociación entre tonos ascendentes y aquellas vocalizaciones que requieren una respuesta por parte del receptor. Marcos (1987), también, halló que los tonos ascendentes

ocurren más frecuentemente asociados con demandas (imperativos), mientras que los tonos descendentes suelen acompañar a las denominaciones (coincidentes con los declarativos aquí planteados). Flax y otros (1991) sostuvieron que los tonos finales ascendentes suelen acompañar a las demandas (aunque se trató de una asociación relativa), y que funciones como la protesta y las demandas de atención tienden a exhibir una frecuencia fundamental más alta que las respuestas o los comentarios no interactivos. Por último, en el caso del italiano, D'Odorico y Franco (1991) hallaron que las vocalizaciones con función asertiva se diferencian prosódicamente de las demandas, las cuales suelen ir acompañadas de un contorno melódico final ascendente y un tono más alto.

3. METODOLOGÍA

La investigación presentada corresponde a un estudio de caso único (N=1) longitudinal de observación naturalista. Se registró la comunicación en el marco del hogar, con los cuidadores habituales del niño.

Se registraron un total de 30 grabaciones, cada una de aproximadamente 30 minutos, lo que arrojó un total de 15 horas de grabación. Las de la primera etapa de la investigación fueron semanales. Las de la segunda etapa fueron mensuales. De este modo, se logró una recogida densa de información en la etapa más temprana del desarrollo comunicativo infantil (catorce-diecinove meses) y una menos densa en la etapa posterior (veinte-veinticuatro meses). El primer periodo se extiende, según la bibliografía, desde los meses en que observamos un tipo de comunicación a menudo referida como prelingüística hasta un momento cercano a la conocida explosión del vocabulario, en que el número de palabras (o protopalabras) aumenta considerablemente. El segundo periodo, por lo menos en el caso del niño aquí considerado, formaría aún parte del periodo de una palabra.

El niño estudiado es el segundo hijo de una familia de clase media alta monolingüe en español. No presentaba ningún trastorno del desarrollo ni había sufrido de ninguna enfermedad que pudiera interferir con la investigación; tampoco presentaba ningún atraso lingüístico, según los resultados del Inventario de Desarrollo Comunicativo MacArthur (López Ornat y otros, 2005), aplicado a los veinte y a los veintiséis meses.

Como parte del proceso de codificación, se descartaron los sonidos puramente vegetativos (risa, llanto, tos, sonidos de esfuerzo, etcétera) y aquellas conductas comunicativas producidas por el niño fuera de la toma de la cámara o completamente de espaldas a esta.

Se codificaron como comunicativas las conductas que cumplieron con las siguientes condiciones (no necesariamente todas a la vez): a) se acompañaba con una mirada orientada al interlocutor; b) se producía de forma contingente respecto de una emisión o conducta inmediatamente anterior del interlocutor; c) producía una reacción o respuesta espontánea en el interlocutor.

En el caso de las unidades vocales, se empleó como unidad de análisis el grupo o ciclo respiratorio, que comprende la emisión producida en una sola exhalación o, si se quiere, entre dos inspiraciones (generalmente manifestadas por pausas en la producción vocal del niño). Presenta límites naturales, producto de la constitución fisiológica del ser humano, y suele coincidir con la delimitación por funciones pragmáticas; la función pragmática suele cumplirse en un grupo o ciclo respiratorio (Karousou, 2003). En el caso de las unidades gestuales, se consideró como unidad cada emisión de un gesto y se asumió que este concluía cuando el niño cambiaba la posición del brazo o la postura particular que constituía el gesto.

Para la presente investigación, se seleccionaron aquellas conductas con función declarativa o imperativa (demandas).

Tabla 1. Funciones comunicativas¹

Función	Descripción
Declarativa	El niño llama la atención del interlocutor sobre algún elemento del entorno (objeto, evento o acción, persona o personaje) a través de la referencia explícita a este —lo cual no implica, necesariamente, el empleo del término léxico convencionalmente usado para referirse a él, sino, más bien, el despliegue de conductas como gestos o la mirada que hagan evidente que el niño se está refiriendo a un evento o entidad específicos— sin que medie intención de que este último lo entregue al niño o lleve a cabo alguna acción con él. El niño busca alinear la atención del adulto con la suya, de modo que ambos compartan un mismo foco en un momento, una especie de sintonización de estados atencionales.
Imperativa (o demanda)	El niño demanda o solicita algo del interlocutor (un objeto, la realización de alguna acción, que le brinde cierta información o que le preste atención). El infante busca, pues, alinear la intención del adulto con la suya con el objetivo de lograr lo que desea. Así, no bastará con que el adulto preste atención a un determinado objeto sino que tendrá que, por ejemplo, entregarlo al infante para que este sienta que la función comunicativa de su expresión ha sido comprendida y, además, satisfecha.

Adicionalmente, cada conducta comunicativa fue analizada con respecto a las siguientes variables.

¹ Todas las tablas y figuras han sido elaboradas por la autora.

Tabla 2. Variables consideradas en el análisis formal

Variables	Valores
Modalidad	Vocal: emisión producida de forma oral.
	Mixta: combinación de una emisión vocal y un gesto.
	Gestual: acción producida con intención comunicativa, que se expresa, típicamente, con los dedos, las manos y los brazos, aunque puede incluir también rasgos faciales o movimientos del cuerpo.
Nivel referencial	Vocalización: emisión cuya forma segmental no se relaciona de forma más o menos consistente con un sentido o contexto de uso particular.
	Protopalabra: emisión que exhibe una relación más estable con un contexto delimitado de uso, semejante al que es abarcado por la palabra adulta correspondiente.
Carácter interaccional	Espontánea: emisión que no se construye como segundo miembro de un par obligado iniciado por el interlocutor.
	Respuesta: segundo miembro de un par obligado pregunta-respuesta o demanda de acción-respuesta iniciado por el interlocutor.
Orientación de la mirada	Objeto o referente: orientada hacia el objeto o persona (distinta del interlocutor) que hace las veces de referente o tema de la interacción.
	Interlocutor: orientada hacia la persona con que el niño está hablando.
	Alterna: orientada alternadamente hacia el objeto/referente tema de la interacción y hacia el interlocutor.
	Nada: distinta de las opciones anteriores, esta mirada parece no dirigirse a nada relacionado con la interacción o a nada en particular.
Contorno final de entonación*	Ascendente: la línea de producción de F_0 , considerada desde el último tono nuclear, describe una subida.
	Descendente: la línea de producción de F_0 , considerada desde el último tono nuclear, describe una caída.
	Suspensivo: la línea de producción de F_0 parece describir una línea llana, prácticamente sin alzas ni bajas.
Frecuencia fundamental (F_0) media	Auditivamente, la F_0 se percibe como altura tonal o entonación (Quilis, 1993). Es un producto de la frecuencia de vibración de las cuerdas vocales, directamente proporcional a la presión del aire proveniente de los pulmones y a la tensión de ciertos músculos laríngeos. Se mide en Hertz (Hz).

* Se consideró que un cambio en la línea del tono debía ser de aproximadamente 20 Hz a partir del último tono nuclear (coincidente con la última sílaba tónica) para ser considerado perceptible y, por ende, relevante. Se empleó, para los análisis prosódicos, el programa PRAAT y se leyó e interpretó cada resultado de forma individual.

4. RESULTADOS

4.1. Comparación de las funciones declarativa e imperativa (o demanda)

Se presenta la distribución de frecuencia de los valores de las distintas variables consideradas: modalidad, nivel referencial, carácter interaccional, orientación de la mirada, y contorno final de la entonación; así mismo, se presentan las medianas de la frecuencia fundamental media. Se analizó la siguiente cantidad (N) de conductas para cada variable en el caso de cada función.

Tabla 3. N (cantidad de conductas) para cada variable y función

	Modalidad	Nivel referencial	Carácter interaccional	Orientación de la mirada	Contorno entonativo final	F ₀ media	Total
Declarativos	1714	1613	1714	1714	1468	1387	1714
Imperativos	1259	1104	1259	1259	880	880	1259

*Aquellos casos en que la cantidad de conductas por variable no coincide con el total correspondiente a la función pueden deberse a dos causas: (1) resultó imposible atribuir un valor a alguna(s) conducta(s) con certeza suficiente o (2) la variable involucra un componente necesariamente vocal, lo que excluye a las conductas puramente gestuales.

Tabla 4. Modalidad de declarativos e imperativos

	Vocal	Mixta	Gestual	Total
Declarativos	71%	23%	6%	100%
Imperativos	60%	28%	12%	100%

Tabla 5. Nivel referencial de declarativos e imperativos

	Vocalizaciones	Protopalabras	Total
Declarativos	56%	44%	100%
Imperativos	69%	31%	100%

Tabla 6. Carácter interaccional de declarativos e imperativos

	Espontáneos	Respuestas	Total
Declarativos	63%	37%	100%
Imperativos	76%	24%	100%

Tabla 7. Orientación de la mirada en declarativos e imperativos

	Obj./Ref.	Inter.	Alter.	Nada	Total
Declarativos	55%	22%	13%	10%	100%
Imperativos	22%	50%	11%	17%	100%

Tabla 8. Contorno entonativo final de declarativos e imperativos

	Descendente	Ascendente	Suspensivo	Total
Declarativos	61%	25%	14%	100%
Imperativos	53%	38%	9%	100%

Tabla 9. F₀ media de declarativos e imperativos

Declarativos	341
Imperativos	390

*Los valores se encuentran en Hertz.

Observamos que ambas funciones son, en el segundo año de vida, predominantemente vocales, si bien la función imperativa (o demanda) presenta un uso gestual mayor que la declarativa (tanto si consideramos las conductas gestuales puras como las mixtas). Los declarativos son más propensos a ocurrir como respuesta que las demandas, de carácter más típicamente espontáneo. En cuanto al uso de la mirada, cada función tiene un perfil bastante claro: los declarativos se caracterizan por orientar la mirada, de forma predominante, hacia el referente (u objeto o persona tema de la interacción) del intercambio comunicativo, mientras que las demandas suelen dirigirla al interlocutor. En tercer lugar, si bien en ambas funciones observamos un porcentaje importante de protopalabras (que, sin embargo, en ninguna de las funciones supera la ocurrencia de vocalizaciones), este es mayor en el caso de los declarativos. Finalmente, con relación al uso de la entonación, notamos la preferencia en ambos casos por el contorno entonativo final descendente, el predominio relativo del ascendente en la función imperativa y la F₀ media más alta en los imperativos frente a los declarativos.

4.2. Desarrollo longitudinal de las funciones declarativa e imperativa (o demanda)

Los análisis presentados a continuación se llevaron a cabo con el estadístico chi-cuadrado. Se analizó el desarrollo longitudinal de las variables consideradas para cada función por separado. Para ello, se dividió el periodo estudiado en cuatro momentos o grupos de edad.

Tabla 10. Grupos de edad considerados en el análisis longitudinal (en cifras)

Grupo de edad	Sesiones de grabación	Edad inicio	Edad final	Rango aproximado
GR1	1-10	1;1;22	1;3;24	14-16 meses
GR2	11-20	1;4;0	1;6;4	16-18 meses
GR3	21-27	1;6;11	1;8;14	18-20 meses
GR4	28-30	1;9;17	1;11;27	21-24 meses

En el caso de la evolución longitudinal de la variable modalidad, en los declarativos, el análisis de chi-cuadrado arrojó un valor de $\chi^2=113,078$, $GL=6$ y $p<0,001$, es decir, se obtuvo un resultado significativo. En cuanto a las conductas con función imperativa, el análisis de chi-cuadrado arrojó un valor de $\chi^2=72,344$, $GL=6$ y $p<0,001$, un resultado nuevamente significativo.

Figura 1. Desarrollo longitudinal de la modalidad en los declarativos

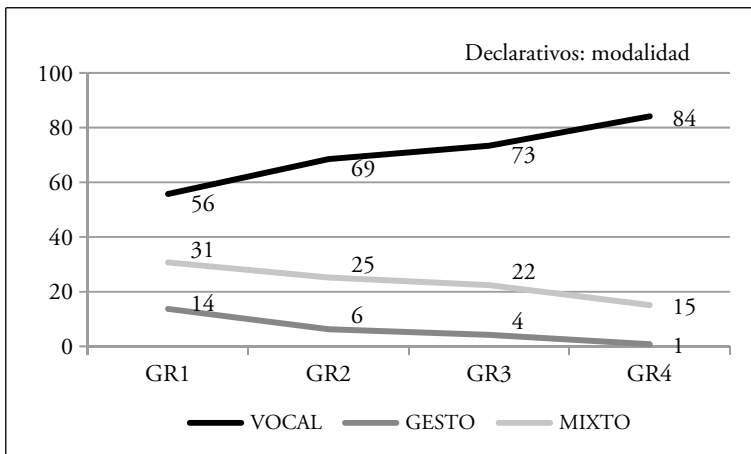
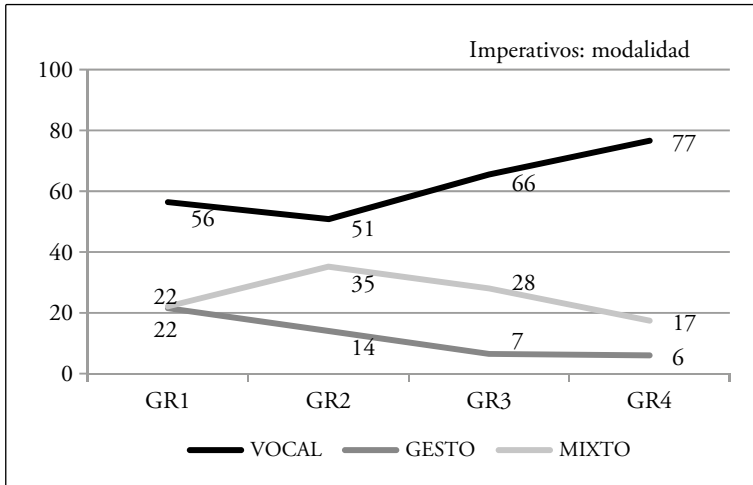


Figura 2. Desarrollo longitudinal de la modalidad en los imperativos



Con respecto al desarrollo longitudinal del nivel referencial de las producciones vocales declarativas, se obtuvieron los siguientes resultados: $\chi^2=122,219$, $GL=3$ y $p<0,001$, un resultado significativo. En el caso de la función imperativa, el análisis de chi-cuadrado arrojó un valor de $\chi^2=107,310$, $GL=3$ y $p<0,001$, nuevamente un resultado significativo.

Figura 3. Desarrollo longitudinal del nivel referencial en los declarativos

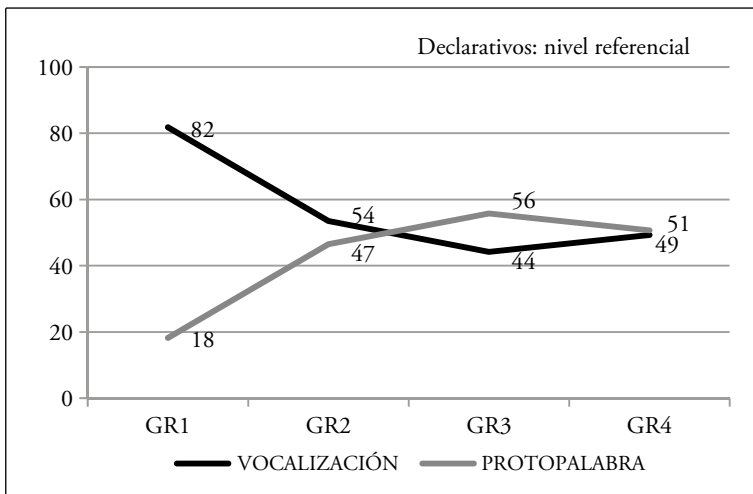
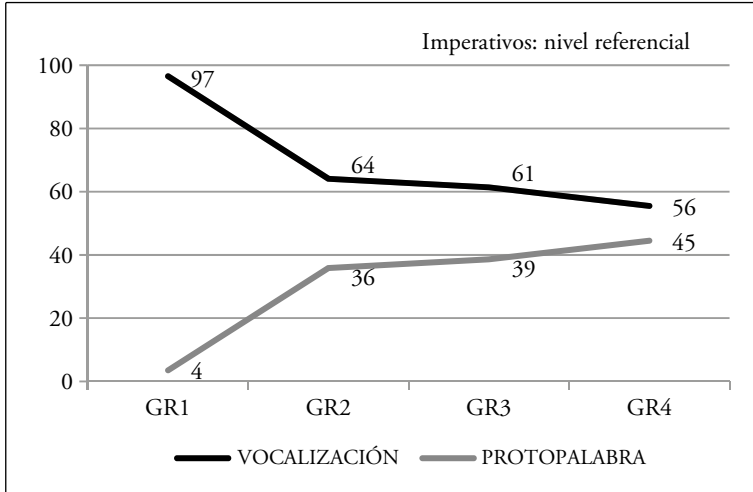


Figura 4. Desarrollo longitudinal del nivel referencial en los imperativos

En el caso del carácter interaccional, observamos los siguientes resultados para los declarativos: $\chi^2=7,990$, $GL=3$ y $p=0,046$, es decir, $p<0,05$, un resultado significativo. En cuanto a los imperativos, los resultados fueron, también, significativos: $\chi^2=12,636$, $GL=3$ y $p=0,005$, es decir, $p<0,05$.

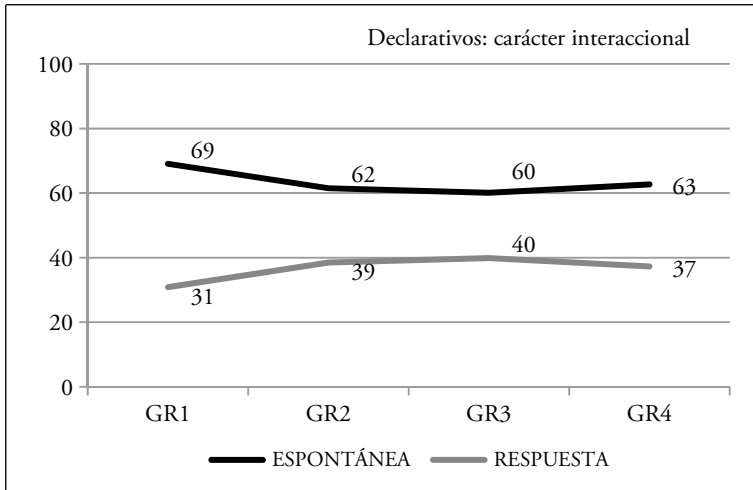
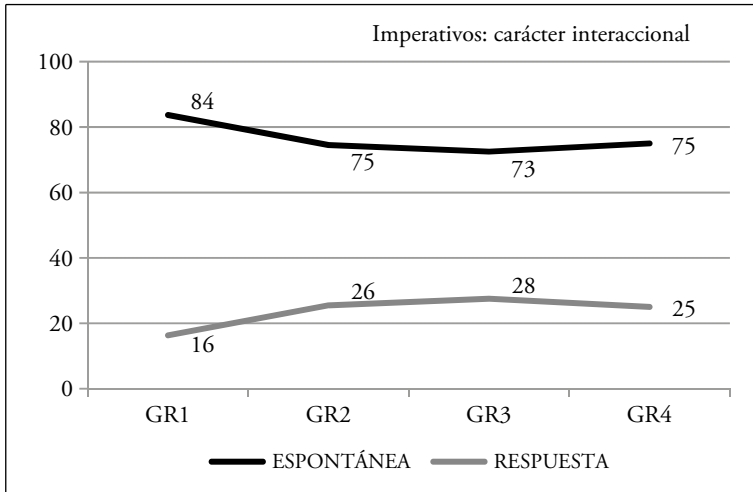
Figura 5. Desarrollo longitudinal del carácter interaccional en los declarativos

Figura 6. Desarrollo longitudinal del carácter interaccional en los imperativos



Se exploró la relación entre la edad y la orientación de la mirada en los declarativos. El análisis de chi-cuadrado arrojó un valor de $\chi^2=82,395$, $GL=9$ y $p<0,001$, es decir, un resultado significativo. En los imperativos, el análisis de chi-cuadrado arrojó un valor de $\chi^2=20,441$, $GL=9$ y $p=0,015$, es decir, $p<0,05$, un resultado significativo.

Figura 7. Desarrollo longitudinal de la orientación de la mirada en los declarativos

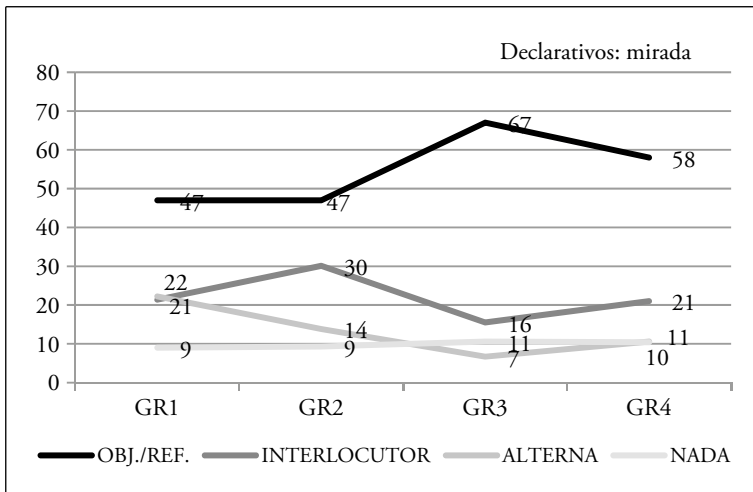
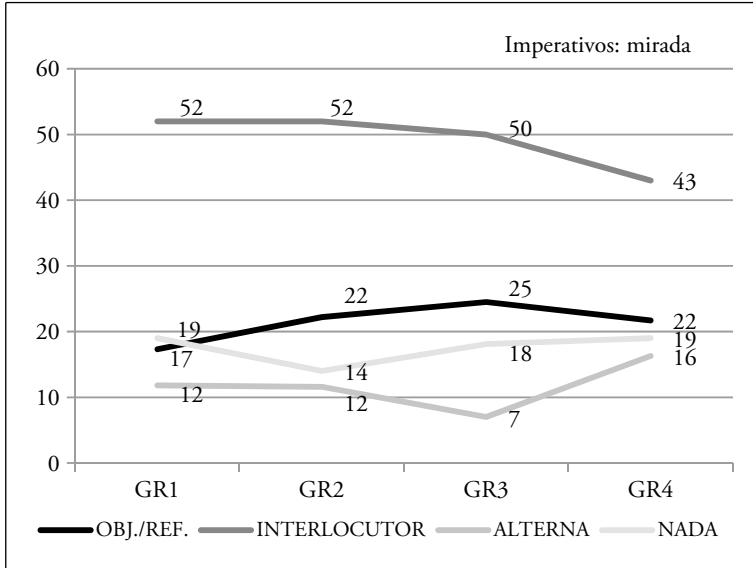


Figura 8. Desarrollo longitudinal de la orientación de la mirada en los imperativos



Exploramos la relación entre el grupo de edad y el contorno entonativo final de las conductas declarativas. El análisis de chi-cuadrado arrojó un valor de $\chi^2=68,316$, $GL=6$ y $p<0,001$, un resultado significativo. En los imperativos, el análisis de chi-cuadrado arrojó un valor de $\chi^2=28,936$, $GL=6$ y $p<0,001$, una vez más un resultado significativo.

