



UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE FILOSOFÍA Y HUMANIDADES
ESCUELA DE FILOSOFÍA

TRABAJO FINAL DE LICENCIATURA EN FILOSOFÍA

Naturalismo e Identidad Personal:

La propuesta narrativista de D. Dennett

Estudiante: Malena León

Directora: Silvia Carolina Scotto

Co-Directora: Laura Danón

Córdoba

2017

Agradecimientos

Aunque esta investigación no tenga un gran valor en sí misma, no deja de ser el resultado de un largo proceso de trabajo, además de significar la finalización de una etapa muy importante de mi vida, por lo que no puedo evitar referirme a algunas de las personas a las que estoy profundamente agradecida, dado que sin ellas ésta no hubiera sido posible.

A mis directoras, Carolina y Laura, por todo lo que me enseñaron y por su apoyo, solidaridad y entrega infinitos. A estas personas excepcionales, con quienes además tengo el placer de compartir el gusto por la filosofía, les debo no sólo la posibilidad de haber podido terminar este trabajo, sino también la de seguir entusiasmada con la idea de dedicar parte de mi vida al estudio de esta disciplina.

A mis compañerxs del grupo de investigación de *Intencionalidad*, de quienes he aprendido un montón de cosas. Pero, especialmente, por sobreponerse a la idea de que en el mundo académico hay una primacía de la competencia, mediante el gesto cotidiano de la solidaridad y el compañerismo.

A lxs docentes de la Escuela de Filosofía que durante el cursado de mi carrera han apostado por entender las aulas como un espacio de discusión genuina y horizontal. Muy especialmente a Marisa Velasco (aunque le molesten las “x”) y Daniel Kalpokas.

A mis compañerxs que participan o participaron de la Asamblea de la Escuela de Filosofía, del “Jardín de Epicuro” u otros espacios de construcción colectiva, con quienes nos hemos animado a preguntarnos qué queremos hacer como estudiantes y futuros profesionales de esta disciplina. En un contexto en el que los estudios universitarios suelen verse como un proceso sumamente individual, agradezco profundamente la posibilidad de haber podido transitar por estos espacios en los que se dejan de lado los intereses personales, y se busca apostar por formas de construcción política más justas y democráticas.

A Valentín, Darío y Franco, por convertirme en su interlocutora, gesto básico que necesitamos para confiar en que tenemos algo para decir. Y a Pedro; si ir a cursar no hubiera sido, todos los días, una cita con nuestra amistad, hubiera sido una tarea que debería haber hecho por obligación y no con tanta alegría.

A mis familias ensambladas, extendidas y postizas; poco convencionales, pero llenas de amor. En especial a mi papá, que me hizo ver desde chiquita que el mundo era un lugar atiborrado de cosas increíbles que descubrir y preguntarse, y a mi mamá, que dice que no entiende absolutamente nada de lo que estudio, y que es la persona más sabia que conozco.

Al Lea, la Debi, las Fer, la Caro, el Mu, la Aye, la Ayle, la Manu, la Juli, la Anto, la Cami, la Clarita, a quienes les debo la existencia misma. En este mundo hecho de cosas como átomos, mesas, termostatos, estados intencionales, movimientos sociales, desórdenes psicológicos, buenos vinos y pasos de danza, mis amigxs son, objetivamente hablando, lo mejor que hay.

Contenido

Introducción.....	3
1.Naturalismos, Intencionalidad y Conciencia	11
El naturalismo de Dennett: Algunas nociones fundamentales	11
La mente naturalizada: La actitud Intencional.....	17
La mente como producto de la evolución: Intencionalidad derivada	33
Teoría de la conciencia: El Modelo de las Versiones Múltiples	39
2.El problema filosófico del “yo” y la propuesta narrativista de Dennett	47
El problema filosófico del “yo”	47
El “yo” según Dennett	55
3.Centros de gravedad narrativos: ¿De qué están hechos?	62
Clases de cosas.....	63
Los objetos intencionales como <i>Abstracta</i>	65
Los centros de gravedad narrativos como <i>Abstracta</i>	68
Clases de psicología intencional.....	75
4.Cómo se constituye el “yo”	78
Explicaciones evolucionistas: Algunos reparos	80
La evolución de la conciencia y el “yo”	85
Yoes múltiples.....	114
5.El acto de narrar: Instintivo y deliberado	117
Exhudando narraciones.....	118
El “yo” como artefacto cultural	126
Conclusiones	145
Bibliografía	152

Introducción

Entonces, ¿qué significa todo esto? La historia de Occidente teorizando sobre el “yo” y la identidad personal es, aunque no exclusiva, sí principalmente, la historia del intento de la humanidad por elevarse a sí misma por encima del mundo natural, y es la historia de cómo ese intento falló. Es otra ilustración, si otra fuera necesaria, de cómo es la soberbia, antes de la caída (Martin y Barressi, 2006, p. 305, mi traducción).

Creo que existe un vago supuesto en los ámbitos locales de discusión filosófica, según el cual prestar atención a algunos aspectos fascinantes de la mente humana, como la capacidad conmovernos, ser creativos, tener experiencias religiosas o sentimientos morales, entre muchas otras, entra en contradicción con un abordaje de la misma en el que se la entienda como parte del mundo natural. Así, concebir lo humano como resultado de un proceso evolutivo parece ir en contra de resguardar aquellas características que nos hacen especiales y que tanto valoramos. En esa línea, cuando se presentan visiones evolucionistas para explicar los rasgos propios de nuestra especie, las mismas son criticadas porque se señala que podrían tener consecuencias éticas o políticas indeseables, y se las rotula, por defecto, de ‘reduccionistas’. Si bien es cierto que hay muchas que efectivamente lo son, y que no debemos dejar de lado una consideración sobre las consecuencias éticas o políticas que tengan las teorías que defendemos, el problema es que la teoría de la evolución de las especies por selección natural es, sencillamente, cierta. Se trata de la única alternativa secular y verdaderamente explicativa con la que contamos para responder a la pregunta de cómo puede haber tanta diversidad entre los seres biológicos, y estar los mismos tan bien adaptados para resolver problemas de supervivencia. En otros términos, es la única que nos permite explicar porqué el mundo natural parece diseñado por una mente inteligente. La pregunta acerca de la causa de la diversidad y el ajuste del mundo biológico ha sido utilizada para defender la existencia de Dios. Esta estrategia ha recibido el nombre de *argumento del diseño*. Una de las más famosas formulaciones de dicho argumento la podemos encontrar en los *Diálogos sobre la religión natural* de Hume (1979). En este texto, si bien los argumentos de Philo, el escéptico, contra Cleantes, el creyente, muestran las innumerables fisuras del argumento del diseño para justificar la existencia de Dios, al final del diálogo Philo termina cediendo y concediéndole a Cleantes que su postura parece la más

adecuada. Dennett (1995), quien afirma que Hume expresa su postura a través de Philo, cree que este final inesperado se debe a que, al no contar con la teoría de la evolución por selección natural, el filósofo moderno *no podía imaginar otra razón* para explicar el presunto diseño de la naturaleza. Así, el aporte darwiniano viene a constituir la única buena explicación disponible, alternativa a la creación divina, de la diversidad y el ajuste presentes en la naturaleza. Precisamente el tipo de explicación con la que Hume, el naturalista, aún no contaba.

Ahora bien, a mi entender, adoptar una perspectiva darwiniana no tiene porqué tener las implicancias negativas que se temían. Al dejar de lado este rechazo *a priori* y considerar las teorías y discusiones evolucionistas en profundidad, encontramos, por un lado, que una perspectiva evolutiva no implica que tengamos que sostener que todos los fenómenos creados por dicho proceso deban ser explicados apelando a su función adaptativa. Por otro, que de aceptar dicha perspectiva tampoco se sigue que las discusiones y dilemas éticos, políticos, antropológicos, puedan ser resueltos apelando a distinciones y conceptos de la biología; en otros términos, que de la reducción ontológica no se sigue la reducción explicativa, algo que se ejemplifica con claridad en la Teoría de los Sistemas Intencionales de Dennett (1987).

Pensadores como Descartes se han sorprendido de nuestras habilidades cognitivas para ordenar nuestra experiencia del mundo de forma racional, mediante categorías conceptuales complejas y útiles. Otros filósofos, como Kierkegaard y Bataille, han hecho más hincapié en aquellos aspectos en los que se presenta una especie de desajuste entre nuestros “apetitos” y nuestras capacidades de intervención, brindando más bien una sensación de que *no estamos hechos* para tolerar la inmanencia a la que estamos condenados. Creo que un abordaje evolutivo de ambas intuiciones (la que señala nuestra adecuación al mundo y la que se fija, más bien, en el desajuste), no las dejaría intactas, pero sería posible y enriquecedor. Con respecto a la primera de ellas, la evolución permite dar cuenta de que a partir de un proceso de sofisticación creciente hayan surgido ciertas capacidades cognitivas extremadamente complejas. Con respecto a la segunda, la perspectiva evolutiva es perfectamente compatible con las visiones que enfatizan que algunas capacidades o características de los organismos no se ajustan a la perfección al tipo de vida que el mundo natural les ofrece (por ejemplo, el anhelo de trascendencia humano con el

hecho de que la vida sea de carácter finito). Creo que es posible compatibilizar una perspectiva evolutiva con este tipo de visiones, dado que el proceso que da origen a la diversidad (la selección natural) es un proceso ciego y sin objetivos, según el cual los rasgos que pasan a formar parte del acervo hereditario de una especie pueden perpetuarse por distintos motivos: por ser adaptativos, por haberlo sido antes o, incluso, por azar (como explicaremos en el capítulo 4).

El párrafo citado en el epígrafe es una de las conclusiones de la *historia intelectual* del “yo” desarrollada por Martin y Barressi y da cuenta de cómo en el seno de esta noción parece residir necesariamente una idea de excepcionalidad humana con respecto al resto de los seres vivos. Creo que el concepto del ‘yo’, emparentado con la idea de ‘alma’, puede ser pensado como uno de los últimos lugares “de refugio donde podía esconderse Dios o los dioses” (usando una expresión de Dupré, 2003), en la medida en que parece que las cosas que constituyen nuestra mismidad (nuestros recuerdos, nuestro temperamento, la posibilidad de tomar decisiones, entre muchas otras cosas) están firmemente asentadas, constituyendo nuestra ‘esencia’. La idea sustancialista de alma permitía explicar esto de forma mucho más satisfactoria, mientras que renunciar a la misma y entender a la especie humana en continuidad con otras especies y a sus rasgos como productos de un proceso evolutivo, exige una explicación mucho más compleja.

Durante el desarrollo de esta pequeña investigación, pude darme cuenta de que yo también suponía que entender las características humanas como resultado de un proceso evolutivo implica pensar que, si bien parte del mismo comparte rasgos y características con la evolución del resto de las especies, en el caso de la humana hay un momento en el que se debe decir “y en este punto, se produce un milagro”; un momento en el cual tiene lugar una suerte de cambio mágico, que nos dota de características esencialmente diferentes que son, además, inexplicables. De este modo, se acepta que compartimos ciertas similitudes con los primates, pero queda resguardado ese “no-sé-qué” que nos hace tan especiales. Con mis primeras lecturas de Dennett pude advertir este supuesto erróneo, y durante buena parte del desarrollo de este trabajo pensé que el mismo podría terminar pretendiendo contribuir a atacar ese espejismo idealizado de lo humano (lo que, por una parte, me complacía, mientras que, por otra, muchas veces me generaba inquietud e incomodidad). Sin embargo, fue a través de Dennett mismo que encontré una propuesta teórica que nos

permite vernos de un modo que impide que dicho espejismo se constituya como tal - dado que se constata que no somos poseedores de una excepcionalidad ontológica (Schaeffer, 2007)-, pero a su vez se ofrece una explicación materialista en la cual la imagen que tenemos de nosotros mismos está enriquecida en un grado tal que no tiene nada que envidiarle a la vieja idea de alma (como intentaré mostrar en el capítulo 5).

Dennett es uno de los filósofos contemporáneos que más seriamente se ha tomado el proyecto de explicar la mente humana desde una perspectiva naturalista. Pero, además, quizás como ningún otro filósofo naturalista, ha pretendido mostrar específicamente cómo el marco teórico proporcionado por la teoría darwiniana puede ayudarnos a comprender no sólo cuáles son los mejores modos de entender las capacidades que nos singularizan como especie, aún siendo criaturas "diseñadas" por la evolución natural, sino también, en ese mismo marco, cómo es posible preservar, bajo cierta interpretación, el rol y el contenido de nuestros conceptos mentales ordinarios.

Ahora bien, centrándonos en el problema principal que aquí nos ocupa, el tratamiento que realiza Dennett del problema de la identidad personal, encontramos que este filósofo comparte la idea de Hume (1739) de que un estudio biológico del hombre no encontrará un correlato físico exacto del "yo". De allí que afirme que el "teatro cartesiano" (es decir, un centro especial en el cerebro en donde se presenta el verdadero flujo de la conciencia, Dennett, 1991a) contemplado por un "yo", no existe. El "yo" es, antes bien, una abstracción definida por un conjunto de atribuciones e interpretaciones, propias y ajenas, que conforman la biografía del cuerpo viviente del que es el "centro de gravedad narrativa" (1988, 1991a, 1992). Se trata, pues, de una abstracción que, además, juega un rol fundamental en la economía cognitiva de dicho cuerpo, dado que el modelo mental que cada uno tiene sobre sí mismo ocupa un lugar privilegiado en relación al resto de las cosas sobre las que construye otros modelos mentales.

Para este autor, la constitución del "yo" es fruto de un proceso que se da en el desarrollo psicológico individual, pero dentro de un marco social. El contar historias y los juegos de simulación serían ingredientes claves en dicho proceso. Estas historias nos definen alrededor de un punto luminoso al que llamamos "yo". En efecto, hay una gran distancia entre nuestro sentido de la identidad y el que tiene una

ameba (el hecho de que la ameba tiene un sentido biológico de identidad se puede constatar porque sigue principios muy básicos como el de no comerse a sí misma); pero dicha diferencia es explicada de un modo naturalista, a partir del comportamiento biológico particular de nuestra especie y a partir de las prácticas, sobre todo lingüísticas, que llevamos a cabo. Es aquí donde el lenguaje juega un rol crucial como herramienta para representarnos a nosotros mismos: en nuestra relación habitual con el entorno nos valemos de narraciones auto-protectoras que vamos tejiendo con otros a lo largo de nuestro desarrollo y, por eso, nuestro sentido del “yo” tiene un carácter eminentemente social.

En este trabajo he pretendido evaluar la potencialidad de la concepción narrativista del “yo” de Dennett a la hora de brindar una explicación de las personas y de la mente humana que no sólo sea compatible con un enfoque naturalista darwiniano, sino que también tome en cuenta los rasgos que la psicología popular reconoce como distintivamente humanos.

Ahora bien, la teoría dennettiana de la identidad personal ha sido objeto de críticas de distintos frentes, que la han presentado como absurda, incomprendible, o contradictoria (Carr, 1998, Di Francesco *et al.*, 2014, Menary, 2008, Schechtman, 1996, Strawson 2005, 2009, entre otros). Creo que al ponerla en relación con otras partes de la teoría de Dennett, la misma puede comprenderse mejor. En particular, considero que es necesario vincular la concepción narrativista del “yo” de Dennett con su teoría de los sistemas intencionales, con su explicación de la conciencia, con el estatus que el autor le confiere a determinado tipo de abstracciones, con su concepción evolucionista sobre el origen del “yo” y con el rol que otorga al lenguaje en la inteligencia. Y eso es lo que intento hacer en este trabajo.

En el **capítulo 1** presento tres tópicos fundamentales de la filosofía de Dennett que, desde mi punto de vista, hacen que la misma constituya un marco particularmente desafiante y, a su vez, atractivo, para plantear el problema del “yo”: su perspectiva naturalista, su teoría de la intencionalidad y su teoría de la conciencia. En primer lugar, planteo los sentidos en los que la perspectiva dennettiana puede ser entendida como naturalista. Además, señalo que ésta se caracteriza por otorgar un lugar privilegiado a la teoría de la evolución. En segundo y tercer lugar, expongo su teoría de los sistemas intencionales y su perspectiva de la intencionalidad derivada. Intento mostrar que las mismas presentan a la mente humana en marcada

continuidad con el resto de las mentes a las que ha dado lugar el proceso de evolución por selección natural, afirmando con ello que no existe una diferencia ontológica, un “hecho profundo último”, que permita hablar de una intencionalidad intrínseca que distinga nuestros contenidos mentales de los que podemos atribuir a otros tipos de mentes o agentes. Por último, realizo una breve exposición de su teoría de la conciencia, en la que Dennett intenta rebatir la idea del “teatro cartesiano” y propone un modelo alternativo, que bautiza con el nombre de “modelo de las versiones múltiples”, según el cual no hay un “centro de mando” en el cerebro. Con esto pretendo señalar que las teorías de Dennett sobre los contenidos mentales y sobre la conciencia evidencian con claridad que estamos ante una filosofía sumamente consecuente con su perspectiva naturalista. Es por esto mismo que dicha propuesta filosófica constituye un marco particularmente desafiante para plantear el problema de la identidad personal, dado que implica un cuestionamiento de las concepciones clásicas en, al menos, dos sentidos fundamentales. Con respecto a los contenidos mentales, implicará dejar de lado la idea de que existen “hechos profundos últimos” que hagan que los humanos tengan creencias y deseos en un sentido completamente diferente a las creencias y deseos que podemos atribuir a otros animales o sistemas. Con respecto a la conciencia, esta perspectiva desenmascara y descarta la visión clásica según la cual hay una especie de proyección interna del flujo de la conciencia en nuestras cabezas.

En el **capítulo 2** dedico un primer apartado a plantear el problema del “yo”, intentando recuperar algunas de las preguntas filosóficas más relevantes y las preocupaciones a las que las mismas se ven vinculadas. Si bien no se trata de un rastreo histórico preciso, sí se intenta resaltar el hecho de que la pregunta ha sufrido modificaciones que acompañaron los cambios de concepciones ontológicas y antropológicas. También, se esboza, a grandes rasgos, cuáles son las principales corrientes actuales para explicar la identidad personal y sus principales inconvenientes. En una segunda instancia, se presenta de manera preliminar la teoría de Dennett y se intenta esclarecer a qué se refiere al hablar de “narraciones”.

En el **capítulo 3** intento responder a la pregunta sobre qué es el “yo” (si es algo), en tanto caracterizado como “un centro de gravedad narrativo” (Dennett, 1988, 1991a, 1992). Para elaborar mi respuesta a este interrogante, apelo a una distinción, que Dennett emplea en otros textos y para otros fines (Dennett, 1987, 1991b) y que

fuera elaborada originariamente por Reichenbach (1938), entre tres clases de entidades: *Concreta*, *Illata* y *Abstracta*. Defiendo la idea de que el mejor modo de realizar una reconstrucción interpretativa de la noción dennettiana de “yo” es entendiendo los centros de gravedad narrativos como *Abstracta* en el sentido de Reichenbach. Para mostrar que ésta es una interpretación plausible del modo en que Dennett entiende al “yo” me apoyo en que este filósofo conoce y emplea la distinción de Reichenbach en otras partes de su obra y en que traza explícitamente la analogía del “yo” con los centros de gravedad que son, justamente, uno de sus ejemplos paradigmáticos de *Abstracta*. Además, según intentaré mostrar, este modo de entender al “yo” permite dar sentido y comprender mejor ciertos aspectos de la teoría narrativista dennettiana.

Una de las particularidades de la teoría de la identidad personal de Dennett es que, en ella, el “yo” es el resultado de un proceso evolutivo. Así, hay, para Dennett, una historia evolutiva del origen del “yo”, que reconstruyo en el **capítulo 4**. En dicha exposición, el origen del “yo” se verá estrechamente vinculado con el de aquello que lo hace posible: la conciencia. El hecho de que ambos fenómenos sean abordados desde una perspectiva evolucionista, conduce a una discusión epistemológica en torno a las explicaciones evolutivas. En primer lugar, examino las críticas que han realizado Gould y Lewontin (1979, 1983) contra el ‘pan-adaptacionismo’ y las consideraciones de Gould y Vbra (1982) sobre las ‘exaptaciones’. Respecto de ambos aportes, se ha considerado que constituyen un problema para las explicaciones de Dennett, bajo la interpretación de que él no reconoce el papel de las ‘exaptaciones’ y de que su teoría constituye una versión de ‘pan-adaptacionismo’. Luego, expongo la explicación dennettiana sobre el origen de la conciencia y el “yo”, intentando mostrar que dichos señalamientos no constituyen un problema para sus explicaciones. Finalmente, señalo que las explicaciones evolucionistas permiten dejar de concebir a los fenómenos que intentan explicar (especies, rasgos morfológicos, comportamientos) como tipos de entidades para las que valen distinciones del tipo “todo o nada”. El fenómeno del “yo” entendido bajo esta perspectiva adquiere un carácter procesual y prescinde de la necesidad de poseer límites claros y bien definidos. Así, se vuelven comprensibles, por ejemplo, algunos trastornos de personalidad, a los que Dennett dedica especial atención para pensar el fenómeno del “yo” y que son abordados en la última sección de este capítulo.

Clark (2002) ha sostenido que la de Dennett es una perspectiva dual, en la medida en que presenta a las mentes humanas como profundamente diferentes de, así como en continuidad con, la de las otras especies. En el **capítulo 5** defiende que el aspecto continuista de esta perspectiva dual también se manifiesta de un modo particular en su teoría del “yo”, dado que Dennett propone una explicación del modo en el que producimos las narraciones que nos constituyen incluyendo componentes de carácter estrictamente biológico. En *Why Everyone is a Novelist* (Dennett, 1988) y *The Origin of Selves* (Dennett, 1989) se compara el tejer narraciones con actividades que realizan otros animales, como las construcciones de represas por parte de los castores, con lo que parece que la primera sería una actividad completamente instintiva y, por ende, se negaría la posibilidad de tener una intervención consciente en su desarrollo. Ahora bien, esta sería una consecuencia que volvería a la teoría de Dennett sumamente anti-intuitiva. Sin embargo, en la segunda parte del capítulo sostengo que una reconstrucción de la teoría de Dennett sobre el rol del lenguaje en la inteligencia, permite negar esta presunta implicancia. Si se toma en cuenta, pues, la posición de Dennett con respecto a la relación entre lenguaje y pensamiento, se puede comprender de qué modo este autor también puede dar cuenta de aquellos aspectos del “yo” que son específicamente humanos.

Si bien las consideraciones que ofrezco a lo largo de este trabajo están lejos de ser exhaustivas y concluyentes, espero que contribuyan a mostrar algunas virtudes de un abordaje evolutivo y narrativista de la constitución del “yo”, que integra aspectos estrictamente biológicos con otros estrictamente culturales, y de este modo pueda darnos herramientas para defender que una buena explicación naturalista de nuestro “yo” ha de apoyarse en una concepción híbrida (biológica y no-biológica) de los seres humanos.

1.Naturalismos, Intencionalidad y Conciencia

El naturalismo de Dennett: Algunas nociones fundamentales

Un gran número de filósofos contemporáneos estaría de acuerdo en que hemos aprendido de los errores del pasado que la filosofía no accede a resultados innovadores si realiza su tarea desde un enfoque axiomático; antes bien, para alcanzar perspectivas que modifiquen productivamente nuestra forma de vincularnos con el conocimiento del mundo en el que vivimos, hay que poner en juego intuiciones, descubrimientos y teorías empíricas, argumentos rigurosos y experimentos mentales. Ese es uno de los (al menos) dos sentidos fundamentales en los que podemos decir que la filosofía de la mente de Dennett se enmarca en una perspectiva naturalista: el que concibe a esta última como una disciplina no apriorística que trabaja de manera cooperativa con las ciencias particulares. Se debe a Quine (1953, 1965) -quien fue su profesor en Harvard- la más importante formulación de esta idea, de la que Dennett, a pesar de haber sido un alumno muy crítico de su profesor, constituye un ejemplo notable: “Mi sentido de que la filosofía está aliada a, y es sin duda una continuación de las ciencias físicas sustenta tanto mi humildad acerca del método filosófico como mi optimismo acerca de su progreso” (Dennett, 1987, p. 18). El espíritu de cooperación que Dennett pregona entre las disciplinas científicas y la filosofía, se puede ver no sólo en la importancia que los resultados experimentales de aquellas adquieren en la formulación y reformulación de sus teorías, sino también en que la comunidad científica es uno de los principales destinatarios de sus escritos.