Figura 9. Desarrollo longitudinal del contorno entonativo final en los declarativos

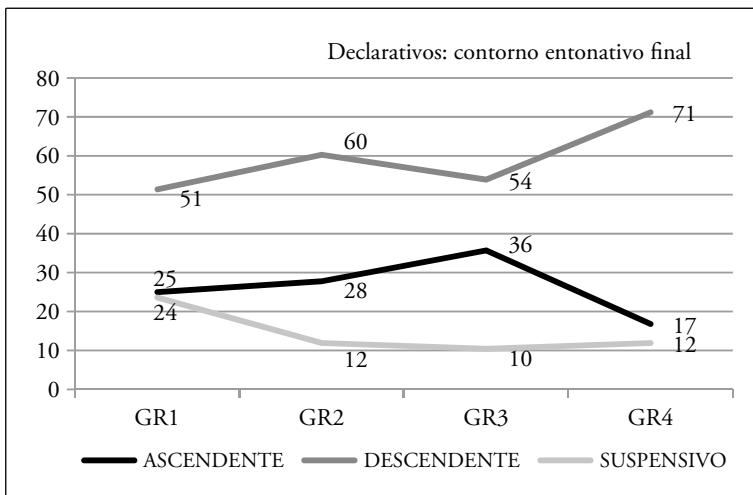
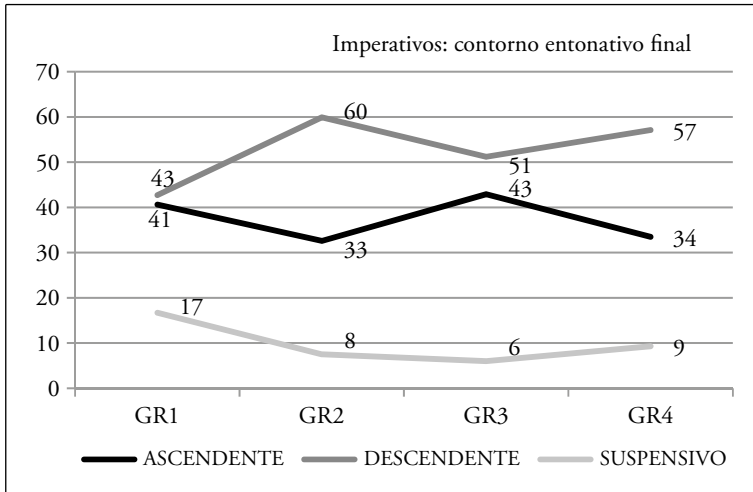


Figura 10. Desarrollo longitudinal del contorno entonativo final en los imperativos



Finalmente, se analizó la evolución longitudinal de la F_0 media con un análisis de Kruskal-Wallis, el cual arrojó los siguientes resultados en el caso de los declarativos: $\chi^2=71,77$, $GL=3$ y $p<0,001$, es decir, un resultado significativo. En los imperativos, los resultados fueron también significativos: $\chi^2=23,63$; $GL=3$ y $p<0,001$.

Figura 11. Desarrollo longitudinal de la F_0 media (en Hertz) en los declarativos

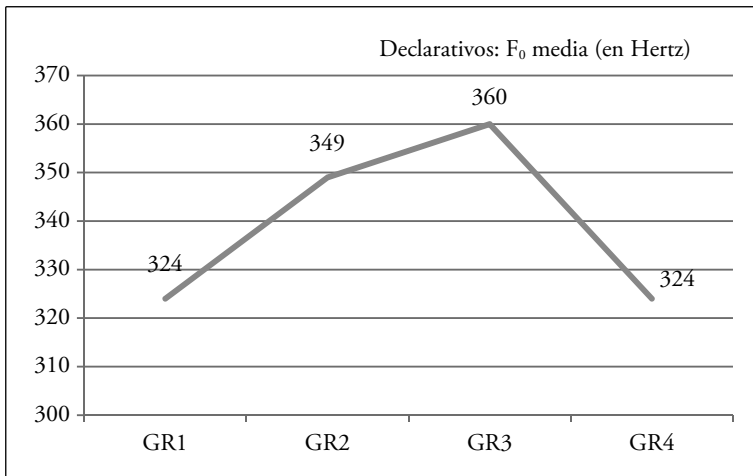
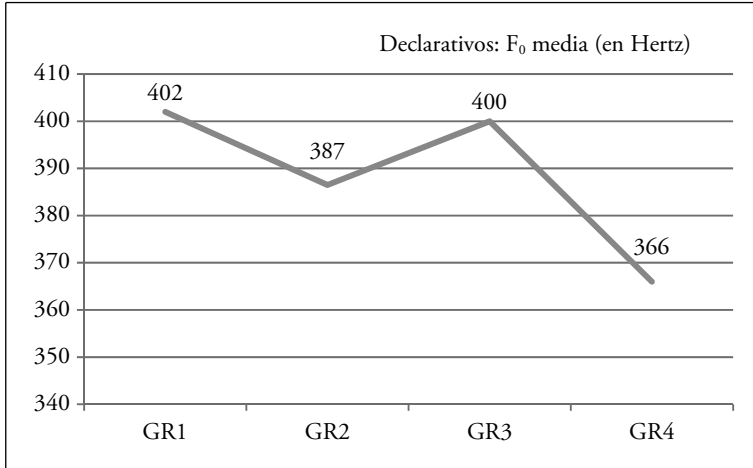


Figura 12. Desarrollo longitudinal de la F_0 media (en Hertz) en los imperativos

5. CONCLUSIONES

5.1. La comunicación en el segundo año de vida: caracterización de las funciones declarativa e imperativa

Desde el primer momento del periodo analizado (catorce-veinticuatro meses), el infante se muestra capaz de expresar intenciones declarativas e imperativas (o demandas): puede, pues, llamar la atención del interlocutor sobre determinado objeto, evento o persona y, también, solicitar o demandar de él un objeto, la realización de determinada acción, que le brinde cierta información o que le preste atención. Más aún, no solo logra transmitir dichas intenciones de forma efectiva —en el marco del contexto natural estudiado y con sus interlocutores habituales— sino que, además, manipula los distintos parámetros de su conducta (por ejemplo, el uso de la mirada, de un determinado contorno entonativo final, etcétera) de forma específica y con este fin.

La clara evidencia de comunicación intencional sistemática y eficaz en este periodo apoya lo propuesto por Dore (1974), Karousou (2003) y Tomasello (2008) con relación al carácter temprano y sofisticado de la expresión de intencionalidad comunicativa en infantes. Así mismo, la caracterización de las funciones comunicativas analizadas revela un panorama que cuestiona seriamente la visión de Jakobson (1968, p. 24) de la expresión «prelingüística» infantil o balbuceo como carente de carácter comunicativo o social; de acuerdo con dicho autor, la producción vocal temprana del niño constituiría una realidad desordenada e hipertrofiada que, aunque precede temporalmente al posterior desarrollo lingüístico, no guarda con este una relación continua sino, más bien, discontinua en muchos casos.

Nuestra clasificación recoge dos tipos de intenciones comunicativas, las declarativas e imperativas (o demandas). Nos alejamos en cierto sentido de lo propuesto por Bates y otros (1975), quienes prefieren las denominaciones de «protodeclarativos» y «protoimperativos». Con la omisión del prefijo «proto» buscamos enfatizar que las intenciones observables en la comunicación infantil son, en realidad, análogas a las del lenguaje adulto, aunque se diferencien de estas últimas por la forma —ciertamente más inmadura y probablemente más variable— a través de la cual se actualizan. Así, el aspecto «proto» se referiría más a la actualización o a la forma de la conducta (aspecto locutivo) que a la intención comunicativa (aspecto ilocutivo). Otra diferencia respecto de la clasificación de Bates y otros reside en la definición que presentan las autoras de los protoimperativos, dentro de los cuales consideran únicamente las demandas de objetos, mientras que aquí hemos considerado una intención como demanda o imperativa ya sea su objetivo un objeto, la realización de una acción, la obtención de la atención o de información a partir del interlocutor. Así, nuestra categoría resulta más abarcadora.

Ambas funciones tienen un perfil definido, caracterizable a partir de las variables consideradas —las cuales representarían distintas habilidades en desarrollo aún. Ello apuntaría a una correspondencia bastante temprana entre intención comunicativa (subyacente), por un lado, y conducta o habilidades por las cuales se expresa en el contexto del hogar (manifiestas), por otro. Así, la intención que motiva o guía al niño en su expresión comunicativa se traduce en un determinado patrón conductual no aleatorio sino complejo y coordinado, con lo que se irían estableciendo asociaciones de forma y función. Por ejemplo, una cierta función motivará con mayor probabilidad un tipo de mirada o de entonación.

Como resultado del análisis global de cada función comunicativa —abstrayendo el factor longitudinal—, observamos que tanto la función declarativa como la imperativa se expresan principalmente a través de la modalidad vocal; sin embargo, las demandas presentan un uso gestual ligeramente mayor (tanto considerando las conductas gestuales puras como las mixtas). Así, el niño en el periodo de una palabra parece inclinarse ya de forma notoria por la expresión vocal frente a la gestual. Ello parece hablar a favor de una sintonización cada vez mayor con los hábitos comunicativos del entorno, que son, justamente, lingüísticos y, en este caso, vocales (podría pensarse en casos distintos en que la lengua del entorno fuera una lengua de señas y, en dicho caso, se trataría de un medio visual o gestual, si se quiere).

Con respecto al nivel de desarrollo referencial o léxico, la función declarativa se revela como más madura que la imperativa. Observamos que la frecuencia de ocurrencia de protopalabras es mayor con función declarativa que demanda. Los niños son más proclives, en este periodo, a usar formas referencialmente más maduras

—o asociadas de forma más estable con un determinado contexto o situación de uso— en situaciones en que, por ejemplo, denominan una imagen o buscan llamar la atención del interlocutor sobre un determinado juguete que cuando, más bien, pretenden lograr algo más del adulto, como, por ejemplo, que les entregue algo o que responda a sus preguntas.

Si consideramos que, al margen de la función para la que se use, la protopalabra debe residir, de alguna forma, en la mente infantil para ser usada, tendríamos que preguntarnos por qué ocurren estas formas más frecuentemente con una intención que con otra. ¿No debería ser igualmente probable la ocurrencia de protopalabras con función declarativa e imperativa? Nuestros resultados sugieren una respuesta negativa. La razón habría que buscarla en el carácter aún relativamente inmaduro de estas formas o representaciones en el niño (asociaciones entre forma y sentido, o forma y uso, si se quiere), más susceptibles a las dificultades y al apoyo provistos por la situación comunicativa o el interlocutor.

Así, la diferencia que observamos podría deberse a la existencia de formatos de interacción promovidos por el cuidador en el marco del hogar que motivan por parte del niño la denominación de objetos del entorno o de imágenes en libros. Así, estas dinámicas podrían empujar el desarrollo léxico de forma importante en el contexto de emisiones con función declarativa. Lo hallado se encontraría en la línea del *Language Acquisition Support System* (LASS) propuesto por Bruner (1983): estos formatos de interacción modelados por los adultos para los infantes constituirían el escenario ideal para el desarrollo comunicativo y lingüístico infantil. Dichas dinámicas o rutinas motivarían la participación del niño en la medida de sus posibilidades, y se adaptarían a su progresivo nivel de desarrollo para brindarle un cierto nivel de predictibilidad que lo oriente en sus intervenciones.

En cuanto al carácter interaccional, los declarativos ocurren con mayor frecuencia que las demandas como respuestas a un enunciado adulto (una pregunta, sugerencia o consigna). Ello podría deberse a los ya mencionados formatos de interacción, por los cuales los adultos, a través de preguntas dirigidas al niño, promoverían la denominación más madura de determinados elementos. En el caso de los imperativos, en cambio, además de no ser tan promovidos a través de rutinas estructuradas como las anteriores, es menos probable que ocurran como respuesta a algo dicho por el adulto; más bien, lo usual es que nazcan de la voluntad del niño, sin mediación de un enunciado previo respecto del cual puedan construirse contingentemente.

Ambas funciones exhiben perfiles distintos en relación con la orientación de la mirada. Así, mientras en el caso de la intención declarativa el niño tiende a mirar al referente relacionado con la interacción en curso, en el caso de la demanda suele mirar, más bien, al interlocutor, quien es, justamente, el responsable de acceder o no

a lo solicitado. Este patrón más o menos distintivo de la mirada podría apuntar a una comprensión del otro como sujeto de atención y de intención. Por un lado, a través de los declarativos, el niño busca orientar la atención del otro hacia determinado objeto de su interés, con el objetivo de compartirlo con él; por otro lado, por medio de las demandas, la intención sería alinear la intención adulta con la propia, de modo que el interlocutor actúe en consecuencia. Estaría emergiendo en el niño una visión del otro como sujeto naturalmente cooperativo, susceptible de ser movido por sus enunciados a actuar de una u otra forma (Grice, 1975).

De forma marginal, notamos que la orientación de la mirada parece coordinarse de cierta forma con la modalidad de expresión. Así, en el caso de los declarativos, si la modalidad es vocal, el niño tiende a mirar al referente de la comunicación, pero si esta es gestual tiende a mirar al interlocutor. En el caso de los imperativos, si bien las miradas al interlocutor son siempre predominantes, en las conductas con modalidad gestual estas superan de forma aún más marcada a las miradas al referente (ver Anexo 1). Esta fina correlación entre modalidad de expresión y orientación de la mirada revelaría un manejo expresivo muy fino por parte del niño: mientras una emisión vocal puede ser percibida (y, por tanto, atendida o acatada) aunque el interlocutor no esté mirando al emisor, en el caso de un gesto el interlocutor debe mirar al emisor para que su intención comunicativa sea realmente comprendida. Así, el niño regula su mirada teniendo en cuenta la modalidad de su expresión y, además, su comprensión de que la capacidad del otro para atender su mensaje depende de cómo lo formule en todas sus dimensiones (Fernández Flecha, 2012).

También, la prosodia permite caracterizar a ambas funciones. En ambas, el contorno entonativo final más frecuente es el descendente, lo que resulta coherente con lo propuesto por Lieberman (1967), de acuerdo con quien dicho patrón entonativo constituiría el modo más natural o fisiológico de la frecuencia fundamental debido a la caída natural en la presión subglotal en el curso de la enunciación (y, así, de la espiración). Este patrón casi «por defecto» de la entonación ocurriría, pues, de forma más natural y, por ello, sería el más frecuente en las etapas tempranas de la comunicación y el lenguaje, en las que el niño no habría ganado del todo control intencional o funcional sobre su producción prosódica. De acuerdo con Lieberman, pues, mientras los contornos descendentes se producirían de forma natural, los ascendentes demandarían un mayor control intencional y serían, por ello, desarrollos posteriores. En contraste con el planteamiento de Lieberman, Snow (2004, 2006 y 2007) propone que tanto los contornos ascendentes como los descendentes serían adquiridos en patrones de desarrollo no lineales casi idénticos en el tiempo.

Esta investigación confirma, para el caso del español, el predominio de los contornos descendentes en la comunicación infantil. El contorno entonativo final

descendente ha sido tradicionalmente asociado con la intención declarativa, con enunciados con sentido completo en el lenguaje adulto (Quilis, 1993) y, también, en la comunicación infantil (Marcos, 1987). Observamos, sin embargo, que los imperativos aquí analizados exhiben un predominio relativo del contorno final ascendente frente a los declarativos, lo que confirma hallazgos previos (Marcos, 1987; Flax y otros, 1991; Fernández Flecha, 2009). Esto podría deberse a la carga interrelativa propia de las demandas (superior a la de los declarativos), en tanto buscan mover al interlocutor a hacer algo de forma más directa.

Finalmente, con respecto a la frecuencia fundamental, los declarativos presentan una F0 más baja que los imperativos, lo cual podría deberse, nuevamente, a su carácter eminentemente más interrelativo que el de los declarativos.

5.2. Desarrollo comunicativo longitudinal: evolución de las funciones declarativa e imperativa

Las distintas habilidades consideradas no progresan con el mismo ritmo de avance ni con el mismo patrón de desarrollo; más bien, cada una parece trazar un camino particular en el marco de cada función. Así mismo, observamos una relación de continuidad entre los momentos menos maduros y los más maduros. Conductas menos y más maduras coexisten en la misma etapa, en un proceso por el cual las versiones más avanzadas van ganando terreno frente a las más bien inmaduras.

En primer lugar, en el segundo año de vida, ambas funciones ven aumentar sus producciones de carácter vocal y reducirse aquellas que presentan un componente gestual (ya sean de modalidad mixta o gestual pura). El desarrollo de la comunicación parece entrañar, pues, un proceso de desgestualización progresiva y un mayor papel de la expresión vocal (coherente con el modelo adulto, que prioriza la expresión lingüística, vocal). Estos resultados apoyan lo planteado en Volterra y otros (2005), de acuerdo con quienes en el segundo año de vida (específicamente hacia los veinte meses) los niños empiezan a priorizar la modalidad vocal o verbal significativamente más que la gestual o la combinación de ambas en la comunicación. En aquel estudio, se propuso que la reducción de las conductas gestuales se relaciona con el aumento del repertorio léxico infantil, para lo que también encontramos apoyo en la presente investigación.

Es claro que en ambas funciones la proporción de protopalabras aumenta de forma sostenida. Además, tanto en los declarativos como en los imperativos, el gran aumento se da entre los dieciséis y dieciocho meses, momento que ha sido tradicionalmente asociado con la conocida «explosión del vocabulario». La reducción en el uso de gestos se relacionaría con este incremento en el uso de producciones referencialmente más maduras que las vocalizaciones. Si se asume que, en momentos más tempranos de la comunicación infantil, los gestos cumplen una función de apoyo

de la expresión vocal —la que, por ser aún inmadura y poco convencional, requeriría el acompañamiento de un gesto que, por ejemplo, precise el objeto al que se refiere o aclare la intención comunicativa del niño—, resulta lógico que, una vez que la expresión vocal deviene más estable y, por ende, clara, la coocurrencia de un gesto devenga más prescindible.

También entre los dieciséis y dieciocho meses presenciamos, en ambas funciones, un considerable incremento de las conductas de tipo respuesta. Hemos aludido ya al importante papel de las dinámicas de interacción entre cuidador e infante en el aprendizaje de palabras; debemos ahora añadir su posible papel como espacios de entrenamiento dialógico o conversacional, por medio del cual el infante se adiestra en cuestiones como el intercambio de turnos o la cooperación conversacional. En el marco de estas dinámicas, el niño sería capaz de producir respuestas a los enunciados del adulto. En este sentido, la especial afinidad de dichos formatos de interacción con la función declarativa —sumada al carácter posiblemente más espontáneo de las demandas frente a los declarativos— explicaría que los declarativos presenten mayor frecuencia de respuestas que las demandas.

En el caso de la orientación de la mirada, no emerge un patrón claro de desarrollo. Sí observamos que, en los declarativos, las miradas dirigidas al referente aumentan. En los imperativos, también se observa un aumento del mismo tipo de miradas aunque ciertamente menos marcado.

Finalmente, con respecto a los aspectos prosódicos, observamos, por un lado, que, en los declarativos, aumentan los contornos entonativos finales descendentes, y se reducen los ascendentes y suspensivos. Sin embargo, no podemos obviar que esto se da por medio de un patrón de desarrollo irregular, caracterizado por aumentos y descensos. En el caso de la F_0 media, esta aumenta para luego caer en el periodo final hasta retornar a los valores iniciales. Por otro lado, en los imperativos, observamos, también, que los descendentes aumentan, y los ascendentes y suspensivos se reducen, nuevamente en un patrón de tipo irregular. Sin embargo, la ocurrencia de contornos finales ascendentes al final del periodo considerado es mayor que la observada en el momento final para los declarativos (aproximadamente el doble). En cuanto a la F_0 media, esta se reduce pero luego de un proceso irregular.

Los resultados parecen apoyar la propuesta de Snow (2006): la producción entonativa atravesaría un proceso de reajuste, regresión y reorganización que se dispararía con la aparición de la comunicación intencional y la producción de las primeras palabras, aproximadamente hacia los doce meses. En ese punto, la producción prosódica infantil, que hasta ese momento habría tenido bases puramente fisiológicas, entraría en un periodo de ajuste que desembocaría, eventualmente, hacia el final del periodo de una palabra (dieciocho-veinticuatro meses), en una producción

prosódica de base funcional o intencional. De este modo, los inicios del desarrollo lingüístico desestabilizarían la producción prosódica del infante, para ir aproximándola progresivamente a la producción de tipo adulto.

En esta misma línea, además, el aumento de la producción de contornos finales descendentes en el caso de la función declarativa podría entenderse como el acercamiento del infante al modelo adulto que, según Quilis (1993), establece una asociación entre los contornos finales descendentes y los enunciados con sentido completo (como los declarativos, por ejemplo).

En conclusión, observamos una relación de continuidad entre los momentos más inmaduros y los más maduros del desarrollo comunicativo y lingüístico. Así mismo, merece la pena destacar la coexistencia de distintos niveles de madurez en el mismo periodo, lo que contrasta con el límite entre lo «prelingüístico» y lo «lingüístico» (tradicionalmente marcado por el aprendizaje de las primeras palabras) propuesto por Jakobson (1968); de acuerdo con dicho autor, no se observaría un patrón de desarrollo claro antes de la aparición de las palabras tempranas, sino que se observaría una relación más bien discontinua. Los presentes hallazgos apoyan la hipótesis de la continuidad, propuesta por Vihman y otros (1985) —quienes defienden la coexistencia evolutiva de palabras y no palabras (vocalizaciones), con patrones de desarrollo semejantes—, y López Ornat y Karousou (2005) —quienes destacan la continuidad evolutiva entre vocalizaciones tempranas y producciones vocales más convencionales o propiamente lingüísticas—.

En el marco de dicha relación de continuidad, la comunicación se desarrollaría a través de un proceso de emergencia progresiva, por el cual devendría cada vez más convencional, más cercana al modelo adulto. La producción comunicativa madura no emergería de forma repentina, con todas las habilidades que implica igualmente maduras a la vez; más bien, las distintas habilidades o componentes implicados atravesarían cada uno un proceso evolutivo propio (con un ritmo y un patrón específico). Este desarrollo de la conducta comunicativa por componentes ha sido ya observado por Karousou (2003). La tarea del niño consistiría en ir regulando su avance en las distintas habilidades necesarias para comunicarse y aprender su lengua; cada emisión supondría el manejo de los recursos disponibles en una edad determinada con el objetivo de expresar una intención específica de la forma más eficaz posible.

6. DISCUSIÓN

Es posible caracterizar de forma clara las funciones declarativa e imperativa en la comunicación temprana. Por un lado, la función declarativa tiene, típicamente, modalidad vocal, una mirada orientada hacia el referente a la comunicación, un nivel referencial variable (más maduro que el de las demandas), un contorno entonativo

final descendente y una frecuencia fundamental media inferior a la hallada en los imperativos. Por otro lado, estos últimos exhiben, generalmente, modalidad vocal, una mirada típicamente orientada hacia el interlocutor del intercambio comunicativo, un nivel referencial variable (menos maduro que el de los declarativos), un contorno entonativo final descendente —aunque con una frecuencia relativamente mayor de ascendentes— y una frecuencia fundamental media superior a la de los declarativos. Adicionalmente, merece la pena destacar una correspondencia más o menos general entre la modalidad de la conducta y la orientación de la mirada: el uso de gestos motivaría por parte del niño una mirada orientada al interlocutor, mientras que una expresión puramente vocal favorecería una mirada dirigida al referente del intercambio, lo que acusa una fina comprensión de los requisitos que deben cumplirse para que la comunicación resulte exitosa.

En el periodo analizado (comúnmente referido como periodo de una palabra), los gestos tienden a ocurrir de forma simultánea con una emisión vocal, en el marco de conductas mixtas (y no tanto como conductas puramente gestuales), lo que refuerza el predominio de la comunicación vocal. La producción gestual parecería servir como apoyo o refuerzo de lo vocal para transmitir una determinada intención. Ello podría ser interpretado como evidencia contra el supuesto predominio gestual en la comunicación temprana —como propone, por ejemplo, Tomasello (1999 y 2008)—, por lo menos en el periodo aquí estudiado.

El aparente retroceso de los gestos apunta a que devendrían cada vez más prescindibles en su papel de apoyo comunicativo conforme el niño progresa en su desarrollo lingüístico y sus emisiones vocales se vuelvan, entonces, más autónomas y autosuficientes. Esto apoya la propuesta de autores como McEachern y Haynes (2004), Volterra y otros (2005) y Crais y otros (2009). Sin embargo, aún es posible pensar que, en etapas previas a la aquí estudiada, ciertas estructuras performativas ocurran fundamentalmente de forma gestual antes de ser transmitidas por medios lingüísticos (Bates y otros, 1975; Carpenter, 2009).

El aprendizaje de palabras sería impulsado por medio de rutinas o formatos de interacción por parte de los padres (al estilo propuesto por Bruner, 1983 y Tomasello, 2001 y 2008). En dichas dinámicas, el adulto participa como moderador del intercambio, estableciendo una especie de plantilla con turnos para la participación de ambos interlocutores, regulando el grado de complejidad del estímulo brindado al niño y de la intervención esperada de él dependiendo del nivel de desarrollo que tenga.

Tanto en el caso de los declarativos como de los imperativos, el principal aumento en la producción de protopalabras se da entre los dieciséis y dieciocho meses, de lo que podemos inferir que estaría teniendo lugar el fenómeno de «explosión

del vocabulario», fenómeno que parecería impulsar —o, por lo menos, acompañar— otros procesos de avance comunicativo como la producción de respuestas. Esta mayor habilidad conversacional se debería, por un lado, a la capacidad de producir más enunciados como respuesta a un enunciado proferido por el adulto y, por otro, a una conciencia más profunda de la importancia de ser cooperativo en la comunicación —lo que supone responder a lo preguntado, sugerido u ordenado por el adulto—, en concordancia con el principio de cooperación de Grice (1975).

Lo observado en relación con el uso de la mirada como herramienta para la comunicación —piénsese en el predominio de la mirada orientada al referente en los declarativos y al interlocutor en las demandas— parece apuntar a una comprensión del interlocutor como sujeto atencional e intencional, cuya atención e intención son percibidas por el niño como potencialmente influenciables (Tomasello, 2008). Su destreza comunicativa se evidencia en las habilidades para llevar a cabo esta tarea con éxito.

Con respecto a los aspectos prosódicos considerados, el contorno entonativo final descendente es, sin duda, el más común. Según parte de la bibliografía sobre el tema, dicho patrón sería el más temprano y recurrente en las primeras etapas de la adquisición del lenguaje por tratarse del tono más natural y fisiológico (Lieberman, 1967). Por otro lado, el uso de los contornos finales ascendentes se revela relativamente mayor en la función imperativa, lo que podría explicarse si asumimos, más bien, que, a partir de cierto punto en el desarrollo, todos los contornos empiezan a estar bajo control voluntario (Snow, 2006). Esta asociación entre juntura terminal ascendente y función imperativa ha sido notada ya por otros autores para el caso de lenguas como el inglés (Marcos, 1987; Flax y otros, 1991).

La producción prosódica parece atravesar un proceso de ajuste, exploración y reorganización, caracterizado por la presencia de sucesivos aumentos y descensos en la ocurrencia de los distintos contornos entonativos. Esto resulta coherente con la teoría de la reorganización prosódica (Snow, 2006): la entonación se hallaría inicialmente bajo control fisiológico o involuntario pero, luego, pasaría a estar bajo control intencional y adquiriría bases funcionales. La exploración, la reorganización y el consiguiente ajuste de las modulaciones prosódicas serían motivados por la adquisición de las primeras palabras (Snow & Balog, 2002; DePaolis y otros, 2008) y se estabilizaría recién hacia el final del periodo de una palabra —en el caso del niño aquí analizado, hacia los veinticuatro meses, cuando empiezan a aparecer las primeras combinaciones de dos palabras—. Esto se apoyaría también en las variaciones longitudinales halladas en la producción de la F_0 media.

En suma, en el periodo estudiado parecen tener lugar distintos procesos generales: (1) un aumento de la producción vocal en detrimento de la gestual; (2) de forma

probablemente relacionada, un aumento progresivo de la producción de protopalabras, producciones léxicas más convencionales que las vocalizaciones; (3) probablemente debido al aumento del vocabulario, un aumento en la producción de respuestas a enunciados adultos, lo que apuntaría a una mayor comprensión de la dinámica conversacional como actividad cooperativa y, a la vez, mayores recursos para ocupar el turno conversacional que le es impuesto por la intervención adulta.

Por otro lado, al interior de cada función comunicativa, se dan relaciones, evoluciones y matices propios que se perderían de considerarse la conducta comunicativa infantil como un todo. Ello refuerza la importancia de estudiar el desarrollo lingüístico temprano desde una perspectiva funcional —es decir, considerando cada función o intención comunicativa como un ámbito del desarrollo— y, además, teniendo especialmente en cuenta que las distintas habilidades (léxica, prosódica, por ejemplo) pueden tener ritmos distintos de avance, estableciendo incluso relaciones entre sí a partir de cierto momento.

En el periodo analizado, el niño se acerca cada vez más a la comunicación lingüística convencional. Es capaz de sostener una conversación ocupando el turno que le corresponde y de transmitir sus intenciones comunicativas casi siempre de forma efectiva a un interlocutor, a pesar de contar con recursos lingüísticos convencionales aún limitados. En este sentido, cobran especial importancia la orientación que da a su mirada y la modulación que hace de los aspectos prosódicos de su emisión, los cuales lo ayudan a dar a entender aquello que busca transmitir. Cada conducta comunicativa producida por el niño no se constituye como una mezcla sin sentido o lógica de los diversos elementos involucrados, sino que revela una fina trama por la cual mirada, modalidad de la conducta, carácter interaccional, entonación, etcétera se conjugan en el intento de transmitir una intención específica, la cual logra casi siempre ser adecuadamente descifrada por el adulto interlocutor.

Finalmente, en el marco de la continuidad observada, la aparición de las primeras protopalabras no implica la desaparición automática de conductas comunicativas más inmaduras como las vocalizaciones o el balbuceo «prelingüístico» (Vihman y otros, 1985, López Ornat & Karousou, 2005), sino que ambos tipos de emisiones coexisten en el marco de ambas funciones comunicativas. El desarrollo comunicativo precedería y prepararía el terreno para el desarrollo más específicamente lingüístico, con el cual coexistiría a partir de cierto punto. El primero evolucionaría por medios diversos —gestuales, conductuales como la mirada, prosódicos, etcétera— y, en un punto del proceso, las conductas vocales devendrían reconocibles como propias y convencionales de la lengua del entorno. Entonces la comunicación se habrá descontextualizado al ser cada vez más lingüística, más convencional. En ese punto, el niño será no solo un comunicador eficaz sino un usuario competente de su lengua.

Anexo 1

Tabla 11. Relación entre modalidad y mirada en declarativos

	Obj./Ref.	Inter.	Alter.	Nada	Total
Vocal	60%	20%	9%	11%	100%
Gesto	32%	45%	18%	5%	100%
Mixto	47%	23%	24%	6%	100%

Tabla 12. Relación entre modalidad y mirada en imperativos

	Obj./Ref.	Inter.	Alter.	Nada	Total
Vocal	19%	52%	6%	23%	100%
Gesto	10%	63%	11%	16%	100%
Mixto	33%	42%	22%	3%	100%

REFERENCIAS BIBLIOGRÁFICAS

- Austin, John Langshaw (1962). *How to Do Things with Words*. Oxford: Oxford University Press.
- Bates, Elizabeth, Luigia Camaioni & Virginia Volterra (1975). The Acquisition of Performatives Prior to Speech. En Elinor Ochs y Bambi Schieffelin (eds.), *Developmental Pragmatics* (pp. 111-126). Nueva York: Academic Press.
- Bates, Elizabeth (1976). *Language and Context. The Acquisition of Pragmatics*. Nueva York: Academic Press.
- Boersma, Paul & David Weenink (s/a). *Praat: Doing Phonetics by Computer* [programa de ordenador]. <http://www.praat.org/>
- Bruner, Jerome (1973). Organization of Early Skilled Action. *Child Development*, 44(1), 1-11.
- Bruner, Jerome (1974/5). From Communication to Language. A Psychological Perspective. *Cognition*, 3(3), 225-287.
- Bruner, Jerome (1983). *Child's Talk. Learning to Use Language*. Nueva York-Londres: W. W. Norton & Company.
- Carpenter, Malinda (2009). Just How Joint is Joint Action in Infancy? *Topics in Cognitive Science*, 1(2), 380-392.

- Crais, Elizabeth, Linda Watson & Grace Baranek (2009). Use of Gesture Development in Profiling Children's Prelinguistic Communication Skills. *American Journal of Speech-Language Pathology*, 18, 95-108.
- Davidson, Donald (2001). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- DePaolis, Rory, Marilyn May Vihman & Sari Kunnari (2008). Prosody in the Production at the Onset of Word Use: A Cross-linguistic Study. *Journal of Phonetics*, 36(2), 406-422.
- D'Oodorico, Laura & Fabia Franco (1991). Selective Production of Vocalization Types in Different Communication Contexts. *Journal of Child Language*, 18(3), 475-499.
- Dore, John (1974). A Pragmatic Description of Early Language Development. *Journal of Psycholinguistic Research*, 3(4), 343-350.
- Dore, John (1975). Holophrases, speech acts and language universals. *Journal of Child Language*, 2(1), 21-40.
- Ellis, Nick (1998). Emergentism, Connectionism and Language Learning. *Language Learning*, 48(4), 631-664.
- Elman, Jeffrey y otros (1998). *Rethinking Innateness. A Connectionist Perspective on Development*. Cambridge-Londres: The MIT Press.
- Fernández Flecha, María de los Ángeles (2009). *Prelinguistic Vocalizations. Relations between Function and Pitch*. En AcquisiLyon 09, Colloque Jeunes Chercheurs en Acquisition du Langage (pp. 1-4). Lyon: publicación digital. <http://elisa.ddl.ish-lyon.cnrs.fr/Colloques/AcquisiLyon/pageweb/Fichier/Fernandez.pdf>
- Fernández Flecha, María de los Ángeles (2012). *Evolución funcional de la conducta comunicativa infantil. Estudio de caso de un niño castellano-hablante*. Tesis doctoral, Universidad Complutense de Madrid.
- Flax, Judy, Margaret Lahey, Katherine Harris & Arthur Boothroyd (1991). Relations between Prosodic Variables and Communicative Functions. *Journal of Child Language*, 18(1), 3-19.
- Franco, Fabia & George Butterworth (1996). Pointing and Social Awareness: Declaring and Requesting in the Second Year. *Journal of Child Language*, 23(2), 307-337.
- Gibb Harding, Carol & Roberta Michnick Golinkoff (1979). The Origins of Intentional Vocalizations in Prelinguistic Infants. *Child Development*, 50(1), 33-40.
- Goldin-Meadow, Susan (2009). From Gesture to Word. En Edith Bavin (ed.), *Cambridge Handbook of Child Language* (pp. 145-160). Nueva York: Cambridge University Press.
- Grice, Herbert Paul (1975). Logic and Conversation. En Peter Cole y Jerry L. Morgan (eds.), *Syntax and Semantics: Speech Acts* (pp. 41-58). Volumen 3. Nueva York: Academic.

- Hallé, Pierre A., Bénédicte de Boysson-Bardies & Marilyn May Vihman (1991). Beginnings of Prosodic Organization: Intonation and Duration Patterns of Disyllables Produced by Japanese and French Infants. *Language and Speech*, 34(4), 299-318.
- Halliday, Michael Alexander Kirkwood (1975). Estructura y función del lenguaje. En John Lyons (ed.). *Nuevos horizontes de la lingüística* (pp. 145-173). Madrid: Alianza.
- Halliday, Michael Alexander Kirkwood (1982). Aprendiendo a conferir significado. En Eric H. Lenneberg y Elizabeth Lenneberg (eds.), *Fundamentos del desarrollo del lenguaje* (pp. 239-267). Madrid: Alianza Universidad Textos.
- Jakobson, Roman (1968). *Child Language, Aphasia and Phonological Universals*. La Haya: Mouton.
- Karmiloff, Kyra & Annette Karmiloff-Smith (2005). *Hacia el lenguaje. Del feto al adolescente*. Madrid: Morata.
- Karousou, Alexandra (2003). *Análisis de las vocalizaciones tempranas: su patrón evolutivo y su función determinante en la emergencia de la palabra*. Tesis doctoral, Universidad Complutense de Madrid.
- Kuhl, Patricia (2004). Early Language Acquisition: Cracking the Speech Code. *Nature Reviews Neuroscience*, 5, 831-843.
- Lenneberg, Eric H. (1975). *Fundamentos biológicos del lenguaje*. Madrid: Alianza.
- Lenneberg, Eric H. & Elizabeth Lenneberg (1982). *Fundamentos del desarrollo del lenguaje*. Madrid: Alianza.
- Leroy, Marie, Emmanuelle Mathiot & Aliyah Morgenstern (2009). Pointing Gestures, Vocalizations and Gaze: Two Case Studies. En Jordan Zlatev, Marlene Johansson Falck, Carita Lundmark y Mats Andréén (eds.), *Studies in Language and Cognition* (pp. 261-275). Nueva York: Cambridge Scholars Publishing.
- Lieberman, Philip (1967). *Intonation, Perception and Language*. Cambridge: The MIT Press.
- Lieberman, Philip (1984). *The Biology and Evolution of Language*. Cambridge: Harvard University Press.
- Liszkowski, Ulf (2008). Before L1. A Differentiated Perspective on Infant Gestures. *Gesture*, 8(2), 180-196.
- Liszkowski, Ulf, Konstanze Albrecht, Melinda Carpenter & Michael Tomasello (2008). Infants' Visual and Auditory Communication When a Partner Is or Is Not Visually Attending. *Infant Behavior & Development*, 31(2), 157-167.
- López Ornat, Susana, Almudena Fernández, Pilar Gallo & Sonia Mariscal (1994). *La adquisición de la lengua española*. Madrid: Siglo Veintiuno de España.

- López-Ornat, Susana & Alexandra Karousou (2005). Las vocalizaciones tempranas (8-30 meses) y su relación con el vocabulario y la gramática. Su medida en el «CDI español»: resultados preliminares. En Ma. Ángeles Mayor Cinca, Begoña Zubiauz de Pedro y Emiliano Díez Villoria (eds.), *Estudios sobre la adquisición del lenguaje* (pp. 401-420). Salamanca: Universidad de Salamanca.
- MacWhinney, Brian (2005). The Emergence of Linguistic Form in Time. *Connection Science*, 17(3-4), 191-211.
- Mampe, Birgit y otros (2009). Newborns' Cry Melody Is Shaped by Their Native Language. *Current Biology*, 19(23), 1-4.
- Marcos, Haydee (1987). Communicative Functions of Pitch Range and Pitch Direction in Infants. *Journal of Child Language*, 14(2), 255-268.
- Mariscal, Sonia y otros (2007). La evaluación del desarrollo comunicativo y lingüístico mediante la versión española de los inventarios Mac Arthur-Bates. *Psicothema*, 19(2), 190-197.
- McEachern, Diane & William O. Haynes (2004). Gesture-Speech Combinations as a Transition to Multiword Utterances. *Journal of Speech-Language Pathology*, 13(3), 227-235.
- Nathani, Suneeti & D. Kimbrough Oller (2001). Beyond ba-ba and gu-gu: Challenges and Strategies in Coding Infant Vocalizations. *Behavior Research Methods, Instruments, and Computers*, 33(3), 321-330.
- Ninio, Anato & Catherine Snow (1988). Language Acquisition Through Language Use: The Functional Sources of Children's Early Utterances. En Yonata Levy, Izchak M. Schlesinger y Martin D. S. Braine (eds.), *Categories and Processes in Language Acquisition* (pp. 11-30). Nueva Jersey: Erlbaum.
- Ninio, Anato & Catherine Snow (1999). The Development of Pragmatics: Learning to Use Language Appropriately. En Tej K. Bhatia y William C. Ritchie (eds.), *Handbook of language acquisition* (pp. 347-383). Nueva York: Academic Press.
- Quilis, Antonio (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Searle, John Rogers (1965). What Is a Speech Act? En Max Black (ed.), *Philosophy in America* (pp. 221-239). Londres: Allen & Unwin.
- Searle, John Rogers (1969). *Speech Acts: An Essay in the Philosophy of Language*. Londres: Cambridge University Press.
- Snow, David (1998). Children's Imitations of Intonation Contours: Are Rising Tones more Difficult than Falling Tones? *Journal of Speech, Language, and Hearing Research*, 41(3), 576-587.

- Snow, David (2004). Falling Intonation in the One- and Two-Syllable Utterances of Infants and Preschoolers. *Journal of Phonetics*, 32(3), 373-393.
- Snow, David (2006). Regression and Reorganization of Intonation Between 6 and 23 Months. *Child Development*, 77(2), 281-296.
- Snow, David (2007). Polysyllabic Units in the Vocalizations of Children From 0;6 to 1;11: Intonation-groups, Tones and Rhythms. *Journal of Child Language*, 34(4), 765-797.
- Snow, David & Heather L. Balog (2002). Do Children Produce the Melody before the Words? A Review of Developmental Intonation Research. *Lingua*, 112(12), 1025-1058.
- Tomasello, Michael (1999). *The Cultural Origins of Human Cognition*. Cambridge-Londres: Harvard University Press.
- Tomasello, Michael (2001). Bruner on Language Acquisition. En David Backhurst y Stuart Shanker (eds.), *Jerome Bruner: Language, Culture and Self* (pp. 31-48). Londres: GBR-SAGE.
- Tomasello, Michael (2003). *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge-Londres: Harvard University Press.
- Tomasello, Michael (2008). *Origins of Human Communication*. Londres-Cambridge: The MIT Press.
- Tomasello, Michael (2009). The Usage-based Theory of Language Acquisition. En Edith Bavin (ed.), *Cambridge Handbook of Child Language* (pp. 69-87). Nueva York: Cambridge University Press.
- Tomasello, Michael & Elizabeth Bates (2001). *Language Development. The Essential Readings*. Oxford: Blackwell.
- Vihman, Marilyn May (1994). When is a word a word? *Journal of Child Language*, 21(3), 517-542.
- Vihman, Marilyn May (1996). *Phonological Development. The Origins of Language in the Child*. Oxford: Blackwell.
- Vihman, Marilyn May y otros (1985). From Babbling to Speech: A Re-assessment of the Continuity Issue. *Language*, 61(2), 397-445.
- Volterra, Virginia, Maria Cristina Caselli, Olga Capirci & Elena Pizzuto (2005). Gesture and the Emergence and Development of Language. En Michael Tomasello & Dan Slobin (eds.), *Beyond Nature-Nurture. Essays in Honor of Elizabeth Bates*. Nueva Jersey: Lawrence Erlbaum Associates.
- Yule, George (2008). *El lenguaje*. Madrid: Akal.

EL CARÁCTER MENTAL DE LOS SISTEMAS FONOLÓGICOS

Jorge Iván Pérez Silva

Pontificia Universidad Católica del Perú

En su relato «Funes el memorioso», Jorge Luis Borges concibe un personaje dueño de una percepción y una memoria infalibles. He aquí algunos ejemplos de lo que Ireneo Funes era capaz de hacer: «Nosotros, de un vistazo» —dice Borges— «percibimos tres copas en una mesa; Funes, todos los vástagos y racimos y frutos que comprende una parra»¹. La *simplicidad* de la imagen visual de tres copas contrasta notoriamente con la *complejidad*, con el abarrotamiento, de una parra. Para nosotros, con solo darle un vistazo, es casi imposible decir cuántas uvas hay inclusive en un racimo; Funes, en cambio, captaba nítidamente los innumerables individuos que componen toda una parra. Aquí otro ejemplo de la extraordinaria percepción de Funes: «Una circunferencia en un pizarrón, un triángulo rectángulo, un rombo, son formas que podemos intuir plenamente; lo mismo le pasaba a Ireneo con las aborascadas crines de un potro, con una punta de ganado en una cuchilla, con el fuego cambiante y con la innumerable ceniza». Comparemos una vez más nuestra percepción plena de una figura geométrica *simple* dibujada en un pizarrón, con la percepción que tenía Funes de una realidad *compleja*, abigarrada de detalles, como las crines de un caballo. Quizá las personas miopes podemos imaginar cómo percibía Ireneo Funes: al sacarnos los anteojos, nuestra imagen del mundo se torna borrosa; las crines de un caballo son indistinguibles y lo mismo ocurre con las hojas de los árboles; al ponernos los anteojos, nuestra imagen del mundo pierde su carácter de pintura impresionista —digamos— para adquirir algo de nitidez; yo imagino la visión de Funes como la que nos darían unos anteojos potentísimos, que no dejaran pasar desapercibido ningún detalle de los objetos, como nos ocurre a simple vista con una circunferencia o un triángulo en un pizarrón.