Lo dicho con anterioridad no implica que Dennett no realice consideraciones importantes sobre la especificidad de la filosofía. Desde su punto de vista, ésta disciplina a menudo no produce resultados firmes y dignos de confianza como lo hace la ciencia en sus mejores momentos, pero puede producir nuevas maneras de ver las cosas, de pensar en ellas, de formular las preguntas y de ver qué es lo importante y por qué.¹

¹ Hace unos meses tuvo lugar en Buenos Aires la reunión anual de la Association for the Scientific Study of Consciousness (Asociación por el Estudio de Ciencias de la Conciencia),

Entre los principales recursos específicos de la filosofía que serán mencionados a lo largo de este trabajo, se encuentran los experimentos mentales, que Dennett también llama “bombas de intuición”. Estos consisten en el diseño de situaciones imaginarias que nos permiten aislar aspectos críticos de las cuestiones filosóficas que están en discusión, ofreciéndonos nuevas formas de escrutarlas conceptualmente. Además de elaborar muchos experimentos mentales propios, Dennett recupera experimentos mentales anteriores, ya sea para aprovecharlos e incluso reformularlos, ya sea para discutirlos. A partir de sus discusiones de muchos experimentos famosos, como el de la habitación china (Searle, 1980) o el de la científica Mary que es ciega a los colores (Jackson, 1986), traza una distinción entre dos tipos de experimentos mentales, uno más útil que el otro. El defecto de los experimentos mentales que son criticados por Dennett (como los citados en la oración anterior) radica en que sólo funcionan si se siguen al pie de la letra todas las instrucciones y se frena la imaginación en el punto en el que se enuncia que hay que frenarla, lo que coincide arbitrariamente con la postura que quien los formula quiere sostener y que, más que arrojar luz sobre la problemática a tratar, sólo desvían la atención (Cfr. Dennett, 1991a, p. 451). A diferencia de estos, los “buenos” experimentos mentales son “bombas de intuición” (de las que brindaremos algunos ejemplos en los próximos capítulos de este trabajo) que ayudan a concebir nuevas posibilidades y permiten que imaginemos algo que antes no veíamos posible; posibilidades que después deberán ser confirmadas por métodos más sistemáticos.

Si bien el empleo de experimentos mentales es un recurso específicamente filosófico, podemos entender que el abordaje que Dennett realiza de los mismos es, a su vez, naturalista por las modalidades que propone para elaborarlos e interpretarlos. Dichas modalidades, dan cuenta nuevamente de una concepción de la filosofía como una disciplina no axiomática que se puede ver en acción, tanto en el desarrollo de bombas de intuición, como de argumentos en general, en los criterios que utiliza Dennett para inclinarse a favor de una u otra postura en los casos en los que hay distintas propuestas en pugna. En lugar de considerar prioritariamente la plausibilidad *a priori* de las mismas, opta por evaluar su plausibilidad en relación a otras cosas que efectivamente sabemos acerca del mundo en el que vivimos. De ahí

de la que Dennett forma parte, por lo que me encontré con la oportunidad de poder hacerle algunas preguntas. En un comentario respecto de la práctica filosófica, me dijo que cuando uno tiene preguntas que resolver debe acudir a la ciencia, pero debe acudir a la filosofía cuando se está preguntando cuáles son las preguntas que hay que formular.

que adquieran fuerza ideas como la de que a veces una imposibilidad de hecho es teóricamente más interesante que una posibilidad en principio (Dennett, 1991a, p. 16), o de que cierta explicación, estrictamente consistente en un sentido lógico, pueda ser descartada en la práctica por carecer por completo de plausibilidad biológica (Dennett, 1991a, p. 227).

Otro motivo, consonante con todo lo anterior, por el cual Dennett merece la etiqueta de naturalista, es que su filosofía asume como punto de partida, según sus propias palabras, la existencia del mundo objetivo, material, tal como este es concebido desde el enfoque de tercera persona de las ciencias físicas (Dennett, 1987, p. 18). Según esta formulación, el naturalismo es entendido como una tesis metafísica de acuerdo con la cual sólo hemos de incluir en nuestra ontología entidades naturales, en contraposición a la admisión de hipotéticas entidades sobrenaturales o no-naturales. En palabras de Diana Pérez (2002), sostener que todo lo que hay es natural, en este sentido, es sostener la que “todo lo que hay forma parte de un sistema espacio-temporal causalmente cerrado” (p. 12). La apuesta de Dennett es, en consonancia con ello, estudiar la mente desde la perspectiva materialista, de tercera persona, de la ciencia contemporánea (Dennett, 1987, p. 20). Ahora bien, aunque estos son los términos empleados por el mismo Dennett, el recorrido que siguen sus argumentos mostrará que no favorece un tipo de explicación reductiva de los fenómenos mentales; por lo que dichas afirmaciones deben ser tomadas como una declaración de principios, según la cual se asumirá que todos los fenómenos naturales son estudiables y que esto debe hacerse de forma tal que no se proponga una ontología contradictoria con la de las ciencias naturales.

En cuanto al estudio de la mente humana, el método privilegiado para la psicología (y para la filosofía de la mente bajo una perspectiva naturalista) ya no sería el clásico método consistente en examinar la mente exclusivamente desde la perspectiva de la primera persona, dado que los fenómenos mentales no pueden ser concebidos como objetos internos, peculiares y separados de los rasgos físicos del cuerpo. Dicho de otra forma, los enfoques naturalistas incorporan los avances científicos bajo el supuesto de que la mente constituye un objeto público y estudiable desde la tercera persona (Cfr. Pérez, 2002). Esto puede contraponerse con la perspectiva de Nagel (1974), quien, al entender a lo mental como un fenómeno inefable y subjetivo, no tiene motivos para pensar que un análisis de tercera persona

pueda arrojar luz sobre el asunto; el supuesto de fondo es la idea de que hay fenómenos perfectamente reales que no pueden ser explicados desde una perspectiva objetiva. En consonancia con esta idea, Nagel (1974) sostiene que la evidencia comportamental no tiene relevancia alguna en la explicación de los fundamentos internos de la conciencia. Dennett, por su parte, no sólo piensa que dicha evidencia es muy importante y útil, sino que además tomará en cuenta los resultados empíricos proporcionados por las neurociencias; es decir, la evidencia objetiva o de tercera persona.

En un contexto general en el que se comenzaban a ejercer fuertes críticas al empirismo radical del S. XX, del que formaban parte Quine y Ryle -quien fue el supervisor del doctorado de Dennett-, nuestro filósofo se posiciona sosteniendo que es necesario introducir un profundo cambio de perspectiva, como lo hicieron también Chomsky y Fodor (véase Brook y Ross, 2002). A diferencia de lo que se sostenía desde empirismo conductista, esta nueva perspectiva reivindica la idea de que es necesario estudiar los procesos cognitivos y neurológicos para lograr una comprensión completa de los fenómenos mentales. Sin embargo, aunque Dennett comienza realizando una fuerte crítica a la concepción de creencia sumamente deflacionada que mantienen sus profesores, finalmente termina ofreciendo una visión mucho más cercana a la de éstos, que a la de los filósofos con los que comparte el podio de los impulsores del cambio.

En el marco de la filosofía de la mente y de la antropología filosófica, el naturalismo desafía pre-conceptos y concepciones enraizadas acerca de la mente humana y, consecuentemente, acerca de la singularidad de los seres humanos frente a otras criaturas. Adoptar esta perspectiva conlleva regularmente la aceptación de la teoría de la evolución darwiniana, que concibe a los seres humanos como una especie emparentada evolutivamente con las demás. Desde esta perspectiva, se nos insta a comparar nuestras características específicas como especie con las de otros animales superiores, sin presuponer singularidades que requieran de un marco metafísico y epistemológico anti-naturalista. En este sentido, un rasgo distintivo de la filosofía naturalista de Dennett es el lugar que en ella ocupa la teoría de la evolución de las especies. Desde su punto de vista, la teoría de Darwin permite ligar el admirable mundo que habitamos -al cual describimos, metafóricamente, como un producto que parece haber requerido, para su producción, de creatividad, ingenio y brillantez- con

el mundo del azar y la ausencia absoluta de diseños. La denominación de Dawkins (1986) del proceso de evolución como un “relojero ciego” hace referencia, justamente, a que la misma permitiría captar esas dos dimensiones del mundo, dado que, por una parte, el proceso de evolución se asemeja a un relojero en la medida en que “produce diseños” de altísimo grado de complejidad, al tiempo que, por otra parte, se trata de un proceso mecánico y sin objetivos, es decir, “ciego”. Así, a través de la teoría de evolución por selección natural deberíamos poder explicar la presencia de complejos “diseños” naturales, sin olvidar, en ningún momento, que no hay ninguna mente inteligente detrás (Dennett, 1995).

Para explicar en qué consistiría ser consecuente con una perspectiva darwiniana en la formulación de teorías, Dennett apela a un término folklórico que aparece en el Diccionario de la Lengua Inglesa de Oxford, que podríamos traducir como “gancho celeste” [*skyhook*], que en el campo de la aeronáutica hace referencia a un invento imaginario que permitiría mantener suspendido un objeto en el cielo (Dennett, 1995, cap: “El ácido universal”). Si bien sus usos en este ámbito tienen que ver con chistes sobre modos en que los aviones podrían permanecer mágicamente en el aire, Dennett señala que su origen puede tener que ver con la idea del *deus ex machina* de los antiguos dramaturgos griegos. Como se sabe, los dramaturgos griegos resolvían los problemas de la trama de una obra haciendo intervenir un dios en la escena y salvando de ese modo sobrenatural la situación problemática. Siguiendo lo que sugiere este caso, emplear un “gancho celeste” sería introducir un elemento externo para resolver un problema de la historia, que a su vez, va en contra de la lógica “interna” de la misma. Una especie de “solución” externa pero también “superior”.

Pasando ahora al contexto de los debates sobre las explicaciones evolucionistas de la mente, apelar a una suerte de “gancho celeste” para explicar algún aspecto complejo de alguna de las criaturas vivas, sería renunciar al intento de elaborar una explicación naturalista. Ahora bien, aunque deberíamos rechazar la apelación a los ganchos celestes no existen, sí podemos disponer legítimamente, en cambio, de “grúas” [*cranes*]. Esta es la metáfora opuesta a la de los “ganchos celestes”, que Dennett introduce para referirse a las explicaciones que consiguen dar cuenta de fenómenos de alta complejidad (y que en ese sentido pueden entenderse

como *elevados*) apelando exclusivamente a los mecanismos no inteligentes de los procesos de la evolución natural.

Intentando brindar una formulación más específica, podemos decir que utilizar un “gancho celeste” sería proponer una excepción al principio darwiniano según el cual los “diseños” del mundo natural son, en último término, el resultado de una actividad mecánica no inteligente y sin motivo, y no de una mente o principio superior. En cambio, las grúas serían sub-procesos de diseño que permiten la aceleración de la lenta y básica selección natural, y que, a su vez, deben poder explicarse como resultado de la misma. Las grúas son, en definitiva, formas de aumentar localmente el poder del proceso de selección natural. Un ejemplo de una grúa es la reproducción sexual, que proporciona un tipo de variabilidad genética que hace que las especies que se reproducen sexualmente tengan una velocidad para moverse en el espacio del diseño mucho mayor que el de las especies que se reproducen asexualmente. Así, la existencia de muchos de los diseños que podemos encontrar en la naturaleza sería prácticamente imposible si la selección natural no hubiera tenido que seleccionar entre individuos que eran producto de la variabilidad genética que sólo la reproducción sexual proporciona. (Otro ejemplo de una importante grúa, que recibe el nombre de “Efecto Baldwin,” será explicado en el capítulo 4).

Explotando las metáforas de Dennett, podemos decir que las grúas son más costosas, menos milagrosas, y necesitan estar bien asentadas sobre tierra firme para funcionar; sin embargo, mientras que “los ganchos celestes son elevadores milagrosos, no soportados e insoportables”, las grúas “no son menos excelentes que los elevadores y tienen la ventaja de ser reales” (Dennett, 1995, p.113). Si bien las grúas encuentran un límite muy rápido si actúan por sí solas, se las puede montar unas sobre otras para potenciar sus capacidades. Aunque la utilización de grúas montadas unas sobre otras es extraña en una construcción, en principio no hay límites para la cantidad de grúas que se pueden emplear unas sobre otras para alcanzar un objetivo (en el capítulo 5 brindaremos un ejemplo de esta estrategia de montar grúas explicativas unas sobre otras al exponer la propuesta de Dennett sobre la “torre de generación y testeo”).

La posibilidad de pensar que una enorme cantidad de grúas montadas unas sobre otras da lugar a la increíble biodiversidad que conocemos, nos permite

sostener que la idea de Darwin es suficiente para dar cuenta de todo el proceso. Esto no significa que cada fenómeno será explicado a partir de su presunta función adaptativa; pero sí que no podemos pensar que, aunque las especies se fueron desarrollando en función de un proceso evolutivo, en algún punto ocurrió un “salto ontológico”, por medio de alguna suerte de milagro, que ubicó a nuestra especie en un lugar especial y diferente con respecto a todas las demás. Concretamente, el contrincante con el que está discutiendo Dennett, en este punto, no es alguien que niegue que el hombre comparte un ancestro común con el mono, sino aquel que piensa que en algún momento del proceso se produce una importante excepción que implica que hay determinados fenómenos que no van a poder explicarse en términos evolutivos. Así, la resistencia a incorporar la dimensión evolutiva de los fenómenos del mundo natural en las explicaciones filosóficas (bajo el supuesto de que se trata de fenómenos “de otro orden”), es una forma de disfrazar la idea del soplo de Dios que da origen a nuestra alma inefable, una idea a la que, al parecer, nos cuesta mucho renunciar.