¹ Todas las citas de «Funes el memorioso» son de Borges (1983). No indico la página de cada cita para facilitar la lectura. El lector interesado podrá encontrar fácilmente en cualquier edición los fragmentos citados dada la brevedad del relato.

Lo que busca Borges con esta ficción es invitar al lector a imaginar cómo sería su relación cognitiva con el mundo si poseyera las cualidades ilimitadas de Funes o, lo que es lo mismo, nos lleva a reflexionar acerca de la manera particularmente limitada en que nuestra memoria y nuestros sentidos finitos nos permiten conocer el mundo. Uno de los cuestionamientos más interesantes que motiva esta ficción es el de nuestras nociones de *individuo* y *clase*. Cito a Borges: «Funes no solo recordaba cada hoja de cada árbol, de cada monte, sino cada una de las veces que la había percibido o imaginado». Además, como su percepción era infalible, cada recuerdo era pleno de detalles, de manera que la siguiente vez que veía lo que para nosotros sería la *misma* hoja, a él le parecía un objeto completamente distinto: más grande o más pequeño, más arrugado o menos brillante, con límites o nervaduras diferentes, etcétera. Más o menos, probablemente, como lo que nos ocurriría a nosotros si viéramos a un niño de un año y no volviéramos a verlo hasta setenta años después: nos resultaría difícil concebirlo como el *mismo* individuo. Dice Borges sobre Funes: «No solo le costaba comprender que el símbolo genérico *perro* abarcara tantos individuos dispares de diversos tamaños y diversa forma; le molestaba que el perro de las tres y catorce (visto de perfil) tuviera el mismo nombre que el perro de las tres y cuarto (visto de frente)». Aquí se nos hace evidente el carácter *construido* de nuestra noción de *individuo*: para nosotros, que tenemos una percepción y una memoria finitas, la imagen de un perro de perfil y la del mismo perro de frente, un minuto después, son tan parecidas que sin mayor inconveniente las consideramos imágenes del *mismo* individuo; la ficción de Funes resalta el hecho de que la noción de *un mismo individuo que permanece el mismo* es una consecuencia directa de la manera finita en que percibimos y recordamos el mundo. Borges lo muestra de una manera más dramática; dice de Funes que «[s]u propia cara en el espejo, sus propias manos, lo sorprendían cada vez».

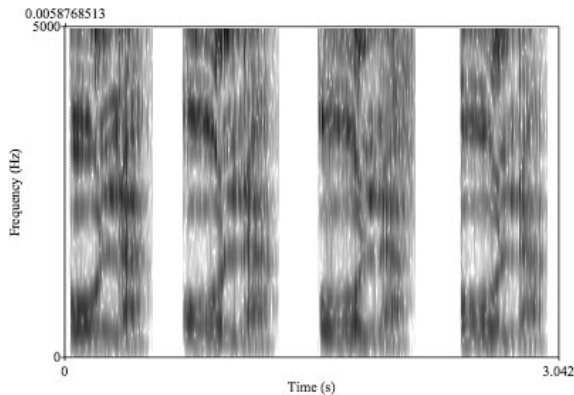
El que a Funes, como leímos líneas atrás, le costara comprender que individuos tan diversos fueran llamados «perros» muestra también el carácter *construido* de nuestra noción de *clase*. En efecto, las clases de individuos que solemos reconocer (y, para las cuales, solemos tener palabras) son constructos dependientes de nuestra percepción finita: si bien podemos reconocer diferencias entre los individuos de una misma clase, encontramos en ellos suficientes similitudes para considerarlos miembros de la misma clase (perros, gatos, sillas, etcétera). Para Funes, esto era muy difícil; según Borges: «[e]ra el solitario y lúcido espectador de un mundo multiforme, instantáneo y casi intolerablemente preciso»; y también: «[e]n el abarrotado mundo de Funes no había sino detalles, casi inmediatos». Nosotros, de manera natural, concebimos la realidad como conformada por entidades que comparten algunas propiedades y se distinguen por otras; por ello, no nos resulta extraño pensarlas agrupadas en especies, géneros o clases. Funes percibe cada entidad con tantos detalles propios que no

consigue encontrar semejanzas suficientes para concebir dos de ellas como *miembros de una misma clase*.

En un trabajito que hicimos mi colega y amigo Alfredo Arnaiz y yo hace algunos años intentamos imaginar cómo percibiría auditivamente el lenguaje humano una persona como Funes². Dado que el funcionamiento de una lengua natural depende (entre otros muchos factores) de nuestra posibilidad de distinguir auditivamente *algunas* propiedades acústicas de los sonidos del habla pero, crucialmente, *no todas*, supusimos que Funes, al tener una percepción infalible, no podría escuchar dos estímulos acústicos distintos como sonidos iguales y, por tanto, no podría utilizar una lengua como lo hacemos nosotros. En efecto, si fuéramos como Funes y pudiéramos distinguir absolutamente todos los detalles acústicos de cada enunciado lingüístico que nos dirigen, no podríamos reconocerlo como una instancia particular de —digamos— el saludo «Hola», la pregunta «¿Cómo te llamas?» o la aseveración «Jorge Luis Borges concibe un personaje dueño de una percepción y una memoria infalibles»; cada uno de estos enunciados sería identificado como un individuo particular, único, y no como un ejemplar de una clase (la palabra «Hola», la oración «¿Cómo te llamas?», etcétera).

Examinemos las propiedades acústicas de cuatro emisiones de la palabra «Hola», producidas por la misma persona, en el espectrograma que aparece a continuación³:

Figura 1. Cuatro instancias de la palabra «Hola» producidas por el mismo hablante⁴.



² Alfredo Arnaiz y Jorge Iván Pérez Silva, 1996. «Una lectura lingüística de “Funes el memorioso”» [manuscrito]. Lima, PUCP.

³ Un espectrograma es un gráfico que representa en el eje vertical las distintas frecuencias que componen las ondas sonoras complejas que sirven como estímulos acústicos. En el eje horizontal se representa el tiempo, de modo que el espectrograma nos permite visualizar cómo cambia la estructura de frecuencias de un enunciado conforme se va articulando (ver Martínez Celdrán & Fernández Planas, 2007).

⁴ Todos los espectrogramas han sido elaborados por el autor [nota de editores].

En la figura 1, podemos observar, representadas gráficamente como manchas en distintos tonos de gris, algunas propiedades acústicas comunes a las cuatro instancias de la palabra «Hola». Resulta claro, sin embargo, que estas cuatro emisiones no son físicamente idénticas entre sí: si bien reconocemos manchas similares en los cuatro enunciados, también comprobamos diferencias entre ellos. Ahora bien, estas diferencias acústicas no resultan ningún obstáculo para que los hablantes reconozcamos con facilidad estos cuatro ejemplares como instancias de la misma palabra. Intuyo, sin embargo, que Funes no habría podido hacer esto, pues se hubiera quedado enmarañado entre miles de detalles acústicos. Su registro de cada enunciado sería tan rico en detalles para él que sería incomparablemente distinto al de cualquier otro enunciado. Para nosotros, sencillamente, son instancias de la palabra «Hola».

En este punto, resulta conveniente preguntarse si la categorización que realizamos de distintos individuos como perros, gatos, sillas, etcétera y la correspondiente a diversos enunciados como instancias de un mismo tipo lingüístico son idénticas. Notemos, para empezar, que las primeras clases dependen notoriamente de una categorización de la realidad que involucra más consideraciones que solo la información de nuestros sentidos. Nuestra concepción del mundo resulta de un largo proceso de aprendizaje en el que desempeña un rol fundamental la interacción con otras personas: si bien nuestro conocimiento del mundo se construye a partir de la información que proporcionan nuestros sentidos, los límites conceptuales son establecidos culturalmente. A partir de la experiencia sensible, acompañada por la formación cultural, vamos construyendo los límites categoriales que dan como resultado nuestra concepción de la realidad: nuestro aprendizaje de qué individuos debemos agrupar como perros y cuáles no, qué individuos son sillas y cuáles no, depende tanto de nuestras capacidades de percepción y de retención de información sensorial, cuanto de consideraciones propias de la sociedad en la que adquirimos nuestro conocimiento del mundo.

Así, por ejemplo, a simple vista, podemos reconocer algunas semejanzas físicas entre mamíferos marinos, como las focas, los lobos marinos, las ballenas y los delfines, y diferenciarlos de mamíferos cuadrúpedos, como los caballos, las vacas y los perros. La primera de estas dos grandes clases quizá podría también subdividirse, a simple vista, en una clase compuesta por las focas y los lobos marinos, y en otra compuesta por las ballenas y los delfines aunque, ciertamente, es discutible que podamos hacerlo sin prestar especial atención a detalles o sin un entrenamiento explícito. Posiblemente, la distinción entre ballenas y delfines pueda basarse en un rasgo notorio como el tamaño, pero habrá una serie de individuos, de tamaño intermedio, que no podremos categorizar sin los límites trazados por la cultura. Y, sin duda, la distinción entre focas y lobos marinos solo podrá establecerse a partir de un aprendizaje más sofisticado.

En esa misma línea, un biólogo marino podrá realizar muchas más distinciones entre todas estas clases, gracias a su «cultura» especializada.

Esta es la gran diferencia entre Funes y nosotros. Si llevamos al extremo la ficción de Borges, tendremos que aceptar que su personaje no podía concebir clases porque su mundo estaba poblado solo por características distintivas: ningún individuo sería lo suficientemente parecido a ningún otro individuo como para considerarlos miembros de una misma clase. Es más, como vimos líneas atrás, en su mundo no habría individuos pues los cambios incesantes de propiedades se sucederían en un flujo continuo de transformaciones.

Ahora bien, para la identificación de diferentes enunciados lingüísticos como instancias de un mismo tipo —por ejemplo, las cuatro emisiones de «Hola» de la figura 1—, también es necesaria la conjunción de la percepción y la memoria con el aprendizaje cultural. La identificación de una señal acústica como un saludo, una pregunta o una aseveración (o cualquier otro acto de habla) depende crucialmente de aprender intersubjetivamente las asociaciones convencionales entre los sonidos de la lengua y las intenciones comunicativas de sus usuarios. El mismo requisito tienen que cumplir las diferentes unidades del léxico: solo podemos interpretarlas (esto es, conocer su significado) si aprendemos socialmente las convenciones que sustentan el vínculo entre su expresión y el contenido que quieren transmitir sus usuarios.

Sin embargo, hay un momento lógicamente previo al reconocimiento de estas asociaciones convencionales: primero es necesario poder identificar la expresión sonora como tal. Para dominar una lengua, el hablante-oyente debe poder reconocer en los estímulos acústicos —siempre físicamente únicos e irrepetibles— los *mismos* sonidos. La condición para esto es no distinguir todas las propiedades sonoras presentes en la señal acústica, es decir, escuchar como *igual* lo que no es físicamente igual. Esto es algo que hacemos las personas normalmente cuando aprendemos nuestra lengua materna: nuestra capacidad para distinguir sonidos del habla es mayor cuando nacemos; durante los primeros meses de vida, nuestra percepción se va adaptando a las distinciones sonoras que son pertinentes para el funcionamiento de nuestra lengua como un sistema semiótico, lo que equivale a decir que vamos *perdiendo* nuestra capacidad para distinguir sonidos del habla, pues nuestra percepción se va sintonizando al inventario cerrado de sonidos que conforma nuestro sistema fonológico (Eimas y otros, 1971; Kuhl, 1991; Kuhl y otros, 2006).

Si bien esta adaptación al medio ambiente lingüístico depende también de la interacción con otros hablantes, aprender a distinguir los sonidos del habla no es como aprender a distinguir ballenas de delfines o focas de lobos marinos. Estas categorías pertenecen a un nivel de conciencia compartido por los miembros de una comunidad, el nivel del pensamiento o conocimiento del mundo. En cambio, los sonidos

del habla pertenecen a un nivel que podríamos llamar «operativo» o «procesual»: el cerebro procesa la información acústica (siempre individualmente diferente) y la organiza en unidades (tipos sonoros), pero no con el fin de llevarlas a la conciencia —pues no es el sonido en sí mismo al que hay que atender en un intercambio lingüístico— sino como parte de las «operaciones» necesarias para la identificación de las unidades lingüísticas mayores (palabras, frases, etcétera) a las que les asociamos un significado.

Así, para reconocer diferentes instancias de la palabra *hola* es necesario poder reconocer los sonidos individuales que la conforman —[ola]—, para distinguirla, por ejemplo, de las palabras *ala*, *hora* y *ole*, cuyos sonidos individuales son, respectivamente, [ala], [ora] y [ole]⁵. Estas distinciones nos resultan claramente perceptibles a los hablantes de castellano, pero no lo son para todas las personas. La diferencia entre *hola* y *hora* no la notan con facilidad, por ejemplo, los hablantes de chino o japonés, pues no es parte del sistema fonológico de su lengua materna. La consecuencia de esto es que escuchan —«categorizan o clasifican auditivamente»— como un mismo sonido ambos estímulos acústicos. Como se puede ver, esta dificultad para distinguir los sonidos [r] y [l] se encuentra al otro extremo de lo que le ocurriría a Funes: mientras que los hablantes de chino o japonés categorizan auditivamente estos estímulos como un mismo sonido, Funes no podría categorizar *ningún* par de estímulos acústicos (por más parecidos que fueran para nosotros) como instancias de un mismo sonido —digamos, [r]— porque todo estímulo acústico sería para él completamente distinto a otro.

Así, pues, de acuerdo con nuestra percepción y memoria limitadas, y en función de los estímulos acústicos propios de nuestro medio ambiente lingüístico, los hablantes-oyentes aprendemos a procesar auditivamente el habla. Propongo que esta organización particular de la materia acústica que realiza nuestro cerebro de acuerdo con su estructura propia y a partir de los estímulos lingüísticos a los que ha estado expuesto desde el nacimiento —organización que determina la forma en que escuchamos— genera *unidades perceptuales* que se distinguen de las ondas sonoras mismas, de la respuesta anatómico-fisiológica del oído, así como de las operaciones físicas del cerebro, de las que emergen. Estas unidades de percepción se distinguen, a su vez, de las *unidades conceptuales* que emergen de la organización de la información que realiza nuestro cerebro cuando categorizamos individuos en clases como perro, gato o silla, lo que, como vimos, resulta en nuestro conocimiento del mundo.

⁵ La representación gráfica de los sonidos individuales como letras entre corchetes es la convención más usual en las ciencias del lenguaje. Las letras pertenecen al Alfabeto Fonético Internacional propuesto por la Asociación de Fonética Internacional (1999).

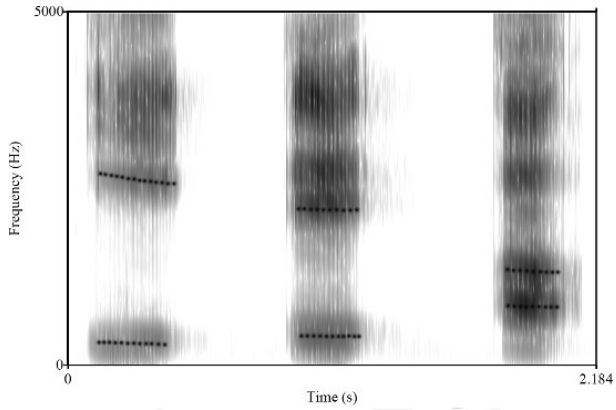
El hecho de que podamos distinguir entidades y procesos físicos probadamente diferentes (estímulos acústicos) que corresponden a una sola unidad perceptual (sonido) resultante del funcionamiento de nuestro cerebro es razón suficiente para describir a aquellos con un lenguaje fisicalista y a esta con un lenguaje mentalista: algo como «dos estímulos acústicos de tales y tales frecuencias son procesados por el cerebro [lenguaje fisicalista] y percibidos como un solo sonido del habla [lenguaje mentalista]». Esto no significa comprometerse con una ontología dualista. Equivale simplemente a reconocer el hecho de que las unidades perceptuales (como las conceptuales) no son reductibles a las propiedades acústicas de las ondas sonoras, sino que resultan de la especial organización que realiza nuestro cerebro de estas⁶.

Ahora bien, desde los orígenes de la lingüística y la fonología modernas, con De Saussure (1945[1916]), Trubetzkoy (1968[1939]) y Jakobson (1968[1941]), se ha subrayado la importancia fundamental que tienen para el funcionamiento de los sistemas lingüísticos, las *diferencias* entre sus unidades. Para De Saussure, en efecto, cada lengua ofrece una organización particular de la materia sonora para la constitución de los signos que utilizan sus usuarios: «Cada idioma compone sus palabras a base de un sistema de elementos sonoros, cada uno de los cuales forma una unidad netamente deslindada y cuyo número está perfectamente determinado. Pero lo que los caracteriza no es, como se podría creer, su cualidad propia y positiva, sino simplemente el hecho de que no se confunden unos con otros. Los fonemas son ante todo entidades opositivas, relativas y negativas» (1945 [1916], p. 201).

El caso de los sonidos vocálicos es particularmente ilustrativo de cómo los sonidos del habla no se definen positivamente sino por oposición. Los sonidos vocálicos se producen modificando, por resonancia, la onda compleja de las cuerdas vocales mediante variaciones en el tamaño de las cavidades faríngea y bucal. La siguiente figura muestra espectrogramas de los sonidos vocálicos [i, e, a] del castellano:

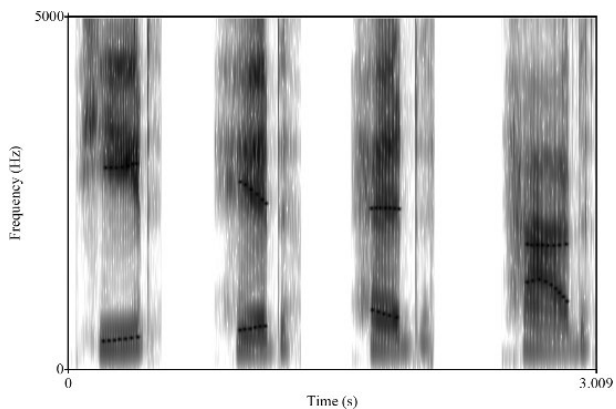
⁶ El recurso a nociones mentalistas para describir la manera en que nuestro cerebro organiza la información del medio ambiente se encuentra en los orígenes mismos de la ciencia cognitiva contemporánea (ver Thagard, 2008; Westbury & Wilensky, 2001). La gramática generativa, gestada desde finales de la década de 1950 del siglo pasado por Noam Chomsky, se sustenta en la idea de que el cerebro humano organiza la información lingüística mediante representaciones abstractas que corresponden —pero, crucialmente, no se identifican con— los estímulos acústicos que les sirven de expresión. Estas representaciones abstractas —de carácter esencialmente operativo o procesual— son de distintos tipos, de acuerdo con el nivel de organización o descripción del sistema lingüístico al que pertenecen: representaciones fonológicas —como las unidades perceptuales de las que hemos venido hablando— morfológicas, sintácticas, etcétera. Ver Chomsky (1986) y Jackendoff (2002), quien engloba los diferentes niveles de procesamiento lingüístico bajo el nombre de nivel «mental-f(uncional)».

Figura 2. Algunas vocales del castellano



Si bien los espectrogramas presentan semejanzas, hay una diferencia acústica que determina que estos estímulos acústicos sean percibidos como vocales diferentes por los hablantes de castellano: las frecuencias de las dos franjas horizontales oscuras inferiores (marcadas con puntos para apreciarlas mejor). Las medidas de estas dos franjas, conocidas como «formante 1» y «formante 2», son, de acuerdo con los estudiosos, las claves acústicas que determinan el timbre o calidad de los sonidos vocálicos. Veamos la siguiente figura, que muestra los espectrogramas de las palabras *heed*, *hid*, *head* y *had* del inglés (los formantes 1 y 2 de las vocales respectivas aparecen marcados con puntos):

Figura 3. Algunas vocales del inglés



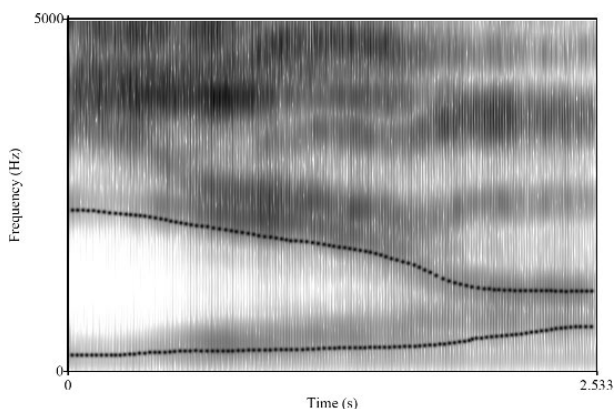
Como en el caso anterior, estos espectrogramas también presentan los formantes 1 y 2 en diferentes posiciones, lo que determina que los estímulos sean percibidos

como sonidos diferentes. Ahora bien, ocurre que los hablantes de castellano con poco entrenamiento en el inglés no distinguimos normalmente todos los sonidos que sí distinguen sin dificultad los anglohablantes. Es conocido el hecho de que solemos escuchar la palabra *hid* bien como *heed*, bien como *head*, y que escuchamos una palabra como *hood* con la misma vocal que *who'd*.

Así, pues, tanto el inglés como el castellano utilizan un número finito de sonidos vocálicos para la construcción de sus signos, pero los sonidos no se corresponden unos con otros y no pueden definirse en términos acústicos independientemente del uso que se les da en cada lengua. Más bien, los sonidos lingüísticos se definen por *oposición* entre ellos dentro de cada lengua porque cada uno corresponde a un rango acústico cuyos límites están determinados por los rangos de los sonidos vecinos.

Esto se puede ilustrar con el siguiente espectrograma, que corresponde a un continuo sonoro que va desde un estímulo acústico perceptible como [i] (al inicio) hasta uno perceptible como [a] (al final):

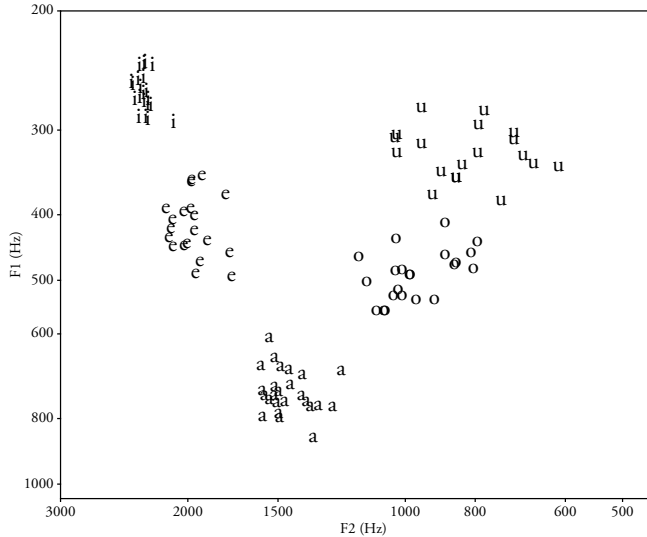
Figura 4. Continuo acústico



La figura 4 muestra cómo los formantes 1 y 2 van acercándose paulatinamente. Los estímulos que se perciben de manera similar a [i] presentan mayor separación entre sus formantes y los que se perciben de manera similar a [a] presentan formantes muy cercanos. Entre estos dos extremos, los estímulos acústicos se perciben de acuerdo con la lengua que uno haya aprendido, dado que, a lo largo de este continuo, diferentes lenguas delimitan distintas unidades sonoras: el inglés distingue, al menos, las cuatro vocales de la figura 3 (si no más) y el castellano distingue las tres de la figura 2.