La mente naturalizada: La actitud Intencional

Como dijimos, una de las marcas distintivas del naturalismo dennettiano es la centralidad que otorga a la teoría de la evolución, según la cual las mentes humanas deben ser entendidas en continuidad con las otras especies. La teoría que permite dar cuenta de dicha continuidad es la Teoría de los Sistemas Intencionales², cuya formulación más completa se encuentra en *La Actitud Intencional* (Dennett, 1987). En ella se brinda una explicación del contenido mental partiendo del supuesto de que la mente es un objeto público y estudiable desde la perspectiva de la tercera persona. La teoría de los Sistemas Intencionales de Dennett guarda un fuerte vínculo con cierta parte del *corpus* de creencias y habilidades que constituyen parte de nuestro saber de sentido común: la *Psicología Folk*. Con esto no queremos decir que estemos frente a algo así como un ‘filósofo del sentido común’; más bien, Dennett considera a la psicología de sentido común como un fenómeno que debe ser tenido en cuenta en cualquier reflexión completa sobre lo mental, sin que eso implique entender que una

² La noción filosófica de *Intencionalidad* hace referencia a la propiedad de ciertos fenómenos de ser *acerca de*, *estar dirigidos a* o representar, otras cosas. En el siguiente apartado haremos una referencia al origen del término.

teoría filosófica debe procurar mantener intactas las intuiciones que de él se desprenden.

Según Dennett, lo que el sentido común dicta no sólo deja sin resolver muchos problemas importantes, sino que, además, muchas veces éste se entrega a persuasivas intuiciones contradictorias entre sí. No hay ningún aspecto del sentido común que no podamos poner en duda: las cosas que parecen evidentes para alguien pueden no serlo para otra persona; además, hay cosas que pueden resultar incuestionables a todos y no sean más que ilusiones cognitivas de nuestra especie, y otras, que simplemente, podrían ser versiones popularizadas de teorías científicas superadas. Las intuiciones de sentido común no son, entonces, consideradas un punto de partida certero, o una buena piedra de toque de nuestras teorías, por lo que no sería problemático arribar a conclusiones que se le opongan. Así, habrá muchas creencias de este *corpus* que serán ciertas y muchas otras que serán falsas, pero - como como es el caso de los ejemplos de la oración anterior- incluso para estas últimas sería posible, en principio, encontrar razones que expliquen su existencia.

Entre nuestras habilidades cognitivas de sentido común más sorprendentes se encuentran aquellas que utilizamos para comprender el comportamiento de las personas. Estamos familiarizados con ellas desde la infancia, las utilizamos casi sin notarlo y con un gran éxito predictivo. Se las ha reunido bajo el nombre de “Psicología *Folk*”, y a los conceptos que forman parte de ella se los ha llamado conceptos “mentalistas”, esto es, conceptos que hacen referencia a estados mentales intencionales como los deseos, las creencias, el dolor, el conocimiento, el temor, la fe, etc. (Cfr. Dennett, 1987, p. 20). En algún punto, la psicología *folk* es análoga a lo que podemos denominar la física *folk*, o física popular, que nos permite predecir cómo funcionan los objetos medianos con respecto a, por ejemplo, los estados de agregación de la materia, sin conocer necesariamente sus leyes científicas, ni realizar cálculos. De ambas perspectivas podemos decir que hay una gran parte que parece ser aprendida, mientras que hay aspectos, en una y en otra, que parecieran ser innatos, al menos para algunos teóricos³. Independientemente de si podemos determinar esto o no, desde la perspectiva dennettiana se sostiene que el hecho de que uno de estos enfoques sea innato o irresistible no implica que sea verdadero. Para adoptar esta postura, según la cual algunas de nuestras creencias innatas

³ Para estudios que defienden que una parte importante de nuestra psicología *folk* es innata, véase Humphrey (1986) y Leslie (1992).

podrían ser falsas, es preciso demostrar tales errores, pero también explicar a qué se podría deber nuestra sensación de certeza. Con la física popular esto resulta evidente cuando observamos algunos de los contra-intuitivos resultados de la física de partículas. En el caso de la psicología popular los ejemplos son más escasos -en alguna medida por el menor desarrollo que ha tenido la psicología como ciencia-; pero sí podemos señalar límites que la misma tiene para tratar con algunos casos “anormales” (el fenómeno del cerebro partido, la negación de la ceguera, el hemi-descuido, los trastornos de personalidad múltiple, entre muchos otros.)

La fe que tenemos en nuestra física y psicología *folk*, que forman parte de lo que Sellars (1963) denominó la *imagen manifiesta del mundo*, no se debe simplemente a una obstinación para mantener las creencias a las que nos hemos acostumbrado, sino a que -dentro de ciertos límites- estas teorías funcionan:

No es accidental que tengamos la imagen manifiesta que tenemos; nuestros sistemas nerviosos fueron diseñados como para hacer las diferenciaciones que necesitamos de forma rápida y confiable, como para colocar bajo rótulos sensitivos únicos, las características comunes pertinentes de nuestro entorno, e ignorar todo aquello de lo que habitualmente podemos desentendernos.
(Dennett, 1987, p. 23)

Desde una perspectiva claramente darwiniana, Dennett sostiene que las creencias y hábitos de cada especie para vincularse cognitivamente con el mundo –en general- pueden ser caracterizados de modo más preciso como “útiles” que como “ciertos”, ya que éste es el aspecto que premia la Selección Natural, pudiendo o no coincidir con la verdad de sus contenidos⁴. Al mismo tiempo, para Dennett la Selección Natural no es omnipotente, por lo que también cabría la posibilidad de encontrar aspectos no utilitarios que de todos modos hayan sido seleccionados; cuestión que será abordada en el capítulo 4. Estas últimas aclaraciones no le quitan potencia a la idea de Dennett de que no nos sería posible enfrentar con éxito la mayoría de las dificultades cotidianas, desde no congelarnos, hasta realizar tareas en equipo, si no contáramos con sistemas eficientes de generación de expectativas y predictibilidad como los que que nos brinda nuestra imagen manifiesta.

⁴ Por ejemplo, podría haber ilusiones cognitivas de una especie que fueran producto de un atajo en el diseño de su sistema cognitivo.

El interés científico/filosófico por la psicología *folk* puede estar guiado por dos motivaciones diferentes. Una de ellas es su existencia como fenómeno, la que -al igual que, por ejemplo, la existencia de las religiones- resulta innegable, con independencia de su verdad. La otra, en cambio, centra su atención en que esta teoría parece verdadera, por lo que se evalúa si es una buena candidata a incorporarse al *corpus* de conocimientos científicos. Sin embargo, puede que tal estudio nos lleve a concluir que se trata de una mala teoría, que no cabe incorporar en su totalidad como conocimiento sobre cómo efectivamente funcionan nuestras mentes. Esta es la tesis que defiende, por ejemplo, Churchland (1986). En contrapartida, Fodor (1985, 1987) es un ejemplo de alguien que piensa que la psicología *folk* es una buena teoría, que debe mantener una relación de continuidad con la psicología científica. Las investigaciones filosóficas de Dennett lo llevan a afirmar que, mientras algunos aspectos de nuestra psicología *folk* pueden ser dejados de lado, la mayor parte de este *corpus* de creencias da cuenta de una habilidad extremadamente útil que tenemos los seres humanos para comprendernos y relacionarnos unos con otros: la actitud intencional.

La actitud intencional es una estrategia que consiste en tratar al sistema cuyo comportamiento se quiere predecir como un agente racional con creencias, deseos y otros estados mentales dotados de “intencionalidad”. Para entender en qué consiste adoptar la actitud intencional podemos comparar dicha estrategia con otras dos a las que la misma está asociada: la actitud física y la del diseño. Todas ellas son estrategias de razonamiento que nos permiten hacer predicciones, en las que el éxito que tengamos en su implementación depende tanto de las estrategias mismas, como de la aplicación efectuada en cada caso.

La *estrategia física* es la que utilizamos cuando nos movemos del lugar en el que estamos para evitar mojarnos porque hemos volcado un vaso de agua cerca de nuestro cuerpo; o cuando intentamos atrapar un objeto sólido que nos arrojan desde un lugar cercano, a diferencia de lo que haríamos si se tratara de una sustancia líquida o gaseosa. Para predecir el comportamiento de un sistema desde esa estrategia es necesario determinar su constitución física, la naturaleza física de los impactos que sufre y utilizar los conocimientos de las leyes de la física para predecir el resultado que se producirá, para cualquier entrada de datos. Si bien no es siempre

viable en la práctica –su utilización es, por lo general más modesta y local-, un principio de las ciencias físicas es que dicha estrategia debería funcionar siempre.

Ahora bien, con frecuencia es conveniente negociar algo de infalibilidad a cambio de practicidad. Siguiendo ese criterio “económico”, la *actitud del diseño* en muchos casos resulta más eficiente, dado que no exige un conocimiento de la constitución física de un sistema, sino que apoya su predicción en la suposición de cómo está diseñado para comportarse en distintas circunstancias. Esta estrategia predictiva se puede aplicar no sólo a artefactos, sino también a muchos seres vivos, dado que, sin caer en una lectura burda y teleológica de la evolución, Dennett recurre a ella constantemente para pensar nuestra realidad como seres que son producto de la evolución y, por ende, *diseñados* por la *Madre Naturaleza* (Dennett, 1987, cap: “Los sistemas intencionales en la etología cognitiva: Defensa del *Paradigma panglossiano*”).

Siguiendo un criterio económico similar al que nos condujo de la estrategia física a la del diseño, podemos suplantarlo a esta última por la *estrategia intencional*. Esta consiste en tratar al sistema cuyo comportamiento hay que predecir como a un agente racional, y deducir qué creencias debería tener ese agente, dado su posición en el mundo y su objetivo. Bajo las mismas consideraciones, se deducen los deseos del agente en cuestión y, finalmente, se infiere cómo actuará para conseguir sus metas a la luz de sus creencias. La atribución de creencias se ve constreñida en función de las capacidades cognitivas (incluidas las perceptuales), las necesidades epistémicas y la biografía del agente. La regla para atribuir creencias contempla el principio de que la exposición perceptiva a un fenómeno nos lleva a adquirir conocimiento sobre el mismo, pero que ese conocimiento se verá consolidado sólo en los aspectos que puedan ser de nuestro interés; por ende, se deben atribuir creencias cuando las verdades pertinentes a su interés o deseo hayan sido asequibles al agente, en virtud de su experiencia. Las creencias falsas requieren una genealogía especial bajo el supuesto general, ya explicado, de que el agente tendrá creencias mayormente verdaderas. La regla para atribuir deseos consiste en atribuir los deseos de aquellas cosas que un sistema consideraría buenas para sí, y de los medios para conseguirlos. Nuevamente, los deseos estrafalarios y perjudiciales exigen relatos especiales. En alguna medida, la atribución de estados mentales depende más de lo que el sistema hace, que de lo que dice.

Ahora bien: ¿con qué nivel de precisión se realizan estas atribuciones de creencias y deseos? Dennett sostiene que el lenguaje hace que nuestros deseos sean mucho más complejos que los que tendríamos en su ausencia; de hecho, muchas veces el entorno socio-lingüístico nos hace especificarlos hasta un punto que nada tiene que ver con su imprecisión inicial. Algo muy parecido ocurre con las creencias: el entorno nos hace formular a través de expresiones verbales precisas, convicciones que no estarían perfectamente adaptadas para ser verbalizadas.

¿Con qué frecuencia adoptamos la actitud intencional? La idea es que utilizamos esta estrategia todo el tiempo. Estamos acostumbrados a hacerlo y lo hacemos de manera automática, por lo que ignoramos el papel que desempeña en muchos de nuestros razonamientos. Así, una de las características de la psicología *folk* es que las explicaciones que construye no sólo descansan en la observación de ciertas regularidades, sino en la presentación de las acciones o elecciones tomadas como siendo razonables bajo tales circunstancias. Hay una apelación irreductible a la racionalidad del agente, sin la cual las explicaciones no podrían construirse. Entendernos los unos a los otros como sistemas racionales significa pensarnos como entidades cuya conducta se puede predecir por el método de atribución de creencias, deseos, y racionalidad, siguiendo un pequeño conjunto de principios rudimentarios. La racionalidad permite ligar lo que se asume que el agente cree y lo que se asume que el agente quiere para predecir su conducta. Se empieza con un supuesto de racionalidad perfecta (de acuerdo con el cual la gente conoce las implicaciones de sus creencias y no cree cosas contradictorias entre sí) que se va revisando según las circunstancias. En ese sentido, la psicología *folk* produce sus predicciones y explicaciones en un sistema normativo, trabajando sobre un supuesto de racionalidad ideal como punto de partida. Es evidente que este supuesto no se satisface enteramente. No obstante, la estrategia funciona bastante bien porque somos bastante racionales.

En función del análisis expuesto, Dennett dice que la psicología *folk* puede ser considerada como un “cálculo racionalista de interpretación y predicción: un método instrumental de interpretación idealizador que ha evolucionado porque funciona y que funciona porque nosotros hemos evolucionado”(Dennett, 1991a, p. 55). Si bien la frase queda justificada por todo lo que acabamos de exponer, resulta enriquecedor aclarar que por la explicación que se da de la estrategia y del modo en el que se la

utiliza, sería más preciso entender la actitud intencional como una destreza antes que como un cálculo, una habilidad que tenemos los seres humanos para relacionarnos entre nosotros y que permite las complejas relaciones sociales y culturales características de nuestra especie. Esto no quiere decir que Dennett piense que nuestro interés en los demás pase exclusivamente por querer saber qué van a hacer; más bien la idea que subyace a esta concepción es la de que nuestra capacidad para interpretarnos los unos a los otros depende de nuestra capacidad para predecir nuestros comportamientos.⁵

La Teoría de los Sistemas Intencionales -de la que aquí hemos ofrecido un mero esbozo- además de ser una teoría que dice mucho acerca del modo en que nos vinculamos los unos con los otros, debe ser entendida como un enfoque novedoso en filosofía de la mente, que resulta incompatible con las categorías ontológicas clásicas que han usado tradicionalmente realistas y anti-realistas a la hora de concebir el modo de existencia que cabe atribuir a los estados mentales. En contraste con la posición de que las creencias son algo subjetivo y misterioso, Dennett se acerca al realismo al abordar a las mismas como un fenómeno perfectamente objetivo. Sin embargo, se diferencia de un realismo *duro*, en el que se asume que tener estados mentales es tener estados internos que se individualizan funcionalmente de modo isomórfico a aquellos⁶; y es por esto que, desde su perspectiva, preguntar si alguien tiene una creencia no es equivalente a “la pregunta de si alguien está infectado por un virus determinado: una cuestión, de hecho, interna, perfectamente objetiva” (Dennett, 1987, p. 26). El modo de discernir las creencias será, en cambio, a partir del punto de vista de alguien que adopta una estrategia predictiva, y cuya existencia puede ser confirmada sólo por una evaluación del éxito de esa estrategia. Este último elemento acerca a nuestro autor al interpretacionismo. Esta ambigüedad con respecto a las etiquetas ontológicas instituidas, da cuenta de que para entender la propuesta dennettiana no se podrá recurrir a una simple clasificación bajo los rótulos de *Realista* o *Interpretacionista*; en su lugar, habrá que atender a los argumentos mismos, permitiendo que, eventualmente, se produzca una reformulación o un abandono de tales categorías. En un texto que Dennett (1991b) destina

⁵ En “Real Patterns”, Dennett (1991b) responde a una objeción a esta idea diciendo que, si bien es cierto que muchas veces la impredecibilidad de las personas es lo que las hace interesantes, esto sólo es posible porque tal impredecibilidad se da en un trasfondo de rutina que tiene un alto componente de predicción, del cual depende toda la interpretación.

⁶ Para un ejemplo de esto ver Fodor (1985, 1987).

específicamente a intentar explicitar mejor su postura sobre la existencia de los estados mentales, inscribe su teoría de los sistemas intencionales dentro de lo que llama un *Realismo Moderado*⁷. Sin embargo, concluye el texto volviendo a desconfiar del valor de tales etiquetas: “¿la opinión que defiendo aquí es un tipo de instrumentalismo o un tipo de realismo? Creo que la perspectiva en sí es más clara que cualquiera de las denominaciones, así que dejaré esta pregunta a cualquiera que se sienta inspirado por ella” (Dennett, 1991b, p. 51).

La pregunta sobre qué son las creencias, es formulada en *La Actitud Intencional* (Dennett, 1987) de manera indirecta, bajo la forma de qué es que alguien tenga una creencia o, en otros términos, qué es que *algo* sea un “verdadero creyente”. La respuesta que da Dennett es que ser un verdadero creyente simplemente significa ser un sistema cuya conducta es predecible exitosamente desde la actitud intencional. Dichos sistemas reciben el nombre de sistemas intencionales. Así, creer real y verdaderamente que *p* no es más que ser un sistema intencional para el cual *p* ocurre como una creencia según lo que dicta la mejor interpretación posible, esto es, según la interpretación que resulte más predictiva (Cfr. Dennett, 1987, p. 39). Si bien la decisión de adoptar la actitud intencional es subjetiva, el éxito o el fracaso que se tenga al ponerla en funcionamiento, son completamente objetivos.

El resultado de este procedimiento arroja un modo de entender las creencias, deseos e intenciones cercano al conductismo: lo que significa que alguien crea que *p* es que esa persona está dispuesta a conducirse de cierto modo en ciertas condiciones. Nótese cómo el fenómeno de la creencia, desde esta perspectiva, deja de estar revestido de un halo misterioso o definido por un criterio introspectivo, y pasa a formar parte de una realidad objetiva y discernible desde la tercera persona.

La idea de que cualquier sistema cuyo comportamiento se pueda pronosticar adecuadamente desde la actitud intencional es un creyente, resulta evidentemente polémica. Sin embargo, antes de realizar evaluaciones al respecto, deberíamos aclarar mejor quiénes serían, desde la estrategia en cuestión, los verdaderos creyentes.

⁷ La etiqueta de *Realismo Moderado* es obtenida en función de una suerte de escala en la que su posición debe ubicarse en un lugar en el que se la entienda como menos realista que el *Realismo de Vigor Industrial* de Fodor y que el realismo de Davidson, pero menos moderado que el interpretacionismo de Rorty (Dennett, 1991b).

Dijimos que los sistemas intencionales son aquellos para los cuales la estrategia intencional funciona con éxito. Según esta definición, y teniendo en cuenta que se pueden realizar predicciones bajo este formato acerca de un espectro amplísimo de sistemas –personas, órganos, computadoras, ranas, termostatos-, ¿qué objetos no serían sistemas intencionales? ¿No corremos el riesgo de terminar afirmando que la estrategia se aplica a todo tipo de entidades? Un ejemplo de cómo la estrategia intencional pareciera poder usarse para cualquier cosa es dado por el mismo Dennett: “Por ejemplo parece que el atril de este salón de conferencias puede ser interpretado como sistema intencional, totalmente racional, creyendo que está colocado actualmente en el centro del mundo civilizado, y deseando sobre todo permanecer en ese centro” (Dennett, 1987, p. 33). Si bien es cierto que la predicción del ejemplo no va a ser contradicha, en la medida en que el atril no va a salir corriendo, la predicción de que el atril va a permanecer en el mismo lugar en el que está, a menos de que alguien lo mueva, podría haber sido obtenida con sólo considerar que se trata de un pedazo de madera atraído hacia el centro de la tierra por la fuerza de gravedad (actitud física), o que fue diseñado para sostener papeles en el lugar en el que se lo deposite (actitud del diseño). Así, podemos decir que no estamos frente a un sistema intencional, dado que la estrategia no aporta ningún poder predictivo adicional al que se tenía antes de emplearla.

Por el contrario, en el caso de los sistemas intencionales la postura intencional es, con frecuencia, la única que nos brinda un poder predictivo que no podríamos obtener de otra forma. Sin embargo, no parece haber una equivalencia absoluta en lo que las distintas personas podrían y no podrían hacer desde la actitud intencional:

Para la gente que tiene una aptitud mecánica limitada, la interpretación intencional de un simple termostato es una muleta práctica y sumamente inocua, pero los ingenieros que hay entre nosotros pueden entender del todo su funcionamiento interno sin la ayuda de esta antropomorfización. (Dennett, 1987, p. 34)

Además de estos ejemplos de la vida real, Dennett utiliza experimentos mentales que incluyen al omnisciente físico laplaceano, o a marcianos superinteligentes que pueden predecir nuestros comportamientos desde la actitud física. La realidad física de todo lo existente nos lleva a sostener esas posibilidades en principio: con respecto al empleo de distintas estrategias para predecir el

comportamiento de un atril y el de una persona, Dennett señala que “Se debe insistir en que esto no supone ninguna diferencia en la esencia, sino simplemente una diferencia que se refleja en nuestra limitada capacidad como científicos” (Dennett, 1987, pp. 33, 34). Ahora bien, si la necesidad de utilizar dicha estrategia depende de las capacidades del intérprete, ¿cómo se puede hablar de un criterio objetivo con respecto a lo adecuado de su aplicación?

La incapacidad de la postura intencional para predecir de formas sumamente precisas termina siendo una fortaleza, dado que permite que la estrategia pueda ser explotada incluso en casos de sistemas muy complejos en los que otro tipo de predicción resultaría de una enorme laboriosidad. Así, la diferencia de complejidad que los sistemas en cuestión guardarían con otros sistemas cuyo comportamiento sí podría ser predicho fácilmente desde otra actitud, no se corresponde simplemente con el grado de ignorancia del intérprete con respecto al sistema, sino con una diferencia objetiva de magnitudes con respecto a la cantidad de información requerida para la predicción, que permiten la elaboración de modelos completamente diferentes.

Creo que en algún punto esta diferencia objetiva de magnitudes de información necesaria para realizar una predicción se puede entender como una cuestión matemática. Pensemos en la notación científica como recurso. El número $3,18 \times 10^{10000}$ se podría escribir, con algo de tiempo, expresando cada una de sus unidades. Dennett, en virtud de su materialismo, concede la posibilidad hipotética de describir desde la actitud física el estado de cosas que puede ser predicho desde la actitud intencional. Sin embargo existe, como en el caso de los números, una diferencia significativa que hace que, de hecho, nunca utilicemos el mismo recurso para expresar el número $3,18 \times 10^{10000}$ que el número 67. Esta diferencia que es netamente de magnitud no es, por eso, poco significativa, ni relativa en absoluto; es en un sentido análogo que Dennett habla de una objetividad en la adecuación de la aplicación de cada estrategia a cada caso. La notación científica, indispensable en un sentido práctico para expresar el número en cuestión, se vuelve tan innecesaria para expresar el 67 como lo es la estrategia intencional para predecir el comportamiento del atril. Esas diferencias de necesidades epistemológicas son el reflejo de otras diferencias de escala en la complejidad de la constitución de los sistemas. Frente a la usual distinción entre las diferencias de grado y de categoría, y la preeminencia

filosófica que se le suele otorgar a las segundas por sobre las primeras, Dennett realiza un análisis en el que un examen atento de las diferencias de grado termina revelando cómo las mismas, al producirse en gran escala, generan a veces diferencias que son de categoría, en el sentido de que se produce un cambio de nivel explicativo, aunque dicho salto resulta explicable.

A su vez, el criterio de atender a la ganancia predictiva que genera la atribución de creencias a un sistema (o no) arroja resultados en los que el conjunto de los verdaderos creyentes es bastante consistente con lo que sabemos acerca del origen de los seres o dispositivos a los que finalmente atribuimos intencionalidad. Atribuir racionalidad a una piedra no sólo sería poco interesante, sino que además no contamos con ningún indicio de que para la misma pueda haber resultado útil haber desarrollado una capacidad que le permita almacenar información sobre el mundo que la rodea para usarla para su propio beneficio, como sí podemos pensarlo de los monos, y, de modo análogo, respecto de algunos robots.⁸ Uno de los puntos distintivos, entonces, de esta teoría de los sistemas intencionales es que sostiene, en contra de parte importante de la tradición (Davidson, 1967, 1975, Brandom, 1994, McDowell, 1994), que las creencias no son una posesión exclusiva de las criaturas con lenguaje. Si bien Dennett, a diferencia de estos autores, no considera que el lenguaje sea condición de posibilidad para tener creencias sin más, años después de desarrollar la teoría de los sistemas intencionales propone una distinción entre las creencias en general - que podemos compartir con sistemas carentes de lenguajes- y las opiniones, que son los estados mentales que sí requieren de una formulación lingüística. Un ejemplo típico de opinión sería, por ejemplo, la de que *el año que viene es bisiesto* (Dennett, 1998). De este modo, puede hacer justicia a la importante diferencia entre nuestras creencias verbalmente articulables y aquellos estados doxásticos (humanos y no humanos) que no tienen ese grado de especificación verbal posible.