Lo más interesante es que las articulaciones particulares de cada sonido no son idénticas, sino que se dispersan en un campo que cubre gran parte del continuo, como muestra la siguiente figura:

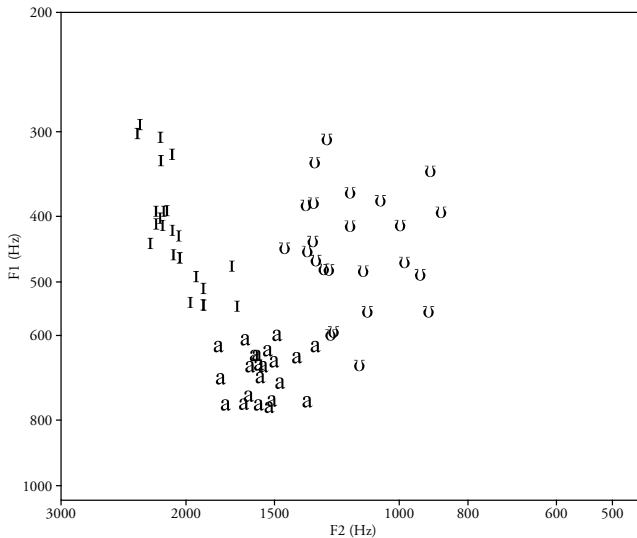
Figura 5. Dispersión de los estímulos acústicos correspondientes a las vocales castellanas



Fuente: Pérez, Acurio y Bendezú (2008, p. 15).

Así, la [i] y la [e] del castellano cubren más o menos el mismo espacio que cubren la [i, ɪ, e, æ] del inglés. Por otra parte, la unidad perceptual [ɪ] del quechua cubre el espacio de la [i] y la [e] del castellano, como se ve en la siguiente figura:

Figura 6. Dispersión de los estímulos acústicos correspondientes a las vocales quechuas

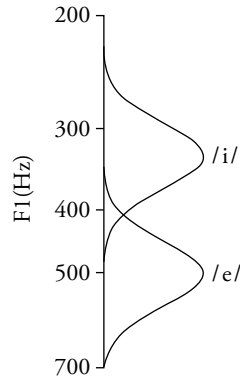


Fuente: Pérez, Acurio y Bendezú (2008, p. 18).

Así, pues, dentro de cada lengua los hablantes suelen producir diferentes estímulos acústicos con la intención de producir una misma unidad perceptual y estos distintos estímulos son, de hecho, escuchados como instancias de un solo sonido por los hablantes de la misma lengua. Las distintas instancias de las vocales inglesas de *heed* o de *hid* no son nunca idénticas acústicamente: lo que las identifica no son sus propiedades acústicas inherentes sino su relación de oposición con sus vocales vecinas. Y lo mismo ocurre con las vocales del castellano: como muestra la figura 5, escuchamos como /e/ los estímulos que no sean tan altos que se confundan con /i/ ni tan bajos que se confundan con /a/; pero no podemos definirlos en términos acústicos absolutos. Esta característica de las lenguas naturales haría imposible que Funes escuchara los sonidos del habla como nosotros: no podría escuchar como auditivamente igual lo que es acústicamente diferente. Para nosotros, en cambio, gracias a nuestras capacidades perceptivas limitadas, resulta fácil el agrupamiento de estímulos acústicos distintos en categorías auditivas⁷.

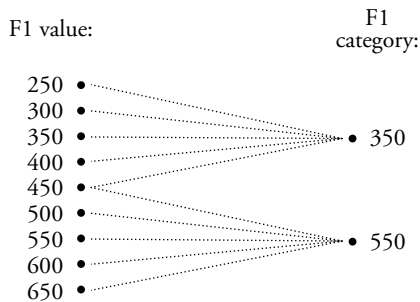
Los estudiosos han mostrado que existe una relación directa entre la manera en que los hablantes producen sus vocales en una comunidad de habla y la manera en que aprenden a percibirlos. En palabras de Escudero, «la experiencia con las formas concretas con que se producen los sonidos resulta en una percepción óptima a partir de la cual los oyentes aprenderán a categorizar la señal hablada en vocales y consonantes que coincidirán con lo producido por el hablante» (2005, p. 52; la traducción es mía). Como vimos, los estímulos acústicos que corresponden a una unidad perceptual no son acústicamente idénticos. El siguiente gráfico muestra, con respecto a la medida del formante 1, que la mayoría de estímulos acústicos que sirven como realización material del sonido [i] en castellano se encuentran alrededor de los 350 Hz, pero que también hay un gran porcentaje de estímulos que se encuentran por encima y por debajo de esta medida.

⁷ En Arnaiz y Pérez (1996) la tesis era que Funes no podría dominar una lengua. Eduardo García Ramírez (comunicación personal), sin embargo, señala correctamente que las habilidades extraordinarias de Funes sí le permitirían calcular los límites de variación correspondientes a las diferentes unidades lingüísticas de una lengua y, por tanto, convertirse en un hablante-oyente de la misma. De todas maneras, nuestra forma de aprender una lengua sería diferente de la de Funes, en tanto nosotros identificamos unidades perceptuales por hipodiferenciación, mientras él lo haría por memorización de los distintos estímulos acústicos que los hablantes usamos de manera equivalente. Agradezco a Eduardo García Ramírez su atinado comentario.

Figura 7. Promedios de producción y distribuciones de [i] y [e] del castellano

Fuente: Escudero (2005, p. 53).

Lo mismo se puede apreciar con respecto al sonido [e], pero en relación con la medida de 550 Hz. La siguiente figura representa la manera en que los oyentes construyen sus categorías auditivas sobre la base de agrupaciones o «mapeos» de estímulos acústicos diferentes en un solo sonido que se distingue de manera discreta o categórica de otro.

Figura 8. Mapeo del continuo acústico a las categorías auditivas discretas

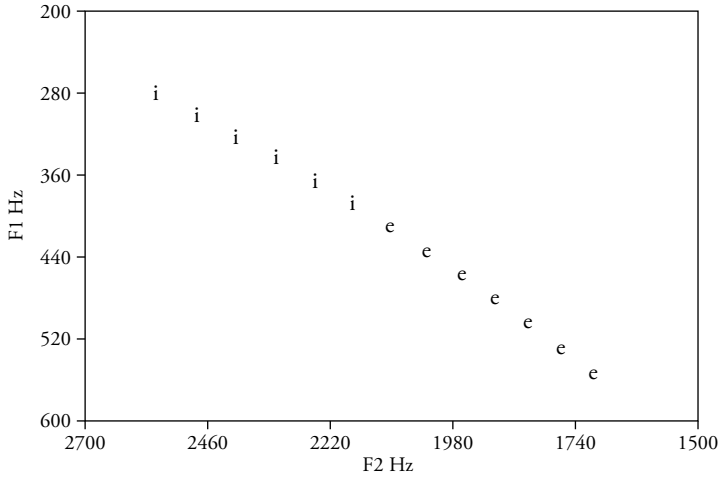
Fuente: Adaptado de Escudero (2005, p. 45).

Así, pues, el resultado final del aprendizaje de una lengua es un estado determinado del cerebro en función del cual procesamos u organizamos los estímulos acústicos. Como consecuencia de esto, los cerebros de hablantes de diferentes lenguas procesan los mismos estímulos acústicos de maneras diferentes. Una metáfora usual de este fenómeno es la del «filtro»: se dice que los hablantes «filtran» los estímulos acústicos a través de las categorías de su sistema fonológico.

Un ejemplo notable de esto es la distinta manera en que los hablantes de castellano y los de quechua percibimos los sonidos vocálicos. Napurí (2012) muestra

que ante un continuo de estímulos acústicos sintetizados, los hablantes nativos de castellano distinguimos las vocales [i] y [e] de manera categórica, de acuerdo con la siguiente figura:

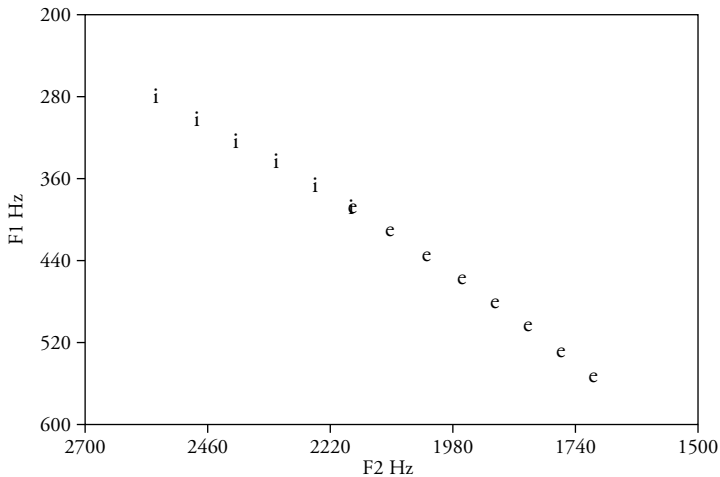
Figura 9. Percepción de hablantes nativos de castellano 1



Fuente: Napurí (2012, p. 56).

Como se ve, algunos hablantes oyentes escuchan los seis primeros estímulos como [i] y los siguientes siete como [e]; otros, en cambio, categorizan el sexto estímulo a veces como [i] y a veces como [e]:

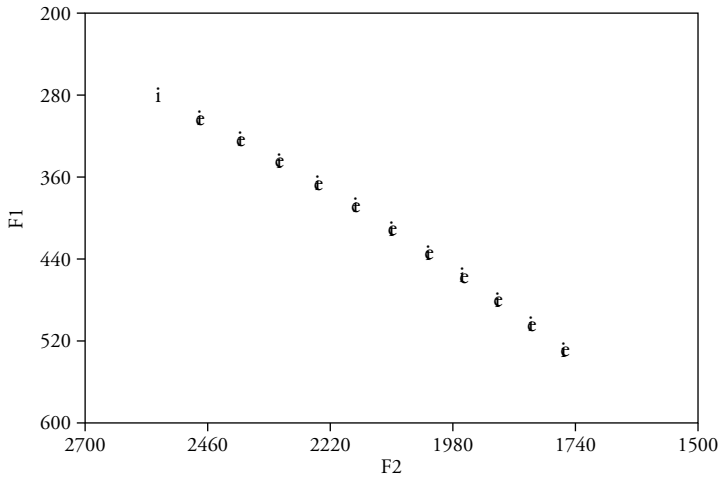
Figura 10. Percepción de hablantes nativos de castellano 2



Fuente: Napurí (2012, p. 57).

Por su parte, los bilingües quechuahablantes que tienen un conocimiento muy incipiente de castellano identifican solo el primer estímulo como [i], mientras que los otros doce los categorizan a veces como [i] y a veces como [e]:

Figura 11. Percepción de bilingües incipientes quechua-castellano

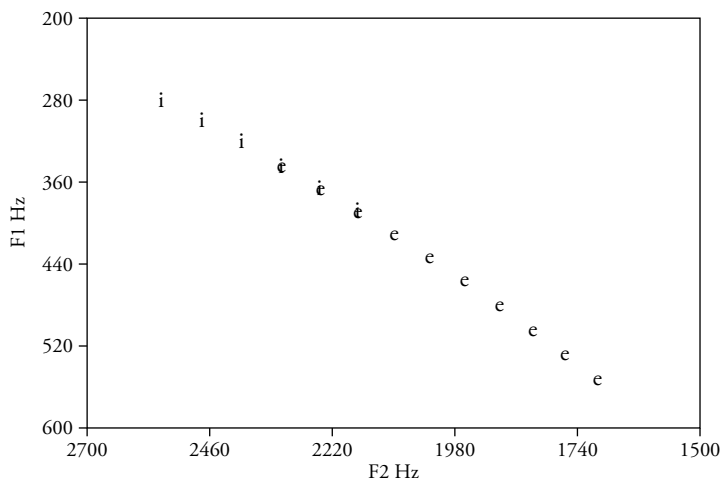


Fuente: Napurí (2012, p. 30).

Esta identificación aleatoria demuestra, de acuerdo con Napurí (2012), que la percepción de los estímulos está siendo «filtrada» por las categorías fonológicas de la lengua materna dominante de los bilingües quechuahablantes; en efecto, como vimos antes, el quechua solo distingue una unidad perceptual allí donde el castellano distingue dos.

Algo que resulta particularmente interesante es que, conforme los bilingües van aprendiendo más y más castellano, van distinguiendo cada vez más los estímulos acústicos. Si bien Napurí (2012) no realiza un estudio longitudinal sobre el desarrollo de la percepción de los bilingües, su investigación transversal, con distintos individuos en diversos niveles de aprendizaje del castellano, nos da una idea clara sobre cómo van construyendo y modificando sus categorías auditivas, hasta que los bilingües avanzados muestran una categorización idéntica a la de los monolingües. La figura siguiente presenta la categorización de un bilingüe quechua-castellano intermedio:

Figura 12. Percepción de bilingües intermedios quechua-castellano



Fuente: Napurí (2012, p. 50).

En conclusión, la manera en que el cerebro humano procesa la información lingüística sugiere la emergencia de unidades perceptuales o auditivas que claramente no son idénticas a sus correlatos físicos. La evidencia empírica, presentada en este trabajo, sobre la distinta manera en que los hablantes de quechua, castellano e inglés perciben los mismos estímulos acústicos muestra la conveniencia de describir los sistemas fonológicos en términos mentalistas, en la medida en que el funcionamiento de los sistemas lingüísticos —en particular, la distinción de palabras— no depende de los estímulos acústicos en sí mismos sino en tanto organizados lingüísticamente por el cerebro en tales unidades perceptuales⁸.

⁸ Estas unidades perceptuales no son las únicas unidades fonológicas que presentan carácter mental. La sutil organización de los sistemas fonológicos incluye unidades menores a las unidades perceptuales —rasgos segmentales y suprasegmentales—, así como unidades mayores a ellas —sílabas, pies, patrones acentuales, entonacionales, etcétera—. A esto habría que añadir la existencia de reglas o restricciones fonológicas, dependiendo del paradigma teórico asumido. Ver Goldsmith (1994).

REFERENCIAS BIBLIOGRÁFICAS

- Asociación de Fonética Internacional (1999). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- Borges, Jorge Luis (1983). *Narraciones*. Bogotá: Oveja Negra.
- Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Nueva York: Praeger.
- De Saussure, Ferdinand (1945[1916]). *Curso de lingüística general*. Buenos Aires: Losada.
- Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk & James Vigorito (1971). Speech Perception in Infants. *Science*, 171, 971-974.
- Escudero, Paola (2005). *Linguistic Perception and Second Language Acquisition*. Utrecht: LOT.
- Goldsmith, John (ed.) (1994). *The Handbook of Phonological Theory*. Cambridge: Blackwell.
- Jackendoff, Ray (2002). *Foundations of Language*. Oxford: Oxford University Press.
- Jakobson, Roman (1968[1941]). *Child Language, Aphasia and Phonological Universals*. La Haya: Mouton.
- Kuhl, Patricia K. (1991). Human Adults and Human Infants Show a «Perceptual Magnet Effect» for the Prototypes of Speech Categories, Monkeys Do Not. *Perception and Psychophysics*, 50, 93-107.
- Kuhl, Patricia K. y otros (2006). Infants Show a Facilitation Effect for Native Language Phonetic Perception Between 6 and 12 Months. *Developmental Science*, 9(2), F13-F21.
- Martínez Celdrán, Eugenio & Ana María Fernández Planas (2007). *Manual de fonética española*. Barcelona: Ariel.
- Napurí, Andrés (2012). *Adquisición de la percepción de vocales del español por bilingües quechua-español*. Tesis de maestría, PUCP.
- Pérez, Jorge, Jorge Acurio & Raúl Bendezú (2008). *Contra el prejuicio lingüístico de la motosidad*. Lima: Instituto Riva Agüero-PUCP.
- Thagard, Paul (2008). *La mente: introducción a las ciencias cognitivas*. Buenos Aires: Katz.
- Trubetzkoy, Nikolai (1968[1939]). *Principios de fonología*. Buenos Aires: Kapelusz.
- Westbury, Chris & Uri Wilensky (2001). La representación del conocimiento en la ciencia cognitiva. Sus implicancias en la educación. En Ministerio de Educación del Perú-MINEDU, *¿Sabemos cómo se aprende? Nuevos enfoques sobre el aprendizaje* (pp. 25-87). Lima: MINEDU.

LA NATURALEZA DE LA FACULTAD DEL LENGUAJE

Liza Skidelsky

Universidad de Buenos Aires – Consejo Nacional
de Investigaciones Científicas y Técnicas

1. INTRODUCCIÓN

La versión más difundida en el ámbito filosófico de la naturaleza de la facultad del lenguaje, postulada por Chomsky, es la versión epistémica que la entiende como el conocimiento proposicional que posee el hablante/oyente acerca de su lengua (Fodor, 1983, Dwyer & Pietroski, 1996; Knowles, 2000). En contra de esto, John Collins (2004, 2006 y 2008) sostiene la versión arquitectural que la entiende como un sistema computacional especificado de manera abstracta. Acuerdo plenamente con Collins en el rechazo de la versión epistémica, de manera que no me ocuparé aquí de ella. Lo que me interesa discutir es su vertiente de la versión arquitectural que sostiene que la facultad del lenguaje es un sistema computacional de la mente/cerebro especificado de manera abstracta, y que esta especificación hace que no pueda concebirse como un mecanismo, esto es, una parte de un nexo causal.

En contraposición, defiendo la idea de que la facultad del lenguaje es un mecanismo causal (aunque idealizado) de la misma naturaleza que los mecanismos de los cuales se pretenden dar cuenta en cualquier teoría científica de la arquitectura cognitiva. En este sentido, intento elaborar un enfoque alternativo en el que tenga sentido concebir al programa minimalista como un modelo cognitivo que describe un mecanismo causal. Aunque el modelo en sí mismo no sea mecanicista, en el sentido de que solo especifica las propiedades funcionales de su objeto de descripción —es decir, es un modelo funcional—, este último —esto es, la facultad del lenguaje— es un mecanismo. Así, sostendré que se puede defender que en psicología cognitiva se trata de dar cuenta de mecanismos cognitivos pero que, sin embargo, hay mecanismos cognitivos que pueden tener una explicación mecanicista o funcional dependiendo del tipo de correspondencia que pueda haber entre el modelo psicológico y el mecanismo cognitivo.

2. LA FACULTAD DEL LENGUAJE COMO SISTEMA COMPUTACIONAL ABSTRACTO

La versión más difundida de la naturaleza de la competencia lingüística, al menos en el ámbito filosófico, es la versión epistémica. Esta versión, que ha sido sostenida principalmente por Fodor (1975, 1983 y 2001), sostiene que la competencia lingüística o lengua-I es una cuestión acerca de lo que el hablante/oyente *sabe* o *conoce*. Dado que el conocimiento es tradicionalmente entendido en términos de actitudes proposicionales (AP), la competencia lingüística consistiría en un conjunto de AP (por ejemplo, creencias) acerca de la lengua particular alcanzada por el hablante/oyente¹. Así, la hipótesis de la Facultad del Lenguaje (FL), según Fodor, y en términos de Collins: «no es una propuesta acerca de la arquitectura de la mente/cerebro, en la que esta tesis ofrece una explicación causal de cómo el hablante/oyente adquiere y mantiene conocimiento del lenguaje y lo pone en uso» (2004, p. 503).

En contra de esta concepción, la versión arquitectural considera que hay que entender la FL como un componente computacional de la mente/cerebro que se encarga del procesamiento de la información lingüística. Collins (2004, 2006 y 2008) defiende una vertiente de la versión arquitectural que si bien entiende a la FL como un sistema computacional, acuerda, sin embargo, con el aspecto de la versión epistémica reflejado en la cita de arriba y, en este sentido, el sistema computacional está especificado de manera abstracta. Esto quiere decir que aunque la FL es un componente de la mente/cerebro, está abstraída de la estructura causal del cerebro. A favor de su versión arquitectural abstracta sostiene, entonces, que en tanto que su descripción no es en términos causales, la FL no es un mecanismo causal. Veamos, en primer lugar, la descripción estándar de la FL, con la cual acuerdo y luego las razones, con las cuales no acuerdo, que ofrece Collins para sostener que puesto que la teoría ofrece una especificación abstracta de la FL, esta no es un mecanismo causal.

2.1. La facultad del lenguaje

La facultad del lenguaje es un subsistema de la mente/cerebro humano. Tiene dos componentes: un sistema cognitivo y sistemas de ejecución (articulatorio-perceptivo y conceptual-intencional). La lingüística se ocupa del estado inicial y el estado estable

¹ Esta es la vertiente más representativa de la versión epistémica. Asumo que esta versión abarca todas las vertientes que consideran que la competencia lingüística o lengua-I consiste (solo o fundamentalmente) en un cuerpo de información lingüística en la mente/cerebro, que puede entenderse desde una estructura de base de datos, pasando por representaciones mentales y AP, hasta «conocimiento». En este sentido, si bien la cuestión de la relación epistémica entre el sujeto y el conocimiento proposicional lingüístico forma parte de algunas vertientes más representativas y de allí la denominación de esta versión, no considero que le sea esencial.

del sistema cognitivo. La Gramática Universal (GU) es el estudio del estado inicial y las «gramáticas» se ocupan del estudio de las lenguas-I, esto es, del estado estable alcanzado por los hablantes, es decir, su competencia lingüística. La GU consiste en un conjunto finito de principios válidos para todas las lenguas-I humanas posibles y parámetros cuyos valores se fijan en función de la experiencia lingüística durante la adquisición del lenguaje. Ambos, principios y parámetros, son parte del equipamiento biológico humano. Las lenguas-I son distintas instanciaciones de los parámetros de la GU². La noción técnica de lengua-I supone una cierta idealización que consiste en hablantes con experiencia uniforme en comunidades lingüísticas homogéneas (sin diversidad de dialectos ni variaciones entre hablantes). La lengua-I es Interna (a la mente/cerebro), Individual (de hablantes particulares) e Intensional (es una función, caracterizada en términos intensionales, que genera descripciones estructurales). La lengua-I (su estado inicial y su estado estable) constituye un objeto real en las mentes/cerebros de los hablantes, el lingüista estudia aquellos aspectos abstractos del cerebro, esto es, los mentales (lingüísticos). En tanto su objeto de estudio es un objeto natural, la lingüística forma parte de las ciencias naturales, dado que es parte de la psicología y esta, en última instancia, es parte de la biología, y encara su objeto con la misma metodología que emplea el resto de las ciencias naturales (tesis del naturalismo metodológico).