Otra metáfora útil para ilustrar que la ganancia predictiva que se obtiene al adoptar la actitud intencional es de carácter objetivo, puede pensarse recurriendo a

⁸ Con respecto a la posibilidad de crear máquinas inteligentes, si bien Dennett aclara que no tenemos motivos en principio para sostener que el contenido mental nuestro diferirá siempre del que se pueda atribuir a un robot super-desarrollado, piensa que con las tecnologías disponibles sería imposible programar una computadora de forma tal que produzca lo que nuestros cerebros orgánicos: el control de la actividad intencional veloz e inteligente de los seres humanos. (Véase Dennett, 1987, cap: “El pensamiento veloz”.)

la idea de un cuadro en el que, a partir de muchas pinceladas discontinuas sobre un fondo ruidoso, se forme la figura difusa de un perro. Funes el memorioso podría acordarse de cada uno de los puntitos y pintarlo exactamente igual. Sin embargo, si tuviera que competir en Pictionary, realizando un dibujo rápido que haga reconocer a su equipo el cuadro del perro, contra alguien que sólo recuerde que los puntitos formaban algo parecido a la figura de un perro, se encontraría en evidente desventaja. Lo que se pierde Funes, además de los puntos de Pictionary, es el perro borroso: ese “modelo” accesible sólo para aquellos que pueden olvidar diferencias (Dennett, 1991b).

Eso es lo que les sucedería a los marcianos super-inteligentes: podrían predecir absolutamente todos nuestros comportamientos desde la actitud física, pero se perderían algo perfectamente objetivo: “Los modelos del comportamiento humano que se pueden predecir desde la actitud intencional y solo desde esa actitud, y que sustentan las generalizaciones y las predicciones” (Dennett, 1987, p. 35). Ellos podrían predecir el comportamiento de una persona que se adelanta de forma bruta en el supermercado para agarrar la última bolsa de carbón, al tiempo en el que justo la está por agarrar lentamente otra persona, que se encontraba eligiendo distintos cortes de carne hace veinte minutos. El marciano podría predecir este comportamiento a partir de una inferencia que toma en consideración los impulsos eléctricos en el cerebro y todos los movimientos de su cuerpo; también podría haber predicho, con el mismo mecanismo, que la otra persona iba a agarrar la bolsa de carbón exactamente al mismo tiempo; pero se encontraría completamente sorprendido ante la capacidad del primero para saber que tenía que apurarse: ¿cómo puede haber predicho este comportamiento de la otra persona, sin tener acceso a sus conexiones nerviosas? Gracias a los modelos del comportamiento humano, en términos de expectativas, intereses y deseos, a los que sólo se tiene acceso desde la actitud intencional. Es por ello que no sería del todo adecuado decir que la posición de Dennett consiste en un eliminativismo acerca de las creencias y los deseos: antes bien, para él la postura intencional es irreductible y no se la puede abandonar sin que se produzca una pérdida.

Lo que la actitud intencional capta puede entenderse como *modelos o patrones* de comportamiento. Estos son los términos que Dennett emplea en *La actitud Intencional* (1987) y en “Patrones Reales” (1991b), respectivamente, y que

considero que deben interpretarse de modo intercambiable. Dennett llama la atención sobre la inconsistencia que alberga la idea de *patrón indiscernible* -análoga a la de *imagen invisible*, o *sinfonía silenciosa*- dando cuenta de esta dimensión dual según la cual si bien los patrones tienen, en alguna medida, una existencia objetiva que los distingue de la simple imaginación, es necesario que haya alguien para captarlos: “Este vínculo laxo pero irrompible con los observadores o perspectivas es, por supuesto, lo que hace que “patrón” sea un término atractivo para alguien ubicado entre el instrumentalismo y el realismo de vigor industrial” (Dennett, 1991b, p. 32, mi traducción). Los modelos que se pueden construir desde la actitud intencional son objetivos, están allí en realidad, pero en algún punto dependen de nosotros, en la medida en que están compuestos, en parte, por nuestras reacciones subjetivas a lo que está ahí. Por otra parte, dado que nadie es completamente racional, ni tiene una memoria perfecta, ni es un observador perfecto, la estrategia en ciertos casos no funciona. En ese sentido, debemos decir que los modelos ordenan las conductas con gran rigor, pero no infaliblemente (o, como dice Dennett en “Patrones Reales”: en algunos casos, los patrones pueden tener cierto porcentaje de “ruido”). Desde el realismo moderado que se está defendiendo, no se podría afirmar que en tales ocasiones hay creencias y deseos que escapan a la estrategia, pero tampoco se cae en un subjetivismo o interpretacionismo puro, porque el éxito o el fracaso de adoptar la actitud en un momento determinado, es una cuestión objetiva. Incluso puede haber interpretaciones rivales muy distintas, y una relatividad del interés de las distintas atribuciones, aunque, nuevamente, cuán buenas sean para predecir, seguirá siendo una cuestión objetiva.

Dennett contempla, inclusive, la posibilidad de que existan interpretaciones rivales, muy diferentes entre sí, que sean igual de exitosas a la hora de predecir (Dennett, 1987, 1991b): “La presencia objetiva de un modelo (cualquiera que sean sus imperfecciones) no excluye la presencia objetiva de otro modelo (cualquiera sean sus imperfecciones)” (Dennett, 1987, p. 38). Es en este punto en el que Dennett (1991b) se define como un realista más moderado que Davidson, y señala que el motivo de su desacuerdo puede entenderse como una diferencia de opinión sobre el alcance del principio de la indeterminación de la traducción de Quine (1960). Según el principio de la indeterminación de la traducción, llevado al ámbito de la interpretación por Davidson (1973), habría algunos casos en los que hay dos interpretaciones rivales, ambas compatibles con la misma evidencia empírica, entre las que es imposible

determinar cuál es la más acertada. El principio profundo detrás de esto es que no hay, más allá de nuestra ignorancia, una interpretación que sea de hecho más acertada que la otra, sino que, sencillamente, hay indeterminación respecto de cuál es más acertada. Ahora bien, Davidson entiende que el fantasma de la indeterminación es prácticamente trivial, por lo que brinda ejemplos como el de la indeterminación que existe entre las afirmaciones rivales de que la temperatura es de 32° Fahrenheit y 0° Celsius, es decir, ejemplos de modos distintos de medir una misma realidad física. En analogía con esto, se debería pensar la indeterminación que Davidson admite en la atribución intencional como involucrando diferentes formas de conceptualizar el mismo hecho. En cambio, Dennett piensa que el problema de la indeterminación podría darse también en casos que incluso lleguen a producir predicciones distintas en casos decisivos. Dado que todas las interpretaciones pueden tener fallas en casos puntuales (que se pueden hacer inteligibles agregando alguna cláusula *ad hoc*), se puede pensar que incluso el fallo de uno de los esquemas para predecir un comportamiento en una situación particular no es necesariamente una prueba de que el esquema sea erróneo: se puede decir que son dos esquemas de interpretación que posibilitan predicciones confiables la mayoría de las veces, aunque estén en desacuerdo en momentos puntuales. De esto no se sigue que no se pueda demostrar que una interpretación es simplemente errónea, sólo se está diciendo que puede haber buenas interpretaciones rivales, con un poco de ruido, frente a las cuales sería imposible decir que una es más verdadera que la otra.

Como dijimos, los modelos o patrones que podemos construir desde la actitud intencional tienen un carácter dual, en la medida en que, tienen un componente objetivo, pero al mismo tiempo, dependen de una capacidad humana de discernirlos (dado que la idea de patrón indiscernible es contradictoria). Las posibilidades de encontrar más de un patrón en una misma realidad pueden variar, por ejemplo, entre un modelo que sea más compacto pero tenga más “ruido”, y uno menos compacto pero con menos “ruido”; las ventajas de uno y otro se pueden sopesar en función de cuán peligrosos son los errores, cuán importante es la velocidad, qué cantidad de recursos se pueden destinar a ese análisis, entre otras variables. Pero lo interesante es que no se trata de elecciones que tomamos de forma deliberada e individual, sino más bien de cambios que se han ido incorporando a nuestro modo de vincularnos cognitivamente con el mundo gracias a procesos de evolución genética

(que moldea nuestros órganos sensoriales) y a procesos de evolución cultural añadidos a aquellos. El producto de estos procesos evolutivos da lugar a la *imagen manifiesta* de la que hablábamos anteriormente que incluye, por supuesto, a nuestra psicología *folk*, y que permite que podamos encontrar cierto tipo de regularidades en el bombardeo de información desordenada que sería de otro modo el mundo para nosotros: “La ontología creada por la imagen manifiesta tiene, por lo tanto, un fundamento muy pragmático” (Dennett, 1991b, p. 11, mi traducción). Así, una de las principales características de nuestra “imagen manifiesta” es que tenemos una facilidad muy pronunciada para percibir patrones que son visibles desde la actitud intencional.

Lo dicho con anterioridad descansa en la idea de que el modelo o patrón captado por la actitud intencional no debe su realidad a un isomorfismo con un hipotético lenguaje del pensamiento. Dennett afirma la existencia del primero, sin comprometerse con la del segundo. Del hecho de que la descripción de la actitud intencional produzca como resultado un modelo verdadero y objetivo en el mundo, no se sigue que el mismo sea “producido por otro modelo real aproximadamente isomórfico a él dentro de los cerebros de seres inteligentes” (Dennett, 1987, p. 43). Con esto se diferencia, a su vez, de un realismo vigoroso como el de Fodor (1985, 1987). Como señalamos con anterioridad, muchas veces el entorno nos hace formular de manera más precisa convicciones que no tenían una formulación lingüística tan ajustada, o especificar deseos que eran más difusos antes de su enunciación. Así, la idea de que es evidente que las creencias y deseos son como oraciones en la mente, podría derivarse de una focalización en el resultado del proceso de verbalización de nuestros estados mentales, más que de la naturaleza última de dichos fenómenos.

Una intuición típicamente realista sostendría, por otra parte, que si bien puede haber interpretaciones rivales que lleven a distintas atribuciones de creencias frente a las cuales no tengamos elementos suficientes para establecer cuál es más adecuada, hay una de ellas que *realmente* es la más adecuada, en virtud de la existencia de ciertos hechos más profundos que determinan los contenidos de las creencias. Dennett se opone a esta intuición, y ese es el modo en que debemos entender el alcance de su forma de entender la indeterminación de los contenidos (cuestión que toma relevancia en su discusión sobre la distinción entre intencionalidad intrínseca y derivada, que será desarrollada a continuación).

Esta forma de entender las creencias influye en cómo se entenderá esa propiedad tan particular que tienen las mismas de ser *acerca de* algo. Pensemos en un dispositivo con respecto al cual emplear la actitud intencional sería, a lo sumo, un atajo práctico, pero no algo objetivamente provechoso. Un termostato, por ejemplo, tiene un “comportamiento” sumamente limitado, que puede ser fácilmente predicho desde la actitud del diseño, pero carece de una conexión con el mundo que le permita enriquecer la semántica de sus “creencias”. Si a ese dispositivo se le agregaran canales de comunicación con el mundo, que fueran haciendo aportes cada vez más específicos a la semántica del sistema, llegaría un momento en el que a ese sistema se le podría otorgar una interpretación semántica prácticamente (aunque no en principio) única. Como consecuencia de ello, se podría decir que esos estados internos son *acerca de la temperatura de esa habitación*. En este planteo encontramos una reformulación de la noción de representación. “Lo que hace que algún rasgo interno de algo sea una representación sólo podría ser su papel en la regulación de un sistema intencional” (Dennett, 1987, p. 41). La representación deja de ser algo de *todo o nada*, no hay un momento mágico que permita identificar nítidamente el inicio de la misma; antes bien, esta posee una naturaleza gradual y puede encontrarse en sistemas de muy distinto tipo. “El termostato tiene una representación mínimamente exigente del mundo (...) Por fin se llega a nosotros. Estamos conectados al mundo de manera tan múltiple e intrincada que casi no hay sustitución posible” (Dennett, 1987, p. 41). No se trata de fenómenos de naturaleza metafísicamente diferente, sino que entre sí guardan diferencias de complejidad, pero en un grado tan grande que pensar la organización interna de algo tan simple como un termostato no nos va a ayudar a entender mejor nuestra organización interna.

Intérpretes de Dennett como Brook y Ross (2002) sostienen que, aunque la teoría de los sistemas intencionales ha sido tomada como el corazón de su aporte filosófico, lo más interesante e innovador que tiene su trabajo se debe más bien a su concepción sobre cómo está delimitado el mundo, según el esquema explicativo que le da forma a la postura intencional. Tal esquema explicativo está moldeado por una particular concepción de la naturaleza de la intencionalidad, así como por un rol sumamente activo de la selección natural para que la misma llegue a ser lo que es.

Si bien, como intentamos mostrar, la aplicación de la estrategia intencional no funciona igual de bien para los distintos tipos de sistemas, y permite marcar diferencias entre el modo en que se puede decir de una computadora, un perro y una persona que creen algo, lo que sí sostiene es que no podemos apelar a hechos profundos sobre los que asentar esas diferencias, hechos que nos permitan sostener que somos capaces de pensar de una forma esencialmente diferente a como son capaces de hacerlo otras criaturas o artefactos. Esto puede explicarse aún mejor a partir de la postura tomada por Dennett con respecto a la distinción entre la intencionalidad intrínseca y la derivada.