Según el programa minimalista, cada lengua-I consiste en un mecanismo computacional y un léxico (Chomsky 1995a, 1995b). La función (intensional) del mecanismo es generar descripciones estructurales de las expresiones lingüísticas a partir de elecciones léxicas. El procedimiento computacional toma como *inputs* rasgos léxicos y sus *outputs* son las descripciones estructurales que son objetos abstractos (simbólicos o formales), que llevan en sí toda la información semántica, sintáctica y fonética necesaria para asociar el sonido y el significado de las expresiones lingüísticas. El léxico es una lista de ítems léxicos idiosincráticos (que no se derivan de principios generales). Los ítems léxicos están caracterizados por una serie de rasgos semánticos, sintácticos y fonéticos. El mecanismo computacional toma ítems léxicos cuyos rasgos fonéticos se computan para dar lugar a una representación fonética (o FF), toma rasgos semánticos para dar lugar a una representación semántica (o FL) y los rasgos formales entran en la computación pero deben ser eliminados (en el camino a la FF) para dar lugar a una representación convergente,

² Los principios son restricciones a las que están sujetas todas las lenguas. Por ejemplo, el objeto de un verbo forma un sintagma con el verbo que no incluye el sujeto o, dicho de otra manera, los objetos se combinan con los verbos antes de que lo hagan los sujetos (requisito de verbo-objeto). Los parámetros, en cambio, se fijan a partir de la experiencia lingüística. Un ejemplo de parámetro es el del sujeto nulo, así hay lenguas que o bien requieren sujeto o bien no lo requieren (confróntese Baker, 2001).

esto es, interpretable por los sistemas de ejecución de la interface³. Son estos sistemas de interface los que imponen condiciones de legibilidad puesto que son los que interpretan el output y siguen sus instrucciones para el pensamiento y la acción. El mecanismo computacional toma ítems léxicos y computa (por medio de la operación principal de *Merge*) sobre sus rasgos hasta un punto (denominado *Spell-Out*) en el que las derivaciones se separan hacia los dos niveles de interfaz. Las operaciones que se realizan entre *Spell-Out* hasta concluir en la FF no son de la misma naturaleza que las que llevan a la FL, se asume que el mecanismo computacional lleva a la FL, mientras que un componente separado, el componente fonológico, es el que realiza las operaciones hasta la FF. Del conjunto de derivaciones convergentes que satisfacen las condiciones de la interface hay un subconjunto de derivaciones admisibles seleccionadas por principios de economía del sistema. De manera que una expresión lingüística E de una lengua L consiste, al menos, de un par {FON, SEM} formado por una derivación convergente óptima. FON(E) es una representación de la FF que contiene información relevante al sonido de E y SEM(E) es una representación de la FL que porta información relevante al significado de E. Se asume que FON(E) y SEM(E) son objetos legítimos que reciben interpretación por los sistemas articulatorio-perceptivo y el conceptual-intencional, respectivamente.

En este sentido, lo que se dice en el modelo minimalista es que la competencia de un hablante consiste en tener una lengua-I que es tener un procedimiento generativo, en la mente/cerebro, que asigna descripciones estructurales a expresiones lingüísticas. Los estados representacionales de la competencia lingüística, o de la lengua-I, son estados que almacenan información lingüística, sobre la que opera el sistema y a la que no se tiene acceso consciente. Los únicos niveles de representación en la teoría minimalista son la FL y la FF que son el *output* del mecanismo de la lengua-I, es decir las descripciones estructurales (los pares {FON, SEM} de cada E de cada L), que consisten, en última instancia, en una disposición de rasgos léxicos sintácticos, semánticos y fonéticos (Chomsky, 1995b, p. 225)⁴. Así, la FL es un subsistema del cerebro que tiene un estado inicial genéticamente determinado y a través de computaciones sobre ese estado inicial más los datos provenientes del entorno manipula esa información de manera de dar lugar a *outputs* que luego pasarán a otros subsistemas que subyacen al uso del lenguaje. De modo que en la descripción estándar de la FL, queda claro que es un sistema computacional de procesamiento de información lingüística.

³ Si bien utilizo las mismas iniciales para la Facultad del Lenguaje y la Forma Lógica, queda clara su referencia en los contextos en donde aparece.

⁴ Chomsky (1995a) conjetura que el mecanismo computacional de la lengua-I es completamente derivacional. En este sentido, duda de que sea apropiado postular niveles de representación como la FL y FF.

De otro modo, entender a la FL solo como un cuerpo de conocimiento o meramente como un conjunto de información o representaciones lingüísticas implicaría la pérdida del poder explicativo del modelo para dar cuenta del apareamiento sonido-significado (tanto en la adquisición del lenguaje como en el funcionamiento maduro de esta capacidad)⁵.

2.2. Especificación abstracta y mecanismo causal

Basado en esta descripción estándar de la FL, Collins (2004) defiende la idea de que aunque esta es un componente del cerebro, en tanto que su descripción no es en términos causales, la FL está abstraída de la estructura causal del cerebro y, en este sentido, no puede considerársela un mecanismo. Lo que intentaré sostener, en cambio, es que no hay razones de peso para descartar que la FL sea un mecanismo que forme parte de la arquitectura causal cognitiva y que, por ende, la teoría acerca de la FL tenga un rol que jugar en una teoría futura integrada de la ejecución lingüística que pueda estudiarse en términos naturalistas. Voy a concentrarme en una inferencia que considero que no está fundamentada. Esta consiste en partir de la afirmación de que la teoría de la FL no ofrece explicaciones causales para concluir que su objeto de estudio, la FL, no es un mecanismo causal. Esta inferencia en el ámbito de la filosofía de la lingüística no es más que un caso de una inferencia más general que suele estar bastante difundida en la filosofía de la ciencia cognitiva. En particular, los defensores de la explicación funcional consideran que la explicación psicológica (o un modelo psicológico) no es una explicación (o modelo) de mecanismos y, por tanto, no describe mecanismos (Fodor, 1968; Cummins, 1983; Weiskopf, 2011).

Considero que esta relación directa entre el tipo de explicación (o modelo) y la entidad que se postule como responsable de los fenómenos cognitivos a explicar (o modelar) no está justificada. Esto es, no parece haber razones de peso para considerar que una descripción abstracta establece de por sí que la descripción es acerca de un objeto abstracto (en contraposición a un mecanismo causal) ni, como se verá en el apartado 3, que una especificación de un mecanismo deba, ella misma, ser mecanicista⁶. En este apartado me ocuparé de analizar las razones esgrimidas por Collins

⁵ Por supuesto que hay razones adicionales a la preservación del poder explicativo del modelo para sostener que la FL no consiste (tan solo o fundamentalmente en) un conjunto de información pero, como he mencionado, no me ocuparé aquí de esas razones. Véase Liza Skidelsky (2013) para un análisis de algunas de ellas.

⁶ Incluso, en este sentido, es perfectamente compatible sostener que cierta teoría o modelo no ofrece una explicación causal de la entidad responsable de los fenómenos cognitivos que se desean explicar pero que, sin embargo, la teoría describe un mecanismo que posee disposiciones causales que permiten insertarlo en un nexo causal.

en favor de su versión arquitectural abstracta de la FL. En el apartado 3, me ocuparé de la discusión más general, en filosofía de la ciencia cognitiva, entre los defensores de la explicación funcional y mecanicista. Intentaré defender mi idea en este ámbito general, mostrando su aplicación directa a la discusión en filosofía de la lingüística. Cabe aclarar, desde el comienzo, que Collins no afirma que su versión arquitectural abstracta sea la que efectivamente sostiene Chomsky, sino que se alinea con los desarrollos chomskianos y es independientemente coherente. Mostraré que hay razones para pensar que la versión arquitectural causal que se desprende de mis consideraciones cumple de manera más satisfactoria con ambos requisitos.

Según Collins, «la facultad del lenguaje es un sistema computacional de la mente/cerebro especificado de manera abstracta» (2004, p. 529). La teoría de la FL ofrece una especificación abstracta de la función en intensión que aparee sonido con significado. Esta función, si bien establece condiciones límites a lo que el cerebro debería hacer de manera de que seamos hablantes/oyentes competentes, no describe la historia causal de la ejecución lingüística (2006). La explicación de cómo se logra el apareamiento sonido-significado no es una explicación causal *cum* neurológica de, por ejemplo, cómo se es capaz de pedir un café, ni es predictiva de esa conducta. Los estados de la FL no son susceptibles de generalizaciones causales ni tampoco contribuyen a la etiología de los actos lingüísticos, de manera que no participan de las conexiones causales que llevan al comportamiento lingüístico.

La FL consistiría en una serie de condiciones *sui generis* satisfechas por la operación de un cerebro normal. Son *sui generis* en el sentido de que especifican un aspecto de un sistema que no sería visible de otro modo. Ese aspecto es la integración recursiva y estructurada de sonido y significado. En este sentido, la facultad da cuenta de la competencia no de la ejecución (distinción introducida por Chomsky, 1965). Así, la caracterización de la facultad apunta a que tener competencia lingüística consiste en poder aparear de manera sistemática sonidos con significados. De modo que, más que explicar ejecuciones (actos de habla o pensamiento), la FL explica cómo es posible la ejecución en sí (Collins, 2004). De manera que la FL estaría explicativa (aunque no causalmente) relacionada con la ejecución así como, salvando las diferencias, el conocimiento de una partitura explicaría la competencia de alguien en una pieza de Beethoven, un conocimiento que se retiene aun si se desarrolla artritis (2004)⁷.

⁷ Nótese la cercanía con la versión epistémica en la idea de considerar a la lengua-I básicamente como «conocimiento».

Según lo que se desprende de estas consideraciones, la distinción entre sistemas abstractos y mecanismos causales estaría dada por, al menos, dos cuestiones relacionadas entre sí: la de las generalizaciones causales y la distinción competencia-ejecución. Comencemos por la primera. Según Collins (2006), el problema con las generalizaciones lingüísticas radica en un profundo aspecto mistericista relacionado con nuestra incapacidad cognoscitiva de capturar por qué el sistema lingüístico se comporta de determinadas maneras, dando determinadas respuestas a determinados estímulos lingüísticos externos. El hecho de que las respuestas lingüísticas estén relacionadas con las intenciones y objetivos humanos parece removerlas de una historia causal.

Ahora bien, se supone que la teoría de la FL está pensada para dar cuenta de la función que proyecta un conjunto de rasgos léxicos en pares de estructura sonido-significado que son el input para los sistemas de ejecución articulatorio-perceptivo y conceptual-intencional externos al sistema de la lengua-I. La teoría en sí misma no está pensada para dar cuenta del comportamiento lingüístico, esto es, los actos lingüísticos efectivos. Dejando a un lado el aspecto mistericista que poco se condice con un naturalismo metodológico, la teoría lingüística chomskiana (o, por dar otro ejemplo siempre mencionado por el mismo Chomsky, la teoría de la visión de Marr) no intenta dar cuenta del comportamiento lingüístico (o visual), mucho menos del comportamiento aún más complejo en el cual las capacidades lingüísticas (o visuales) están involucradas.

El comportamiento lingüístico parece depender de consideraciones que van más allá del ámbito de la psicología computacional. Hay un aspecto relacionado con el punto de vista de los agentes, sus intereses y demás, que interviene en el comportamiento lingüístico que parece que va más allá de una investigación centrada en mecanismos cognitivos. Esto es porque el objeto de estudio de estas investigaciones son los subsistemas cognitivos, cuyos *outputs* tienen que integrarse de manera de dar lugar al pensamiento y a la acción. Dar cuenta del pensamiento y de la acción es una tarea sumamente compleja para la cual no hay, actualmente, principios generales ni generalizaciones plausibles y mucho menos, lo que esperan los filósofos, leyes *ceteris paribus*. La explicación del complejo comportamiento humano puede ser un desiderátum por tener en cuenta en la psicología computacional, pero su cumplimiento va más allá de esta ciencia. Un objetivo razonable de la psicología, y observable en la práctica psicológica, es dar cuenta de cómo funcionan las capacidades cognitivas, es decir, dar cuenta de los mecanismos subpersonales (Skidelsky, 2006). El hecho de que no haya generalizaciones causales lingüísticas que abarquen el comportamiento lingüístico (o, quizás, nunca las haya), no establece de por sí que los estados lingüísticos no tienen entonces ningún rol causal. Si nos atenemos al mecanismo

mismo de la lengua-I, seguramente es posible establecer generalizaciones que reflejen las disposiciones causales que describen cómo funciona este mecanismo. Puede que no estén expresadas en términos cognitivos, puede que sean solo en términos físicos, pero eso no hace menos causal a la FL en tanto tal⁸.

La idea de que los estados lingüísticos no forman parte del nexo causal de la mente/cerebro también se relaciona, tal como se mencionó, con la distinción competencia-ejecución. Collins (2008) sostiene que la teoría de la FL se ocupa de especificar la función intensional que los sistemas neuronales de producción y comprensión subyacentes respetan. Esta función codifica pares de representaciones fonológicas y semánticas que son explicativas en relación a los juicios de los hablantes/oyentes, entre otras cosas. Collins afirma que: «Esta concepción se opone a un modelo de “procesamiento” de la facultad del lenguaje en la medida en que especificar la función no es especificar *cómo* el hablante/oyente procesa material lingüístico» (2008, p. 5).

Ahora bien, el cuidado de Chomsky en insistir que su ámbito de estudio es la competencia lingüística (es decir, la lengua-I) no es porque no considere que la FL deba formar parte de una teoría de la ejecución sino porque cree que es el único aspecto de una futura teoría integrada de la ejecución que puede estudiarse en términos naturalistas. Algunas afirmaciones típicas de Chomsky que dan a entender que la FL es un componente del mecanismo que subyacería a la ejecución lingüística son las siguientes: «el estudio de modelos de ejecución que incorporan las gramáticas generativas puede ser un estudio fructífero; más aún, es difícil imaginar cualquier otra base sobre la cual una teoría de ejecución pudiera desarrollarse» (1965, p. 15); «Si aceptamos, como lo hago la aseveración de Lenneberg en el sentido de que las reglas de la gramática están involucradas en los mecanismos del procesamiento, entonces es de esperarse que la evidencia referente a la producción, el reconocimiento, el recuerdo y en general el uso del lenguaje, será pertinente (en principio) a la investigación de las reglas de la gramática» (1983, p. 214); la facultad del lenguaje está «dedicada al uso e interpretación del lenguaje [...] asume estados que varían en maneras limitadas con la experiencia. Interactuando con otros sistemas (cognitivo, sensoriomotor), estos estados contribuyen a la determinación del sonido y el significado de las expresiones» (2000, p. 168). Así, «cada expresión lingüística [el par FF, FL] generada por la lengua-I [...] incluye instrucciones para los sistemas de ejecución en los cuales

⁸ La explicación funcional (a la que, como se verá en la sección 3, podría corresponder la explicación ofrecida en la teoría de la FL) está en consonancia con la idea de que no hace falta leyes *especiales* para la explicación psicológica. Si de lo que se trata es de dar cuenta de por qué un mecanismo tiene cierta capacidad, entonces no hace falta apelar a leyes causales en el nivel computacional. Sin embargo, si de lo que se trata es de explicar la transición de estados en un sistema computacional, entonces ahí sí se baja al nivel físico para dar cuenta del paso de la causa al efecto (confróntese Cummins, 1983).

la lengua-I está incrustada. Es solo en virtud de esta integración en tales sistemas de ejecución que este estado cerebral califica como un lenguaje» (p. 27). De manera que, en varias ocasiones, Chomsky da a entender que la FL formaría parte de los mecanismos de ejecución del lenguaje.

Sin embargo, según Chomsky, un estudio naturalista «es una empresa humana particular que busca una clase especial de comprensión, obtenida por humanos en algunos pocos dominios cuando los problemas pueden simplificarse lo suficiente» (1995a, p. 10). La lingüística naturalista se encarga de dar cuenta de un ámbito bastante restringido de fenómenos cuya formulación es bastante precisa y cuyos resultados pueden someterse a evaluación empírica. En consecuencia, se ocupa del sistema cognitivo, esto es, de su estado inicial (GU) y su estado estable (lengua-I). Deja a un lado los sistemas de ejecución que constituyen el uso del lenguaje porque son de una complejidad extrema, en los que intervienen múltiples factores que no pueden ser aislados y simplificados para su estudio y, en esa medida, son poco propicios para una investigación naturalista (Chomsky, 1991, p. 40). Así, se encara el estudio de la FL como un sistema independientemente constituido, de manera que su explicación no es reducible a los sistemas de ejecución al tiempo que habría construcciones bien formadas que serían inutilizables por estos sistemas. Este sentido de autonomía de la FL, sin embargo, está constreñido al hecho de que la FL satisface las condiciones de legibilidad impuestas por los sistemas de ejecución. De manera que la distinción entre competencia y ejecución es más bien de orden metodológico y no proyecta ninguna distinción sustancial entre, lo que podría ser, conocimiento y mecanismo, o entre un sistema computacional abstracto o un mecanismo computacional causal.

Tal como Matthews recuerda, la distinción no alude a dos partes de la mente, sino que hay una sola FL que es «la facultad cognitiva de la producción y comprensión del lenguaje» (2008, p. 48). Esta facultad puede estudiarse desde dos perspectivas: acerca de qué es una capacidad (competencia) o cómo se logra esa capacidad y cómo se la utiliza en el curso de la producción y comprensión (ejecución). La función en intensión que va de ítems léxicos a pares de sonido-significado simplemente abstrae de especificaciones concretas relacionadas con la ejecución y, agregaría, de las propiedades estructurales relacionadas con su base física de realización. Hay un sentido en el que la mayoría de los modelos de la ciencia cognitiva están formulados en términos de funciones intensionales. En este sentido, la descripción del mecanismo de la FL es una descripción acerca de un mecanismo idealizado como toda idealización en ciencia. Por ejemplo, el modelo de producción de oraciones de Garrett (1982) —de cuyo estatus como modelo de ejecución nadie duda— está en el mismo nivel abstracto que el modelo chomskiano, esto es, está expresado en un diagrama de flujo abstraído

de cualquier realización física⁹. Y, en este sentido, tampoco tiene restricciones que dependen de tareas *on-line* o de factores como la memoria. Los modelos de las capacidades cognitivas en ciencia cognitiva, en general, tampoco los tienen (piénsese, por ejemplo, en el modelo de la visión de Marr, 1982). ¿Esto quiere decir que ningún modelo en ciencia cognitiva intenta dar una explicación causal de la arquitectura cognitiva? Si esto es (muy improbablemente) así: ¿Qué sería entonces una descripción de la arquitectura cognitiva causal? ¿Solo los modelos *neurofisiológicos* describirían mecanismos causales? Volveré sobre esta cuestión en el apartado 3.

El hecho de que la función en intensión abstraer de especificaciones concretas relacionadas con la ejecución, no implica que las gramáticas no especifiquen (de manera intensional) los pares que los hablantes/oyentes computan en el procesamiento del lenguaje. Aunque esta función (pares de sonido-significado) no es ninguna de las que se computan en el procesamiento (que son proyecciones de sonidos a significados o vice versa), la primera especifica la segunda, aunque bajo idealización y aproximación, en el sentido de que especifica el dominio y rango de la función que se computa en el curso de la comprensión y producción del lenguaje, de modo que el procesamiento debe respetar estos apareamientos. En este sentido, «las gramáticas *son* hipótesis empíricas acerca de procesos psicológicos, aunque hipótesis articuladas en un nivel de abstracción significativa» (Matthews, 2006b, p. 465). El propio Chomsky (1983) considera que las gramáticas están especificadas en el nivel computacional de Marr (1982), de modo que las teorías de la competencia serían teorías de nivel superior acerca de mecanismos psicológicos (confróntese Matthews, 2006a).

Las consideraciones realizadas en este apartado parecerían ir o abrir el camino a considerar el modelo minimalista como uno que permite ofrecer, o al menos podría integrar, una explicación causal de la FL. El hecho de que no puedan establecerse generalizaciones causales que abarquen el comportamiento lingüístico (ya sea por un naturalismo internalista o un mistericismo) no es óbice para la existencia de generalizaciones causales, aunque idealizadas, que abarquen el comportamiento del mecanismo específico subpersonal de la lengua-I. Y si bien el modelo no es en sí mismo utilizado como modelo de procesamiento debido a que está formulado en un nivel de «abstracción significativa», constituye una hipótesis empírica acerca de procesos cognitivos causales. Lo que quiero indagar en el apartado siguiente es en qué sentido un modelo de «abstracción significativa» o una especificación abstracta en términos funcionales de la FL puede considerarse una especificación de un mecanismo cognitivo y por qué a pesar de que la entidad responsable del fenómeno a dar cuenta es un mecanismo, su explicación no es mecanicista.

⁹ Aunque véase en el apartado 3 la dificultad de establecer a qué tipo de modelo funcional podría pertenecer el modelo minimalista chomskiano.

3. MODELOS FUNCIONALES Y MECANISMOS CAUSALES

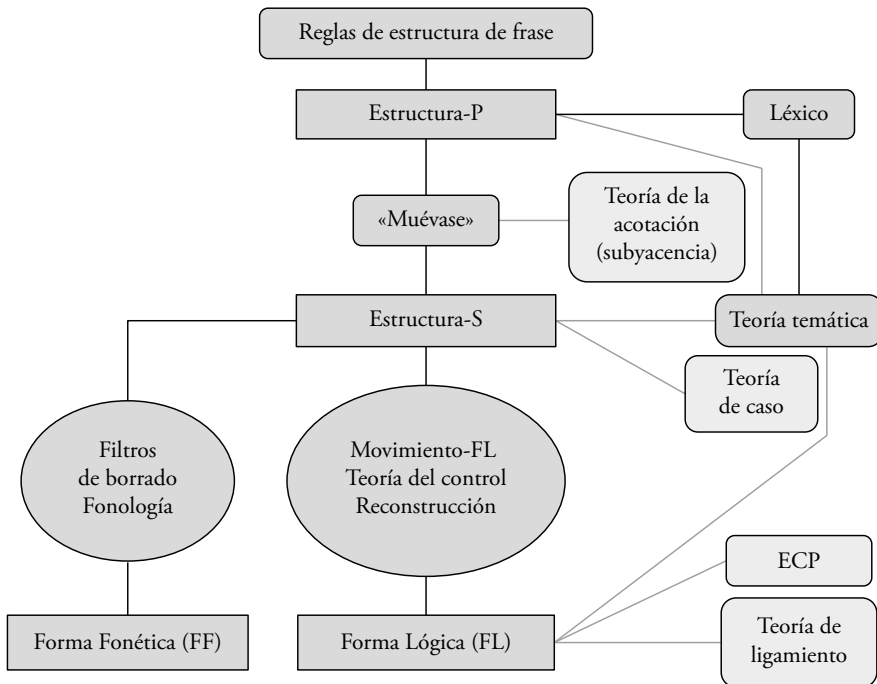
Aparentemente, todo el punto de la inferencia errónea que va de afirmar que puesto que la teoría de la FL no es una teoría acerca de aspectos causales/estructurales de la mente/cerebro, entonces no describe un mecanismo causal, radica justamente en el hecho de pensar que dado que solo describe propiedades funcionales de la entidad responsable de los fenómenos cognitivos a dar cuenta, entonces esta entidad no puede ser un mecanismo causal. Esta inferencia, tal como adelanté en el apartado titulado «Especificación abstracta y mecanismo casual», parece ser un caso de la inferencia general sostenida por algunos de los defensores de la explicación funcional en ciencia cognitiva en contra de los que adhieren a la explicación mecanicista. Los defensores de la explicación funcional consideran que la explicación psicológica no captura mecanismos (Fodor, 1968; Cummins, 1983 y Weiskopf, 2011). Su argumento se basa en que puesto que la explicación psicológica a partir de modelos cognitivos no es mecanicista, sino funcional, entonces no se trata de dar cuenta de mecanismos (es decir, entidades que poseen partes componentes, actividades y rasgos organizacionales).

Según los defensores de la explicación funcional, los modelos cognitivos intentan dar cuenta de una capacidad cognitiva en términos de las propiedades funcionales de un sistema y su organización. Para ello postulan un sistema representacional y un conjunto relevante de procesos que operan sobre el mismo, y recursos disponibles que interactúan con las operaciones¹⁰. Se da cuenta de la capacidad en cuestión en términos de la organización de subcapacidades, en otras palabras, de cómo está organizado el sistema en una secuencia de subprocesos que permiten a partir de sus *inputs* dar lugar a sus *outputs*. Los modelos cognitivos son un tipo de análisis o explicación funcional (Weiskopf, 2011). El *explanandum* de la explicación funcional, al igual que el de la mecanicista, es el hecho de que un sistema S tiene la capacidad para F (Cummins, 1983), y se explica esa capacidad en términos de sus propiedades funcionales y su organización. Siguiendo a Piccinini y Craver (2011), se pueden distinguir en la literatura sobre la explicación psicológica tres tipos de explicaciones funcionales dependiendo del tipo de propiedades funcionales que invoquen. El *análisis de tarea* consiste en la descomposición de una capacidad en subcapacidades y su organización, en este caso no se explica la capacidad para F en términos de la capacidad para G de las partes de S, sino en términos de la actividad de S mismo (Fodor, 1968 y Cummins, 1983). El *análisis funcional de estados internos* apela a la organización de estos estados y sus interacciones (Fodor, 1968). Y la *boxología* apela en la descomposición a procesos y organización de componentes funcionalmente individuados (1968).

¹⁰ Sigo a Weiskopf (2011) en la idea de que para los propósitos presentes alcanza con considerar a un modelo como un tipo de representación (que puede ser simbólica, gráfica, etcétera) de algún aspecto del mundo. Los modelos tienen una organización en términos de entidades, procesos y actividades que se corresponden de alguna manera con el mundo.

Ahora bien, el modelo de la FL, sin duda, especifica la capacidad del sistema cognitivo de la FL para aparear sonido-significado en términos de sus propiedades funcionales. Las propiedades funcionales de un sistema consisten en las «actividades o manifestaciones de sus poderes causales, disposiciones o capacidades» (Piccinini & Craver 2011, p. 291). Lo que no parece muy claro es el tipo de modelo funcional al que correspondería el modelo de la FL. Si se piensa en el modelo, anterior al minimalismo, de Principios y Parámetros (1981) articulado en torno a un subcomponente del sistema de reglas de la gramática y un subcomponente de principios, parece que es un modelo boxológico. En el primer componente se puede distinguir, a su vez, un léxico, la sintaxis (componente categorial y componente transformacional), el componente de la FF y el componente de la FL. El segundo abarcaría los principios de las teorías de la acotación, temática, rección, ligamiento, caso y control. Las cajas de la figura 1 (los rectángulos rectos) serían los componentes individuados funcionalmente en virtud de los *outputs* que producen dados ciertos *inputs*, y las flechas, como en todo modelo boxológico, reflejarían las relaciones funcionales/causales.

Figura 1



Extraída y traducida de las diapositivas de la clase de Hans C. Boas (2008). Construction Grammar in the 21st Century. Curso de doctorado Expanding Construction Grammar and Frame Semantics. <http://www.hf.uib.no/forskingskole/CxG%20in%20the%2021st%20Century.pdf>

Sin embargo, no queda claro que el modelo minimalista (Chomsky, 1995b), descrito en el apartado «La facultad del lenguaje», que está motivado por principios de diseño mínimo y óptimo, pueda configurarse en términos de un modelo boxológico. En favor de la lectura *boxológica*, se puede pensar que, al igual que en los modelos boxológicos, se postulan componentes que consisten en un sistema representacional (dado por las representaciones del léxico), y un conjunto relevante de procesos que operan sobre el mismo —en particular *Merge*— y recursos disponibles que interactúan con las operaciones. En este sentido, se da cuenta de la capacidad en cuestión en términos de la organización de subcapacidades, es decir, cómo está organizado el sistema en una secuencia de subprocesos que permiten a partir de sus *inputs* dar lugar a sus *outputs*.

Pero también es posible interpretar el modelo minimalista en términos de *análisis funcional de estados internos*. Dado el enfoque de la lengua-I como estados internos de la mente/cerebro del hablante/oyente, estos estarían individuados por sus propiedades funcionales, esto es, en relación a sus *inputs*, *outputs* y a otros estados internos. En este sentido, el estado inicial de la FL (la GU), los estados intermedios del procesamiento y el estado final alcanzado (la lengua-I), que serían estados globales del sistema, se explicarían en función de estados internos de los componentes del mecanismo computacional postulado. Si se está pensando en la especificación funcional de la FL en términos de un programa que está representado o corre en una computadora (Chomsky, 1983, p. 201), es natural pensar a los estados internos de esta como estados de sus componentes (por ejemplo, de su procesador o de su memoria). Así, la carga explicativa recaería en los componentes funcionales, al igual que en el análisis funcional en términos boxológicos.

Sin embargo, si se toman en cuenta las afirmaciones de Chomsky respecto de la teoría de la FL, parecería que está pensando más en una explicación funcional del tipo de *análisis de tarea*. Así, afirma: «Podemos conceder existencia a las estructuras postuladas en los estados inicial, intermedio y estable en precisamente el mismo sentido en que concedemos existencia a un programa que creemos está representado de alguna manera en una computadora o que postulamos para explicar la representación mental de un objeto tridimensional en el campo visual» (1983, p. 201). En este sentido, la cita se ajustaría a la descripción que realiza Cummins (1983) de las explicaciones psicológicas, denominadas *teorías de propiedades*. El objetivo de una teoría de propiedades es responder la pregunta ¿en qué consiste para el sistema S tener la propiedad P (o más precisamente, instanciar la propiedad P)? Para ello se apela a las propiedades de los componentes de S y cómo están organizadas.

La idea principal es explicar una capacidad cognitiva compleja de procesamiento de la información (que son propiedades disposicionales complejas) descomponiendo

el sistema en subcapacidades más simples y menos sofisticadas y problemáticas que la capacidad que se quiere explicar tal que la manifestación de las capacidades analizadoras (especificables en un programa) equivale a la manifestación de la capacidad analizada. Así, el primer paso del análisis funcional consiste en el análisis de la capacidad cognitiva P y el segundo en dar cuenta de la instanciación en S del análisis de P. De manera que si S posee los componentes organizados de la manera especificada en el análisis, entonces S posee P. La explicación funcional concluye cuando S ejecuta el programa especificado por el análisis (S tiene que tener la estructura necesaria para cumplir las instrucciones del programa). De manera que todo el peso explicativo de la explicación funcional radicaría, en última instancia, en la especificación de las capacidades analizadoras en el programa.

En este sentido, se puede pensar que cuando se describe el programa ejecutado por una computadora no se estaría describiendo ningún componente de la misma, de modo que no se daría ninguna información acerca del mecanismo subyacente. Sin embargo, la cita anterior de Chomsky prosigue: «La evidencia que conllevan hipótesis empíricas como esta podría derivar de muchas y variadas fuentes. En última instancia esperamos encontrar evidencia referente a los mecanismos físicos que llevan a cabo el programa [...]» (1983, p. 201). En este sentido, cuando se habla de una computadora que produce cierto comportamiento en virtud de ejecutar un programa, se quiere decir que sus estados son parte de algún componente de la computadora, puesto que si no estuvieran físicamente instanciados, no podrían ejecutarse y por ende no se generaría el comportamiento en cuestión. La afirmación de Chomsky alude a la idea de que la teoría de la FL es una especificación abstracta de un *sistema físico*. En este sentido, tal como sostiene Piccinini, «una descripción abstracta de un sistema físico no es una descripción de un objeto abstracto sino más bien una descripción de un sistema concreto que omite ciertos detalles» (2010, p. 11).

De esta manera, sea cual fuera la lectura del tipo de explicación funcional que ofrece la teoría de la FL, parece inevitable que esta se comprometa, al menos, con *componentes funcionalmente individuados* en virtud de sus actividades o manifestaciones de sus poderes causales, disposiciones o capacidades. Piccinini & Craver (2011, p. 291) sostienen que la «explicación mecanicista es la explicación de capacidades (funciones, conductas, actividades) de un sistema como un todo en términos de algunos de sus componentes, sus propiedades y capacidades (incluyendo sus funciones, conductas o actividades) y la manera en que están organizadas conjuntamente» (Bechtel & Richardson, 1993, Machamer, Darden & Craver 2000; Glennan, 2002). Este tipo de explicación involucra dar cuenta de un conjunto de fenómenos identificando el mecanismo que los produce, de manera que en el *explanans* de la explicación tiene que figurar la descripción del mecanismo responsable de los fenómenos, en este

caso, cognitivos. El *explanans* es así un modelo de las partes componentes, las actividades y los rasgos organizacionales del mecanismo que subyace al fenómeno del cual se quiere dar cuenta. Los componentes de un mecanismo tienen tanto propiedades estructurales (ubicación, forma, orientación y organización de los subcomponentes) como funcionales («sus actividades o manifestaciones de sus poderes causales, disposiciones o capacidades», Piccinini & Craver, 2011, p. 291). En este sentido, las partes reales o los componentes de un mecanismo pueden identificarse tanto por sus propiedades estructurales, como la organización espacio-temporal, el tamaño, etcétera, como también por sus propiedades funcionales, en términos de sus roles funcionales/causales.

Si bien en los modelos cognitivos se individualizan componentes funcionales, de modo que las subcapacidades están organizadas componencialmente (distintos niveles representan su propia información y procesos propios), no obstante, según Weiskopf, estas subcapacidades no son subpartes o componentes *reales* del sistema. Los modelos cognitivos suelen «postular elementos que no tienen un “eco” mecanicista: no se proyectan en partes del sistema realizador de ninguna manera obvia o directa» (2011, p. 332). Así, puede haber un mecanismo neural que subyace a los modelos cognitivos, pero los modelos cognitivos no capturan ese mecanismo. Lo que capturan es un nivel funcional de abstracción, no son modelos de «mecanismos reales en el mundo», de manera que «los modelos psicológicos no describen mecanismos» (p. 336).

Según Weiskopf, los componentes, las actividades y la organización que figuran en los modelos funcionales, a diferencia de los mecanicistas, no tienen una relación directa de correspondencia con los componentes, las actividades y la organización del sistema en que se realizan. De modo que se relacionan de manera sumamente indirecta, en el sentido de que «no se puede de una manera simple o llana leer la presencia del estado de nivel superior a partir del estado de nivel inferior» (p. 328). La naturaleza de la proyección entre las propiedades funcionales y estructurales es a menudo opaca ya sea porque no hay un sistema físico bien demarcado para descomponer en partes o no tenemos idea de qué partes u operaciones usar en las descomposiciones. Sin embargo, hay técnicas que permiten ofrecer enfoques indirectos¹¹. Los modelos cognitivos que hacen uso de estas técnicas pueden, según Weiskopf, ser perfectamente

¹¹ Weiskopf (2011) menciona la reificación (postular, por ejemplo, entidades con características de un objeto estable y durable, cuando de hecho no existe algo como eso), la abstracción funcional (descomponer un sistema en subsistemas en función de lo que hacen y no de su correspondencia con la organización del sistema en el cual se realizan) y la ficcionalización (incluir componentes en el modelo, que se sabe que no corresponden a ningún elemento en el sistema modelado pero que cumplen un rol esencial para que el modelo sea operativo).

adecuados o legítimos aun cuando en el nivel neurofisiológico no haya entidades con las características postuladas por el modelo. Lo importante para ser modelos adecuados es que satisfagan los criterios habituales para una explicación adecuada¹².

De modo que los modelos cognitivos difieren de los mecanicistas en la manera en la que se corresponden con el sistema modelado. Mientras que en un modelo mecanicista (completamente desarrollado) todos los elementos que figuran se corresponden con partes reales del mecanismo, en los modelos cognitivos «no todo lo que cuenta como un componente desde el punto de vista del *modelo* se parecerá a un componente en el *sistema* modelado mismo —al menos no, si nuestra noción de componente se basa en una entidad física distintiva relativamente localizada como una columna cortical, cadena de ADN, ribosoma, o canal iónico—» (Weiskopf, 2011, p. 332). Un eco de este argumento se encuentra en Collins cuando sostiene que claramente hay estados y procesos cerebrales que realizan la lengua-I y la teoría de la FL es, en un cierto grado de abstracción, acerca de esos procesos pero, sin embargo, «no imaginamos una proyección nítida entre los elementos y operaciones postulados de la FL y los procesos físicos —no hay una cinta en la cabeza con símbolos en ella—» (2006, p. 498).

Ahora bien, este argumento de Weiskopf —que sostiene que si los componentes y actividades del modelo cognitivo se corresponden de manera indirecta con los componentes y actividades del sistema modelado, se sigue que los modelos cognitivos no capturan la estructura mecanicista de partes «reales», y en este sentido, no describen mecanismos— supone ligar la noción de «mecanismo» con «localización física bien demarcada» de partes. Ejemplos de componentes reales son los que se han mencionado antes: una columna cortical, cadena de ADN, un ribosoma, etcétera. Sin embargo, también se ha mencionado que se pueden dar definiciones de qué es un mecanismo tanto en términos funcionales como estructurales. De hecho, Glennan afirma que un mecanismo es «un sistema complejo que produce ese comportamiento por la interacción de un número de partes, en donde la interacción entre las partes puede caracterizarse por generalizaciones que están relacionadas con cambios, son invariantes y directas» (2005, p. 445). Tomando en cuenta la distinción de Bechtel & Richardson (1993) entre localización (la identificación de partes del mecanismo por sus localización espacial) y descomposición (la identificación de partes del mecanismo por sus relaciones funcionales), la definición de Glennan está expresada en términos de relaciones funcionales entre partes. De manera que, en las propias palabras de este autor, «es la estructura funcional revelada por la descomposición lo que es

¹² Criterios de apoyo evidencial, corrección representacional, explicación genuina (no meramente descriptiva o predictiva), y plausibilidad y consistencia con el resto de nuestro conocimiento.

constitutivo del mecanismo» (p. 447)¹³. Por otro lado, se pueden tener definiciones de qué es un mecanismo que enfatizan sus propiedades estructurales. Así, los componentes o partes simplemente son las entidades en un mecanismo físico. Por ejemplo, si se trata de dar cuenta de los potenciales de acción, los componentes del mecanismo serán los canales de Na^+ y K^+ , iones y cadenas de proteínas (Craver, 2006). Mientras que los componentes estructurales se especifican utilizando técnicas anatómicas que permiten caracterizar diferentes regiones del cerebro en función del tipo de neuronas y sus relaciones, los funcionales se especifican en términos de sus efectos bajo ciertas circunstancias.

Asimismo, según Piccinini y Craver, la idea de partes reales en sentido estructural no implica que estas sean «localizables espacialmente de manera nítida, que tengan una única función, que sean estables y no cambien, o que carezcan de relaciones de retroalimentación dinámicas o complejas con otros componentes. De hecho, un componente estructural puede estar tan distribuido y difuso como para desafiar una prolija descripción estructural [...]» (2011, p. 291). En este sentido, la mayoría de los mecanicistas no defienden relaciones de correspondencia isomórficas sino que los modelos que representan mecanismos mantienen relaciones de similitud en varios grados y aspectos (Glennan, 2005, p. 449). No obstante, Weiskopf (2011) argumenta que debilitar la restricción de organización espacial permitiendo que esta esté distribuida en partes no estrictamente localizables, vacía de contenido a la noción de «mecanismo» puesto que imposibilita identificar mecanismos y sus características particulares. Considero que la cuestión de la organización espacial claramente delimitada como criterio primordial para la identificación de mecanismos no se puede dirimir con independencia de la evidencia empírica. De hecho, la evidencia empírica actual parece indicar que muchas funciones cognitivas están localizadas en vastas áreas. Sin embargo, esto no obstaculiza su investigación. En este sentido, no parece que esta evidencia vacíe de contenido a la noción de «mecanismo», siempre y cuando haya un conjunto de criterios que al actuar conjuntamente puedan servir de guía para delimitar mecanismos.

De manera que, contrariamente a la posición de Weiskopf, puede haber explicaciones psicológicas de mecanismos cognitivos que sean completamente funcionales, en el sentido de que solo describan las propiedades funcionales del mecanismo, haciendo abstracción de las propiedades estructurales. La insistencia de Weiskopf en que las explicaciones funcionales no capturan mecanismos va ligada a su idea de que *mecanismo*, en el contexto de esta polémica, es sinónimo de *mecanismo neural*.

¹³ Otra definición de «mecanismo» en términos puramente funcionales puede encontrarse, por ejemplo, en Piccinini (2008).

Así, se entienden afirmaciones como la siguiente: «puede haber un sistema neural mecanicista subyacente, pero esta estructura mecanicista no es lo que los modelos cognitivos capturan» (2011, 333). Este tipo de afirmación se asemeja en mucho a las afirmaciones de Collins respecto de la FL. Collins afirma que «la lingüística no está en el negocio de caracterizar un proceso físico» (2006, p. 497) y que «la gramática no es una descripción de ningún proceso físico, sino que el output de la gramática debe estar estructurado de tal manera que un proceso físico pueda generar las estructuras» (p. 498). Así, según Collins, si bien hay estados o procesos del cerebro que realizan la competencia lingüística y la teoría de la FL es, en cierto modo, acerca de esos procesos, en la medida en que es una caracterización abstracta no capturaría un mecanismo.

Sin embargo, hasta aquí no parece haber razones de peso para considerar, tal como sostienen los defensores de la explicación funcional, que la abstracción funcional no capture propiedades o componentes reales de un mecanismo (cognitivo o neural o como se lo considere) y, consecuentemente, no habría razones de peso para afirmar que la teoría de la FL entendida en términos funcionales, no describe un mecanismo. Aunque, tal como mencioné al comienzo, no es mi intención mostrar que este enfoque arquitectural sea el que de hecho asume Chomsky (sino que se alinea con sus desarrollos), es inevitable acudir a algunas, de entre muchas, de sus afirmaciones que remarcan claramente que considera a la FL como un mecanismo. Por ejemplo, «la lingüística es [...] el estudio abstracto de ciertos *mecanismos*, de su desarrollo y maduración» (1983, p. 201; las cursivas son mías). Asimismo, «nosotros nos ocupamos de las condiciones abstractas que deben satisfacer *mecanismos* desconocidos» (p. 211; las cursivas son mías).

Lo que intenté mostrar hasta aquí es en qué sentido se puede considerar que una especificación abstracta en términos funcionales, en sentido general y aplicada a la FL, puede considerarse una especificación de un mecanismo cognitivo. Ahora bien, el hecho de que la entidad responsable de la producción de los fenómenos a explicar sea un mecanismo (o, al menos, que no haya razones de peso para considerar que no lo es) no implica, a diferencia de lo que sostienen los defensores de la explicación mecanicista, que la explicación deba ser mecanicista. Piccinini y Craver consideran que la explicación psicológica es un tipo de explicación mecanicista. Dado que las propiedades funcionales son un aspecto de la explicación mecanicista, el análisis funcional no puede ser autónomo sino que constituye una forma elíptica de explicación mecanicista (lo que denominan «esbozo de explicación mecanicista») en la que faltan algunos aspectos estructurales que al llenarse convierten a la explicación en una mecanicista completamente desarrollada. Según esta concepción, puesto que «el *target* de descripción en cada caso es un mecanismo» (2011, p. 290), tanto si se pone énfasis en la descripción de sus propiedades estructurales como funcionales, la explicación es mecanicista.

Sin embargo, siguiendo a Weiskopf (2011), si lo que determina que los modelos funcionales y mecanicistas sean tipos distintos de modelos es la relación de correspondencia indirecta o directa, respectivamente, que hay entre los componentes, las actividades y la organización postulados en el modelo y los del sistema modelado, entonces se podría sostener que aunque la entidad responsable sea un mecanismo, dada la relación indirecta que se da entre el modelo cognitivo y el mecanismo, no puede considerarse que la explicación sea mecanicista. De manera similar, el hecho de que la FL sea un mecanismo no implica que la explicación brindada por la teoría de la FL sea mecanicista, esto es, sea una explicación funcional a la que le faltan detalles estructurales (es decir, un esbozo de explicación mecanicista) que una vez llenados la convertirán en una explicación mecanicista completamente desarrollada.

La mayoría de las veces, los componentes, las actividades y la organización que figuran en los modelos cognitivos no tienen una relación directa de correspondencia con los componentes, las actividades y la organización del sistema físico en el que se realizan. De modo que se relacionan indirectamente, en el sentido mencionado de que «no se puede de una manera simple o llana leer la presencia del estado de nivel superior a partir del estado de nivel inferior» (Weiskopf, 2011, p. 328). Si hay partes nítidamente localizables, entonces es posible, como en muchos casos en neurobiología, diseccionar el mecanismo en partes por medio de métodos observacionales y experimentales. En el caso en que no haya partes nítidamente localizables, como en los mecanismos cognitivos de nivel superior, se pueden identificar mecanismos y estudiar sus propiedades por medio de métodos de inferencia indirecta, como los que menciona Glennan (2005), que testean por ejemplo la organización temporal de los procesos o se pueden variar las condiciones (precipitantes, de inhibición, de trasfondo, etcétera) del fenómeno a dar cuenta.