La mente como producto de la evolución: Intencionalidad derivada

La noción de *intencionalidad* a la que se apela en filosofía de la mente para explicar la mente humana, sus operaciones y sus productos, es una noción medieval reintroducida por Brentano (1874). La misma hace referencia a la propiedad distintiva de ciertos fenómenos que consiste en ser *acerca de* o estar *dirigidos a* o *representar* otras cosas: estados mentales, expresiones lingüísticas, cierto tipo de comportamientos y de entidades (cuadros, libros, artefactos, etc.) manifiestan ese rasgo en algún grado. Brentano había enfatizado, sin embargo, que la intencionalidad era la marca distintiva e irreductible de los fenómenos mentales (y sólo de ellos, pues sostenía que no hay fenómenos físicos que sean intencionales), sentando las bases de enfoques que, aunque se declaren naturalistas, comparten rasgos importantes con los de la tradición filosófica clásica y moderna, básicamente no naturalistas. Esta perspectiva es mantenida contemporáneamente por filósofos de la mente como Searle (1983), Dretske (1985) y Fodor (1987), quienes siguen sosteniendo que existe una distinción profunda entre la “intencionalidad intrínseca” u “original” y la “intencionalidad derivada”. La primera hace referencia a aquella intencionalidad que es exclusiva de las mentes humanas, que no está dada por nada ni nadie, pues las mentes humanas son una especie de punto cero, las dadores de significado (idea que Dennett (1987, p. 255) compara ingeniosamente con el motor inmóvil aristotélico, que sólo es movido por sí mismo). Ésta se diferencia de la “intencionalidad derivada”, propia de los artefactos que hemos construido (incluidas las expresiones lingüísticas) y que sólo poseen intencionalidad porque nosotros se la hemos otorgado. Dennett se propone rebatir los fundamentos de esta distinción, que se ha

mantenido implícita durante años, y que resulta profundamente intuitiva (Dennett, 1987). El problema filosófico de fondo es si, independientemente del poder que cualquier intérprete u observador pueda tener para descubrirlos, hay efectivamente hechos profundos que determinan los contenidos intencionales humanos, lo que los hace ontológicamente diferentes a los que podemos atribuir, por ejemplo, a un texto. Para explicitar las posturas en disputa, Dennett (1987, cap: “La evolución, el error y la intencionalidad”) expone algunas de las consideraciones que se desprenden del examen de un ejemplo paradigmático de intencionalidad derivada, que referiremos a continuación.

Un caso paradigmático de intencionalidad derivada sería el tipo de intencionalidad que tiene, por ejemplo, una máquina expendedora de gaseosas, que acepta o rechaza una moneda determinada. Ahora bien, si la máquina acepta un pedazo de metal de similares características que la moneda en lugar de la moneda, ¿estamos frente a un caso de falsa representación? ¿Podemos hablar de intencionalidad en los aparatos? En los artefactos diseñados, el significado de los “estados” de los aparatos se debe a la intención de los diseñadores, por eso sólo se trata de una intencionalidad derivada. Así, teniendo en cuenta el objetivo para el que fue diseñada, podemos decir que cuando la máquina acepta una moneda (lo que la llevará a entregar la gaseosa) entra en un estado *Q*, que “significa” algo como “percibo una moneda auténtica de *X* país, por el monto *Y*”, y que cuando entra en el estado *Q* frente a otro pedazo de metal, está “cometiendo un error”. Ahora bien, si la misma máquina es trasladada a otro país, en el que debe reconocer otras monedas, de curso legal en ese nuevo entorno, por el monto a pagar, los casos de error ya no serían los mismos. Dichos casos de error pueden ser enunciados de muchas formas distintas. Parece haber, entonces, un amplio grado de indeterminación en la interpretación de lo que los estados de la máquina significan. Si pensamos que al aceptar una moneda del país 2 está significando “moneda del país 1”, la aceptación debe ser entendida como un error. Pero también podemos pensar que está significando “moneda del país 1, o moneda del país 2”, o bien “pedazo de metal cilíndrico de un diámetro de 13 mm y un peso de 2 g”, etc.; no existiendo la posibilidad optar por sólo una de las interpretaciones. Para decidir cómo interpretar el error, podemos utilizar un criterio adaptativo: deja de ser relevante identificar la intención de quienes diseñaron la máquina, y comienza a importarnos pensar en las intenciones de quienes la mudaron de entorno: en su nuevo nicho fue seleccionada

por su capacidad para reconocer las monedas del país 2, en virtud de ciertas constricciones en el diseño que la hacen funcionar bien desarrollando su nueva tarea. Aún bajo esta interpretación, la indeterminación permanece y existen casos en los que resulta difícil pensar cuál debería ser el criterio en torno al cual interpretar el contenido (si la máquina hubiera sido enviada a otro país por error, pero luego de un tiempo su nueva utilidad hubiera sido reconocida y se hubiera convertido en un motivo para conservarla, ¿cuándo exactamente cambiaría el significado de lo que máquina reconoce con acierto y lo que acepta erróneamente?). Este grado de indeterminación, que permanece en el caso de la máquina, no parece un gran problema: no tenemos la sensación de carecer de la capacidad para conocer algo relevante, ya que no hay ningún hecho más profundo a descubrir que determine cuál sería el contenido intencional de la máquina al aceptar o rechazar un objeto; lo cual es equivalente a decir que no se trata de un caso de intencionalidad intrínseca. Las consideraciones realizadas hasta aquí no entran en conflicto con la posición de los defensores de la existencia de la intencionalidad intrínseca, más bien explicitan la diferencia entre la intencionalidad humana y la que atribuimos a artefactos, e incluso sirven para marcar los rasgos específicos de dicho fenómeno. Dennett disiente con sus críticos en que él piensa que lo mismo que ocurre en el caso de la máquina, también vale para nosotros. Los defensores de la intencionalidad intrínseca pensarían que en nuestro caso esos hechos más profundos sí existen, con independencia del poder que cualquier observador o intérprete pueda tener para conocerlos. Desde la visión de dichos contrincantes, incluso en los casos en que es imposible discernir epistémicamente entre dos interpretaciones rivales sobre nuestros estados mentales, hay una diferencia ontológica: nosotros nos encontramos *realmente* o en un estado o en otro.

Dennett sostiene que esta distinción no puede sustentarse, puesto que toda forma de intencionalidad es derivada. En definitiva, con este argumento, Dennett “explícitamente y de manera muy firme, sienta su postura contra todas las variedades de la idea de que el contenido mental puede ser intrínseco a los estados cerebrales” (Brook y Ross, 2002, p. 8). Para sostener esto se apoya en argumentos evolucionistas. Una de las implicancias más importantes de la teoría de la evolución por selección natural es que permite explicar que exista diseño sin diseñador (Dennett, 1995, 2001a, 2001b, 2006). Sin embargo, a pesar de que considera que la Selección Natural es un proceso mecánico, sin objetivos, y sin una visión del futuro,

Dennett defiende que tratar a la Madre Naturaleza desde la actitud intencional puede ser, en ocasiones, muy útil. Desde esa perspectiva podemos decir que la Naturaleza *discrimina, reconoce, aprecia y elige* entre muchas características y relaciones, probando ser exquisitamente sensible a las “razones de ser” de sus diseños (sus famosas razones de ser de “flotación libre”⁹). Tratar al proceso de evolución desde la actitud intencional permite que nos veamos a nosotros mismos como artefactos diseñados por ella.¹⁰ A partir de esta visión, podemos reconocer que hay objetivos, sin suponer que están representados en la mente de nadie: “cuando la selección natural selecciona puede “elegir” un diseño determinado *por una razón más que por otra* sin si quiera representar conscientemente- io inconscientemente!- la elección o las razones” (Dennett, 1987, p. 265).

Por otro lado, desde la postura de los contrincantes, la intencionalidad de ciertos organismos biológicos (los seres humanos) sería intrínseca por ser independiente de las intenciones o propósitos de un organismo externo -a diferencia de, por ejemplo, un robot-; sin embargo, aplicando la actitud intencional a la Madre Naturaleza, podemos decir que, dado que ella ha “diseñado” los organismos para que sus estados mentales les permitan relacionarse con su entorno de cierta forma, la intencionalidad de aquellos sí depende de las “intenciones” y “propósitos” de la Madre Naturaleza. Bajo esta nueva formulación, nuestra intencionalidad parece tan derivada y, por ende, tan sometida a la indeterminación de la interpretación como la de los artefactos creados por nosotros. “De manera que si ha de haber alguna

⁹ En muchos estudios sobre comportamiento animal los etólogos se enfrentan a dilemas sobre cómo interpretar determinadas estrategias ingeniosas. Por ejemplo, hay un pájaro que cuando un depredador se aproxima a sus polluelos, realiza un extraño baile simulando tener un ala rota, presuntamente para que el depredador calcule que le conviene perseguir a la madre que a los huesudos polluelos. Cuando el depredador está lo suficientemente cerca, la madre escapa volando, dándoles tiempo de huir también a sus hijos. Hay experimentos que nos permitirían apoyar en distinto grado las hipótesis acerca de cuánto de este comportamiento es entendido o pensado por el pájaro. La postura de Dennett al respecto es que, incluso si descubriéramos que el pájaro se comporta de un modo completamente instintivo y sin tener ningún tipo de noción de la utilidad de su conducta, eso no iría en detrimento de que podamos decir que el extraño baile sigue teniendo una “razón de ser”, por la cual fue seleccionado naturalmente, y es por ello que, incluso en estos casos, la pregunta *por qué* no pierde pertinencia. Así, la noción de razones de ser de “flotación libre” es no representacional: “La naturaleza apreció estas razones sin representarlas” (Dennett, 1987, p. 281).

¹⁰ Dennett señalará que el límite para aplicar la actitud intencional a la Madre Naturaleza se encuentra en que ésta carece de capacidad representacional y, por ende, no posee el poder de anticipación que nosotros sí tenemos. (Véase Dennett, 1987, cap: “La evolución, el error y la intencionalidad”)

intencionalidad original –original sólo en el sentido de no derivar de ninguna otra fuente ulterior- la intencionalidad de la selección natural se merece el honor” (Dennett 1987, p. 281). De donde resulta que los rasgos intencionales originados gracias a la evolución natural no pueden ser explicados de manera consecuente en función de propiedades intrínsecas, sino que deben ser comprendidos como rasgos que dependen de la adopción de una estrategia atributiva.

En consonancia con algunos aspectos señalados en el apartado anterior, podemos decir que la teoría de la mente de Dennett muestra cómo las consideraciones evolucionistas llevan en una dirección opuesta a la de las posiciones realistas acerca del contenido, que suponen que los mismos son algo completamente determinado e independiente de la interpretación. Desde esta visión se siguen dos implicancias que obligan a dejar de lado importantes supuestos de muchas de las teorías filosóficas de la mente (y del significado). La primera de ellas consiste en que si el contenido de nuestros pensamientos es derivado, es algo sobre lo que no tenemos ninguna autoridad especial, por lo que tendremos que abandonar la idea del acceso privilegiado de la primera persona (y otras tesis relacionadas: infalibilidad, incorregibilidad, asimetría). La segunda, que acabamos de explicar, es que no existen hechos más profundos que fijen o determinen los significados de nuestros pensamientos, una piedra de toque que permita distinguir nuestra intencionalidad de la que podemos encontrar en dispositivos con mera intencionalidad derivada.

Permítasenos una disquisición más sobre el efecto de las consideraciones evolucionistas para la teoría del significado. Para hacerlo, tomemos el caso de una reacción de “llamada de atención” (de la que volveremos a hablar más adelante) que, en los peces, consiste en un estado de alerta desencadenado a raíz de una percepción de patrones visuales que presentan simetría con respecto a una eje vertical (como lo es la cara de un depredador vista de frente: una boca al medio, ojos y aletas a los lados ubicados de forma simétrica) ¿Cómo debe interpretarse el contenido del estado mental de los peces en estos casos? En otras palabras, cuando el pez huye frente a una simetría ecológicamente insignificante (como un alga que casualmente es simétrica), ¿está cometiendo un error? Esto último dependerá de si consideramos que su dispositivo de sensibilidad a estos patrones debe entenderse como un detector de *depredadores mirando de frente*, o como un detector de *depredadores, o comida, o pareja mirando de frente*, o si se le agregan la cantidad de disyuntores necesarios,

incluyendo los ecológicamente irrelevantes, entendiendo que el dispositivo fue diseñado de este modo dado que resulta conveniente en términos de costo y beneficio incluir ese pequeño grado de imprecisión. Ahora bien, ¿cuál es exactamente el contenido? En lugar de intentar responder a esa pregunta, la postura dennettiana conduce, como es de esperar, a abandonar la búsqueda de un contenido mental preciso e independiente de la interpretación. Esta estrategia podría tener aceptación para el caso de la detección de patrones con simetrías verticales, dado que a lo único que se puede apelar para responderla es a la función ecológica del dispositivo. Lo mismo ocurre con los artefactos creados por el hombre. Así, vemos cómo desde consideraciones evolucionistas y de ingeniería “inversa” o “retrospectiva”¹¹, el contenido intencional a atribuir estaría ligado a la función para la que fue diseñado.