Esto mismo es lo que parece sugerir Chomsky (1983) cuando trata la cuestión de la evidencia empírica indirecta que se puede obtener en relación a la teoría de la FL, estableciendo una analogía con el problema de determinar la naturaleza de las reacciones termonucleares en el interior del sol. Aun cuando las técnicas disponibles solo permitieran estudiar la luz que emite en sus capas exteriores, los científicos podrían conjeturar la fusión de elementos ligeros que forman elementos más pesados que convierten masa en energía y así dar cuenta de las reacciones termonucleares ocultas. La evidencia ofrecida por la luz emitida por la periferia solar es evidencia indirecta. No hay manera de instalarse en el interior del sol para obtener pruebas directas. De manera similar, la investigación de la FL hace uso de evidencia indirecta (tiempo de reacción, reconocimiento, juicios de informantes, etcétera) esperando que esta refiera a mecanismos subyacentes. La diferencia en la analogía entre las postulación de entidades y procesos físicos, para el caso del sol, y la especificación

abstraída de propiedades estructurales del cerebro para el caso del lenguaje no constituye, sin embargo, una diferencia de principio. Simplemente, por razones éticas no podemos explorar los mecanismos por experimentación directa, pero esto no trae ningún problema particular para la investigación de las capacidades cognitivas (en todo caso, sería análogo al caso del astrónomo que solo puede estudiar las emisiones de luz provenientes de la periferia del sol).

De modo que el tipo de explicación no dependería del tipo de entidad responsable del fenómeno por dar cuenta (es decir, si es o no un mecanismo) sino del tipo de acceso epistémico (directo o indirecto) a dicha entidad. Si la relación de correspondencia o proyección entre las partes, las actividades y la organización postuladas en el modelo es directa en relación a las partes, las actividades y la organización del mecanismo, entonces, hay buenas chances de obtener una explicación mecanicista. Esto es lo que se espera que ocurra, en general, en los casos de múltiples niveles de mecanismos neurobiológicos o de mecanismos cognitivos de nivel básico. Si, en cambio, la correspondencia es indirecta, como en el caso de los mecanismos cognitivos de nivel superior, la explicación funcional parece ser la más adecuada. Esto no significa necesariamente que la explicación funcional sea un esbozo de explicación mecanicista que una vez llenados los aspectos estructurales se convierta en una explicación mecanicista completamente desarrollada. Esto equivaldría quizá a una especie de reducción, en este caso, del nivel lingüístico al neurofisiológico. Puede que esos aspectos no puedan ser llenados por razones de acceso epistémico, dado justamente que «no se puede de una manera simple o llana leer la presencia del estado de nivel superior a partir del estado de nivel inferior» (Weiskopf, 2011, p. 328) o que la «reducción es una bestia rara» (Collins, 2006, p. 501) o que la evidencia de la que dispondremos, ya sea por razones éticas o de otra índole, será siempre indirecta (Chomsky, 1983).

4. CONCLUSIÓN

Lo que intenté mostrar es que la inferencia que parte de la afirmación de que la teoría de la FL no ofrece explicaciones causales de la mente/cerebro para concluir que su objeto de estudio, la FL, no es un mecanismo cognitivo causal está injustificada. Esta inferencia en el ámbito de la filosofía de la lingüística no es más que un caso de una inferencia más general que suele estar bastante difundida en la filosofía de la ciencia cognitiva por parte de los defensores de la explicación funcional. En este sentido, intenté defender la idea de que es perfectamente aceptable sostener que cierta teoría o modelo solo ofrece una explicación que alude a las propiedades funcionales de la entidad responsable de los fenómenos cognitivos que se desean explicar pero que, sin embargo, el objeto de estudio de esa teoría es un mecanismo que posee tanto

propiedades funcionales como estructurales que manifiestan disposiciones causales que permiten insertarlo en una red causal. Este es el sentido en el que creo que la versión arquitectural causal de la FL es más iluminadora de la concepción chomskiana de que la lingüística es el estudio abstracto de mecanismos de la mente/cerebro y que el modelo de la FL es una hipótesis empírica acerca de procesos psicológicos. En consecuencia, no hace falta negar que la explicación cognitiva capture mecanismos, como hacen Weiskopf y Collins, para defender alguna forma de explicación funcional psicológica o especificación abstracta ni insistir en que la explicación psicológica es mecanicista porque de lo que se trata es de dar cuenta de mecanismos, como hacen Piccinini y Craver¹⁴.

REFERENCIAS BIBLIOGRÁFICAS

- Baker, Mark (2001). *The Atoms of Language. The Mind's Hidden Rules of Grammar*. Nueva York: Basic Books.
- Bechtel, William & Robert Richardson (1993). *Discovery Complexity: Decomposition and Localization as Scientific Research Strategies*. Princeton: Princeton University Press.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1983). *Reglas y representaciones*. México DF: FCE.
- Chomsky, Noam (1991). Linguistics and Cognitive Science: Problems and Mysteries. En Asa Kasher (ed.), *The Chomskyan Turn* (pp. 26-53). Oxford: Basil Blackwell.
- Chomsky, Noam (1995a). Language and Nature. *Mind*, 104(413), 1-61.
- Chomsky, Noam (1995b). *The Minimalist Program*. Cambridge: The MIT Press.
- Chomsky, Noam (2000). *New Horizons in the Study of Language and Mind*. Cambridge: The MIT Press.
- Collins, John (2004). Faculty Disputes. *Mind & Language*, 19(5), 503-533.

¹⁴ Versiones anteriores de partes de este trabajo han sido presentadas en 2011 en el Coloquio Instituto de Investigaciones Filosóficas (UNAM)-SADAF (Buenos Aires) y XX Jornadas de Epistemología e Historia de la Ciencia (Córdoba), y en 2012 en el IV International Workshop Language, Context and Cognition (Pisac) y el II Congreso Latinoamericano de Filosofía Analítica (Buenos Aires). Agradezco los comentarios de José Ahumada, Axel Barceló, Natalia Buacar, Maite Ezcurdia, Toni Gomila, Sabrina Haimovici, Eleonora Orlando y Rob Stainton. Asimismo, agradezco especialmente los comentarios escritos de Ángeles Eraña y Max Freund a versiones expuestas en esos eventos, y el de Mariela Destéfano a una versión anterior. Este trabajo se ha beneficiado con el apoyo financiero de los proyectos UBACyT 20020110100188 (2012-2015), PICT-2010 0545 de ANPCyT (2011-2014), y PIP-CONICET 2531 (2009-2011).

- Collins, John (2006). Between a Rock and a Hard Place: A Dialogue on the Philosophy and Methodology of Linguistics. *Croatian Journal of Philosophy*, 6, 471-505.
- Collins, John (2008). Knowledge of Language Redux. *Croatian Journal of Philosophy*, 7, 3-42.
- Craver, Carl (2006). When Mechanistic Models Explain. *Synthese*, 153(3), 355-376.
- Cummins, Robert (1983). *The Nature of Psychological Explanation*. Cambridge: The MIT Press.
- Dwyer, Susan & Paul Pietroski (1996). Believing in Language. *Philosophy of Science*, 63(3), 338-373.
- Fodor, Jerry (1968). *Psychological Explanation*. Nueva York: Random House.
- Fodor, Jerry (1975). *The Language of Thought*. Nueva York: Harper & Row.
- Fodor, Jerry (1983). *The Modularity of Mind*. Cambridge: The MIT Press.
- Fodor, Jerry (2001). Doing Without what's Within: Fiona Cowie's Critique of Nativism. *Mind*, 110(437), 99-148.
- Garret, Merrill F. (1982). Production of Speech: Observations from Normal and Pathological Language Use. En Andrew Ellis (ed.), *Normality and Pathology in Cognitive Functions* (pp. 19-76). Londres: Academic Press.
- Glennan, Stuart (2002). Rethinking Mechanistic Explanation. *Philosophy of Science*, 69, 342-353.
- Glennan, Stuart (2005). Modeling Mechanisms. *Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 445-464.
- Knowles, Jonathan (2000). Knowledge of Grammar as a Propositional Attitude. *Philosophical Psychology*, 13(3), 325-53.
- Machamer, Peter, Lindley Darden & Carl Craver (2000). Thinking about Mechanisms. *Philosophy of Science*, 67, 1-25.
- Marr, David (1982). *Vision*. San Francisco: W. H. Freeman.
- Matthews, Robert (2006a). Could Competent Speakers Really Be Ignorant of Their Language? *Croatian Journal of Philosophy*, 6(3), 459-69.
- Matthews, Robert (2006b). Knowledge of Language and Linguistic Competence. En Ernest Sosa y Enrique Villanueva (eds.), *Philosophy of Language: Philosophical Issues*, 16 (pp. 200-220). Oxford: Blackwell.
- Matthews, Robert (2008). Epistemic Heresies: Reply to John Collins' «Redux». *Croatian Journal of Philosophy*, 22, 45-55.
- Piccinini, Gualtiero (2008). Computers. *Philosophical Quarterly*, 89(1), 32-73.

- Piccinini, Gualtiero (2010). Computation in Physical Systems. En Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/fall2010/entries/computation-physicalsystems/>
- Piccinini, Gualtiero & Carl Craver (2011). Integrating Psychology and Neuroscience: Functional Analyses as Mechanism Sketches. *Synthese*, 186, 283-311.
- Skidelsky, Liza (2006). Personal-Subpersonal: The Problems of Inter-level Relations. *Protosociology*, número especial: *Compositionality, Concepts and Representations II: New Problems in Cognitive Science*, 22, 120-139.
- Skidelsky, Liza (2013). Competencia lingüística, actitudes proposicionales y representaciones. En Adriana Gonzalo, Griselda Parera y Cintia Carrió (comps.), *El camino desde Syntactic Structures. Un recorrido histórico y filosófico por la lingüística chomskiana* (pp. 145-164). Buenos Aires: Universidad Nacional del Litoral.
- Weiskopf, Daniel (2011). Models and Mechanisms in Psychological Explanation. *Synthese*, 186, 313-338.

SOBRE LOS AUTORES

Andrés Abugattás es licenciado en filosofía por la PUCP. Estudia en la maestría de filosofía de la misma universidad, donde también es asistente de docencia. Se interesa por la filosofía del lenguaje y de la mente.

Clark Barrett es un antropólogo biológico especializado en la evolución de la cognición humana. Sus investigaciones utilizan comparaciones interculturales para examinar hipótesis en torno a mecanismos cognitivos especializados en el aprendizaje, la toma de decisiones e inferencias en relación con la búsqueda de comida, el reconocimiento de peligro, la teoría de la mente y el razonamiento moral. Su trabajo teórico examina la evolución de la modularidad y el rol de los procesos de desarrollo en la formación de la cognición de dominio específico. Actualmente, es profesor asociado del Departamento de Antropología de la Universidad de California en Los Ángeles.

Ricardo Braun se graduó *summa cum laude* en filosofía por la Universidad de Texas. Es magíster y candidato a doctor en filosofía por la PUCP y licenciado en educación por la Universidad Cayetano Heredia. Realizó estudios doctorales de filosofía en Queen's University, en Ontario, y ha sido investigador visitante en la Facultad de Psicología de la Universidad Católica de Lovaina. Es profesor asociado de la Universidad de Lima en la Facultad de Psicología. Sus temas de interés son la epistemología, la filosofía de la ciencia y la filosofía de la psicología. Ha publicado los libros *El problema mente-cuerpo: un enfoque neurofisiológico* (1989), *Elementos de lógica simbólica* (1992), *Introducción a la psicología* (1996) en autoría con Diógenes Rosales, y *¿Qué soy yo? Una introducción a la filosofía de la mente y de la psicología* (2009) y numerosos artículos sobre el problema de la mente en diversas revistas especializadas.

Peter Carruthers es profesor de filosofía en la Universidad de Maryland, miembro asociado del Programa de Neurociencia y Ciencia Cognitiva y del Comité para la Filosofía y las Ciencias en la misma universidad. Sus investigaciones se centran en la filosofía de la mente, la filosofía de la psicología y las ciencias cognitivas. Sus contribuciones más importantes tratan la teoría de la conciencia, el conocimiento de nuestras propias actitudes proposicionales, el rol del lenguaje natural en la cognición humana y la modularidad de la mente. Sus libros más recientes son: *The Opacity of Mind: An Integrative Theory of Self-Knowledge* (2011), *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought* (2006), *Consciousness: Essays from a Higher-Order Perspective* (2005) y *The Nature of the Mind: An Introduction* (2004).

Paola Cépeda es licenciada en Lingüística Hispana, y magíster y doctora en Lingüística, por la PUCP. Es docente a tiempo completo del Departamento de Humanidades de esta misma universidad. Sus intereses giran en torno de la lingüística teórica y el estudio formal de las lenguas, en especial, el castellano; el desarrollo de habilidades letradas académicas; y la relación entre el lenguaje, el cerebro y la mente. Forma parte del Grupo Textus, dedicado a la investigación en competencias comunicativas, y del Grupo Lingüística y Educación, interesado en el estudio del lenguaje y sus aplicaciones en diferentes niveles de formación educativa. Ha sido coordinadora del Grupo Mente y Lenguaje entre enero de 2010 y julio de 2011.

Robin Dunbar es profesor de Psicología Evolucionista en el Magdalen College de la Universidad de Oxford. Fue elegido miembro de la Academia Británica en 1998. Sus intereses principales de investigación se centran en la evolución de la socialización con énfasis particular en los primates y humanos. Sus contribuciones más conocidas son la denominada « hipótesis del cerebro social», la teoría del *gossip* en la evolución del lenguaje y el «número de Dunbar» (el límite en el número de relaciones que un individuo puede manejar). Sus libros científicos más conocidos son *Grooming, Gossip and the Evolution of Language* (1997), *The Human Story* (2004), *How Many Friends Does One Person Need? Dunbar's Number and Other Evolutionary Quirks* (2010) y *The Science of Love and Betrayal* (2012).

César Escajadillo es magíster en Filosofía por la PUCP. Actualmente se desempeña como profesor de los cursos de Teoría del Conocimiento en la Universidad Peruana de Ciencias Aplicadas, Historia de la Filosofía Contemporánea en la Universidad Antonio Ruiz de Montoya y Advenimiento del Sujeto en el Centro de Psicoterapia Psicoanalítica de Lima. Es autor del libro *Pensamiento y acción. La filosofía peruana a comienzos del siglo XX* (2009) con Pablo Quintanilla y Richard Antonio Orozco. Sus intereses de investigación se centran en la filosofía del lenguaje y la mente, el pragmatismo y la epistemología. Desde enero de 2013 es el coordinador del Grupo Mente y Lenguaje.

Silvia Español es doctora en Psicología por la Universidad Autónoma de Madrid. Actualmente se desempeña como investigadora independiente del Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina. Su área de especialidad es el desarrollo sociocognitivo en la infancia temprana. Sus trabajos se ubican en la frontera entre la psicología cognitiva del desarrollo, la psicología de la música y la filosofía de la mente. Es profesora de las maestrías en Psicología Cognitiva de la Universidad de Buenos Aires y en Psicología Cognitiva y Aprendizaje de la Facultad Latinoamericana de Ciencias Sociales (FLACSO). Cuenta con numerosas publicaciones sobre su área de especialidad. Es miembro de la Sociedad Argentina de Análisis Filosófico (SADAF) y vicepresidenta actual de la Sociedad Argentina de Ciencias Cognitivas para la Música (SACCOM).

María de los Ángeles Fernández Flecha es licenciada en Lingüística por la PUCP y doctora *cum laude* por el programa de Intervención en el Lenguaje de la Universidad Complutense de Madrid. Es docente de cursos de teoría del lenguaje y metodología de la investigación del Departamento de Humanidades de la PUCP. Sus intereses académicos y de investigación abarcan el estudio del proceso de adquisición de lenguas maternas y, de forma específica, el desarrollo ontogenético de la comunicación en el periodo prelingüístico y lingüístico temprano. Colabora con el Grupo de Investigación Equipo de Investigación sobre la Adquisición del Lenguaje (EQUIAL) de la Universidad Complutense de Madrid.

Antoni Gomila es catedrático de Psicología Básica del Departamento de Psicología de la Universidad de las Islas Baleares. Anteriormente trabajó como profesor titular de Lógica y Filosofía de la Ciencia en las Universidades de Salamanca y La Laguna. De formación inicial filosófica, su interés se ha centrado en las cuestiones teóricas de la psicología, como la naturaleza del significado y la representación mental, la expresión y la intersubjetividad, o la explicación de la conducta humana, tomando siempre como punto de partida una perspectiva evolutiva. Fue becario posdoctoral Fulbright en la Rutgers University, donde fue discípulo de Jerry Fodor (1990-1991), y disfrutó de un año sabático en la School of Psychology de la University of St. Andrews (2009-2010).

Marcos Herrera es psicoanalista y lingüista. Es licenciado en psicología clínica por la PUCP y obtuvo *summa cum laude* en el grado de doctor en Lingüística Románica por la Albert-Ludwigs Universität Freiburg, con una beca del Servicio Alemán de Intercambio Académico. Actualmente desarrolla paralelamente su actividad académica como profesor asociado a medio tiempo del Departamento de Humanidades de la PUCP y su actividad profesional como psicoanalista miembro de la Sociedad Peruana de Psicoanálisis. Es autor de varios artículos publicados en revistas especializadas.

Carla Mantilla es magíster en estudios psicoanalíticos teóricos por la Universidad de Londres y tiene un diploma de estudios superiores de filosofía por la PUCP. Es licenciada en Psicología Clínica por esta misma universidad y candidata a psicoanalista del Instituto Peruano de Psicoanálisis. Es docente de la especialidad de Psicología y de la maestría en Estudios Teóricos de Psicoanálisis en la PUCP. Asimismo, se desempeña como psicoterapeuta clínica en práctica privada. Actualmente sus intereses de investigación giran en torno a la eficiencia de los procesos terapéuticos con adultos y a la interface entre psicoanálisis y filosofía de la mente. Ha sido la coordinadora del Grupo Mente y Lenguaje en el año 2009.

Gabriel Martínez Vera es bachiller en Lingüística en la PUCP. En el año 2011 obtuvo el Diploma en Humanidades por la Universidad Antonio Ruiz de Montoya y el Diploma de *EFL Teachers* del Centro de Idiomas de la Universidad del Pacífico. Actualmente trabaja como profesor de inglés y lenguaje en la Universidad del Pacífico y asistente de docencia de los cursos de Argumentación, Estructura del Lenguaje e Introducción a la Lengua y la Literatura en la PUCP. Ha publicado diversos artículos en la revista *Sorda y Sonora*.

Luis Manuel Olguín es bachiller y magister en lingüística por la PUCP, donde también es asistente de docencia. Sus intereses de investigación se centran en la relación entre lenguaje y cognición desde la lingüística cognitiva, la semántica cognitiva y la psicología discursiva.

Diana Pérez es doctora en Filosofía por la Universidad de Buenos Aires (UBA). Es profesora de los cursos de Metafísica y de Fundamentos de Filosofía en el Departamento de Filosofía, de las maestrías en Psicología Cognitiva de la UBA y Psicología de la Música en la Universidad Nacional de La Plata. Es investigadora independiente del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Ha publicado dos libros, cuatro compilaciones y numerosos artículos sobre temas de filosofía de la psicología y metafísica de la mente en revistas especializadas del país y del extranjero. Ha presentado numerosos trabajos y conferencias en el país y en el exterior y ha sido galardonada con el premio Bernardo Houssay 2003, el diploma al mérito Konex 2006 y el diploma de reconocimiento UBA 2008. Actualmente, es presidenta de la SADAF.

Jorge Iván Pérez Silva es doctor en Lingüística por la Universidad de Cornell y profesor principal del Departamento de Humanidades de la PUCP. Sus intereses académicos incluyen el contacto lingüístico, la fonología, la sintaxis y la semántica. Ha publicado los documentales audiovisuales *Los castellanos del Perú* (2004), *Las lenguas del Perú* (2008) y *Perusuyunchikpa qichwasiminkuna: Los quechuas del Perú* (2011),

con Luis Andrade, Roberto Zariquiey y Virginia Zavala, así como los libros *Manual de gramática del castellano* (2004), con Karen Coral; *Contra el prejuicio lingüístico de la motosidad* (2008), con Jorge Acurio y Raúl Bendezú, y *Las lenguas del Perú* (2009) con Luis Andrade.

Pablo Quintanilla es PhD en Filosofía por la Universidad de Virginia y MA en Filosofía por el King's College de la Universidad de Londres. Actualmente, es profesor principal de filosofía en la PUCP. Sus áreas de especialización son la filosofía de la mente y del lenguaje, la epistemología, la teoría de la acción y el pragmatismo clásico y contemporáneo. Es autor de una serie de artículos académicos sobre temas de su especialidad y del libro *La comprensión del otro. Ensayos sobre interpretación y racionalidad* (de próxima aparición). Con César Escajadillo y Richard Antonio Orozco es autor de *Pensamiento y acción. La filosofía peruana a comienzos del siglo XX* (2009). Ha editado el libro *Ensayos de metafilosofía* (2009) y coeditado *Desarrollo humano y libertades: una aproximación interdisciplinaria* (2009). Fue coordinador del Grupo Mente y Lenguaje (2008-2010).

Carolina Romero es licenciada en psicología clínica por la PUCP y sigue la formación para psicoterapeuta psicoanalítica en el Centro de Psicoterapia Psicoanalítica de Lima. Asimismo, ha realizado estudios avanzados de neurociencias en el Instituto de Neurociencias Aplicadas del Perú. Actualmente trabaja como psicoterapeuta en práctica privada y en el centro de psicoterapia con niños Vinculare.

Liza Skidelsky es profesora del Departamento de Filosofía de la UBA e investigadora independiente del CONICET. Es doctora en Filosofía por la UBA. Ha realizado estudios postdoctorales en la Universidad de Nueva York y ha sido investigadora visitante en la Universidad de Massachusetts y en el Centro de Ciencias Cognitivas de la Universidad de Rutgers. Sus áreas de investigación son la metafísica y filosofía de la mente, y la filosofía de la ciencia cognitiva. Dirige proyectos y becarios de investigación de la UBA, CONICET y la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT). Ha coeditado *Cuestiones mentales. Debates filosóficos contemporáneos* (2009) y *Conceptos. Debates contemporáneos en filosofía y psicología* (2010), así como numerosos artículos en libros y revistas especializadas.

Se terminó de imprimir en
los talleres gráficos de
Tarea Asociación Gráfica Educativa
Psje. María Auxiliadora 156, Breña
Correo e.: tareagrafica@tareagrafica.com
Teléfono: 332-3229 Fax: 424-1582
Se utilizaron caracteres
Adobe Garamond Pro en 11 puntos
para el cuerpo del texto
abril 2014 Lima - Perú