Sin embargo, otra ilusión que Dennett intenta tirar por la borda en el mismo capítulo es la de la función determinada: la idea de que toda función portaba en su génesis la idea clara y precisa de su razón de ser. Los estudios evolucionistas van en una dirección contraria a dicha idealización, mostrando cómo la Madre Naturaleza más que una optimizadora, es una “satisfactora” que hace lo que puede con lo que va encontrando a la mano. Uno de los científicos que más se ha dedicado a resaltar esta idea es Stephen Jay Gould (1980), de quien volveré a hablar en el capítulo 4 de este trabajo. Ahora bien, dado que no hay ningún hecho profundo que pueda resolver la cuestión de para qué sirve algo, no puede haber ninguno que pueda responder el qué significa. “No hay más fundamentos sólidos para lo que podríamos llamar la funcionalidad original que la que hay para su descendiente cognitivista, la intencionalidad original” (Dennett 1987, p. 284). Dennett le reconoce a Gould que la Madre Naturaleza funciona como una chapucera oportunista más que como un ingeniero ideal. Pero donde Gould “encuentra una paradoja en esta mezcla de chapucería y teleología” (Dennett 1987, p. 283), Dennett sostiene que no hay ninguna paradoja, sino simplemente indeterminación funcional. La función no está determinada por una representación de la Madre Naturaleza sobre lo que quería crear, sino en nuestras interpretaciones de para qué sirven ciertas características, donde dichas interpretaciones parten de asumir un marco de optimalidad que no

¹¹ Esta expresión es tomada por Dennett (1987) del ámbito de la ingeniería, como herramienta de análisis que consiste en tomar como punto inicial de estudio al producto final, y procurar averiguar las posibles razones por las cuales este fue diseñado. El adaptacionismo es el método que resulta de esta aplicación en la biología.

tiene que corresponderse con los fenómenos reales de la evolución por selección natural.

En función de todo lo expuesto, podemos decir que de esta concepción se desprende la idea de que no hay un salto ontológico que implique una aparición milagrosa de un reino de lo mental que establezca una ruptura “de todo o nada” entre nosotros, por una parte, y posibles robots, artefactos complejos o animales no humanos, por otra. En palabras de Clark, la postura intencional implica que “humanos, perros, insectos e incluso meros termostatos son todos capaces de creer y desear, fundamentalmente, en el mismo sentido teórico” (Clark, 2002, p. 187). Compartimos con ellos la ausencia de una intencionalidad intrínseca. Tal posición entra en conflicto con muchas de nuestras intuiciones, todas las cuales parecen indicar que hay una diferencia profunda en el modo en que nosotros significamos y el modo en que lo hace un artefacto que no posee ningún tipo de vida interior verdadera. Por más que estos últimos dispongan de mecanismos de entrada y salida y de información, no hay ningún intermediario que experimente, tome decisiones, sea el beneficiario y el último responsable de lo que “significan” sus estados intencionales. En cambio, tendemos a pensar que esa sí sería una buena descripción de lo que ocurre en la mente humana. En el apartado siguiente realizaré una aproximación a la propuesta teórica de Dennett sobre cómo entender a la mente humana desde uno de sus aspectos fundamentales: la conciencia.

Teoría de la conciencia: El Modelo de las Versiones Múltiples

Dennett realizó, por otra parte, importantes contribuciones a los estudios sobre la conciencia. Si bien no abordaré ese tópico de manera exhaustiva en este trabajo, señalaré algunos aspectos básicos sobre su concepción de la conciencia, dado que se relacionan en forma directa con el tema central a desarrollar: su concepción del “yo”. En su examen crítico de modelos y teorías clásicas y contemporáneos sobre la conciencia, y en la elaboración de su propio modelo alternativo, llamado “Modelo de las Versiones Múltiples”, se evidencian con claridad los resultados de la cooperación entre ciencias y filosofía pregonada por Dennett. La novedosa visión que ofrece del flujo de la conciencia, complementaria a una forma de entender el funcionamiento del cerebro, adquiere una fuerte relevancia en el marco del planteo del problema del “yo”.

Según Dennett (1989, 1991a) una vez que hemos abandonado el dualismo ontológico, que permitía identificar al “yo” con la sustancia pensante con la que se vincularía nuestro cuerpo, se ha llegado a pensar que la alternativa materialista al eliminativismo con respecto al “yo”, es la de identificar al mismo con una parte puntual en el cerebro (por ejemplo, Popper y Eccles, 1977). La concepción dennettiana del funcionamiento del cerebro impide la aceptación de esta forma alternativa de salvar de manera materialista la excepcionalidad de nuestros “yoes” humanos, dado que en el cerebro no podemos encontrar nada que se parezca a un “yo”.

Con respecto a la conciencia, Dennett mantiene que hay que estudiar este fenómeno desde un abordaje científico de tercera persona, que él mismo formula y denomina *Heterofenomenología* (1991a, cap: “Un método para la fenomenología”). El método heterofenomenológico puede verse como una aplicación especial del método general de interpretar los comportamientos por medio de atribuir estados mentales desde la actitud intencional. En este caso no se trata de atribuir los estados mentales que les dan sentido a nuestros comportamientos, sino aquellos que *nos parece* que dan lugar a nuestro comportamiento, y en este aspecto, es heredero de la concepción fenomenológica de los estados conscientes Sin embargo, el método no propone que hay que abordar esos relatos como si necesariamente refirieran a cómo se dan efectivamente los eventos en el cerebro. Es una fenomenología, porque sólo refiere a cómo las cosas se nos aparecen, pero es “hetero” porque la propuesta consiste en que los relatos fenomenológicos de un sujeto en particular sean identificados y descriptos por un tercero, de modo tal que a este fenómeno de carácter subjetivo, se le añadiría un componente objetivo, propio de las ciencias.

La teoría de la conciencia elaborada por Dennett pretende brindar una explicación materialista de la misma. Así, uno de sus objetivos es proporcionar una alternativa a las *explicaciones dualistas de la conciencia*, según las cuales la mente no es igual al cerebro, en el sentido de que no está compuesta de sustancia ordinaria, sino de un tipo especial de sustancia. Por el contrario, Dennett asumirá como punto de partida una posición *materialista*, según la cual sólo hay un tipo de materia, la sustancia física de las ciencias naturales (física, química, biología). Dicha perspectiva materialista propone que, en algún sentido, la mente *es* el cerebro. Sin embargo, el filósofo norteamericano también evitará caer en una concepción igual de errónea que

el dualismo, pero materialista, que bautiza como *Materialismo Cartesiano*: “la visión a la que se llega cuando se rechaza el dualismo de Descartes pero no se consigue abandonar esa imagen de un teatro central (aunque material) a donde ‘todo acude’” (Dennett, 1991a, p. 121). Su concepción alternativa queda plasmada en su *Modelo de las Versiones Múltiples*, cuya formulación más completa puede encontrarse en *La Conciencia Explicada* (Dennett, 1991a). En lo que sigue señalaré algunos aspectos relevantes de su crítica al *Materialismo Cartesiano* y de su *Modelo de las Versiones Múltiples*, en función de lo desarrollado en dicho libro.

Dennett señala que la idea de una mente consciente siempre está asociada con la de un punto de vista, como si la misma fuera equivalente a un observador que recoge una cantidad limitada de información. En un sentido práctico, este punto de vista puede ser considerado como un punto que se mueve en el espacio-tiempo. Este es el supuesto que manejamos cuando consideramos a un observador consciente en dimensiones ordinarias. Sin embargo, cuando se intenta localizar ese punto de vista dentro del individuo, los supuestos que funcionan bien a gran escala, empiezan a fallar. Lo cierto es que no existe un solo punto del cerebro al cual acuda toda la información. Este hecho tira por la borda algunas de las intuiciones más básicas que tenemos acerca de cómo se da nuestra percepción, según las cuales hay un momento preciso en el que un determinado fenómeno que observamos pasa a formar parte del flujo de nuestra conciencia.

La postura que adopta Dennett rechazando la existencia de aquel punto central en el cerebro entra en claro conflicto con la teoría de la conciencia de Descartes y su propuesta sobre la glándula pineal como pórtico del alma. Si bien podemos considerar que el dualismo ha sido rebatido, muchos materialistas olvidan que una vez que se abandona la idea de la *res pensante*, ya no hay lugar para ningún centro funcional en el cerebro. El *Materialismo Cartesiano* es, entonces, para nuestro filósofo, la concepción según la cual, de formas más o menos explícitas, muchas teorías materialistas actuales siguen comprometidas con la idea de que hay un teatro central, aunque material, en el cerebro, en donde se “proyecta” el flujo de la conciencia (el *Teatro Cartesiano*). Otra arista de esta suposición es la de que habría un observador dentro del cerebro, que le da sentido al teatro. La idea de fondo que la sustenta es que hay una línea de meta crucial, en alguna parte del cerebro, que señala el lugar preciso en el cual la llegada de un estímulo significa la “presentación” en la

experiencia, el acceso a la consciencia. Por el contrario, Dennett considera que si bien hay un sentido en el que el cerebro puede considerarse el cuartel central, no hay otro cuartel central dentro de él y, por ende, no hay un observador dentro del cerebro.

En base a la argumentación anterior, el filósofo propone abandonar tales hábitos del pensamiento y buscar nuevas imágenes para pensar la consciencia, en función de lo que sabemos acerca del cerebro. El modelo sustituto que nos propone será el de las *Versiones Múltiples*, según el cual todas las variedades de la percepción (y en realidad la actividad mental en general) se llevan a cabo a través de procesamientos paralelos de interpretación y elaboración de los estímulos sensoriales de entrada.

La información que ingresa al sistema nervioso se encuentra sometida a un proceso de edición y corrección. No experimentamos lo que ocurre en nuestros órganos sensoriales, sino el producto de muchos procesos interpretativos: los procesos de compilación. Lo novedoso del *Modelo de Versiones Múltiples* es que sostiene que los procesos de detección de rasgos o de discriminación tan sólo tienen que efectuarse una vez. Esto quiere decir que cuando una porción especializada y localizada en el cerebro ha llevado a cabo la "observación" (interpretación) de un rasgo determinado, el contenido informativo queda fijado y no tiene porqué ser llevado a otra parte para ser "re-discriminado". En términos generales, "(E)l proceso de discriminación no conduce a una representación del rasgo discriminado en beneficio de los espectadores de Teatro Cartesiano, porque no hay ningún Teatro Cartesiano" (Dennett, 1991a, p.127).

Una de las implicancias del *Modelo de las Versiones Múltiples* es que si quisiéramos fijar un instante de procesamiento en el cerebro, como si fuera el paso a la consciencia, la decisión sería arbitraria. Si bien esos procesos de fijación de contenido que están distribuidos espacio-temporalmente en el cerebro podrían localizarse con precisión, los mismos no marcan el comienzo del contenido de la consciencia. Sigue siendo una posibilidad abierta que un contenido fijado de este modo no se haga consciente, y es un error preguntarse cuándo se vuelven conscientes los que lo hacen. Esos contenidos identificados producen algo parecido a un flujo, que de algún modo está sujeto a un proceso continuo de edición distribuido en el cerebro. Dicho flujo se parece a un relato sólo por su multiplicidad, pero no por su linealidad: "en cualquier intervalo de tiempo existen múltiples "versiones" de

fragmentos narrativos en diferentes estadios de edición y en diferentes puntos del cerebro" (Dennett, 1991a, p. 127). Así, este modelo nos evita caer en la idea de que tiene que haber una versión canónica o real del flujo de la conciencia.

El modelo también contradice la idea de que en algún momento se produce una corrección o "relleno" de las percepciones que quedaron con huecos. Para ilustrar esta idea podemos recuperar un ejemplo imaginario brindado por Dennett (1991a), en el cual alguien entra a un habitación cuyas paredes están forradas con un cuadriculado de retratos idénticos de Marilyn Monroe. Como prueban estudios empíricos, el reconocimiento de un retrato no es absolutamente inmediato: es necesario focalizar la mirada en él, observar rasgos y, de este modo, se vuelve posible identificar de forma más o menos precisa la imagen. Sin embargo, no se podría decir que el tiempo que el sujeto tarda en llegar a la conclusión de "esta pared está repleta de retratos idénticos de Marilyn" haya sido suficiente para que realice esta operación con cada uno de los retratos. En realidad, lo hizo con uno o dos, y luego realizó una inferencia. Según lo que indican estudios científicos sobre el funcionamiento de la visión, la misma está lejos de permitir realizar distinciones mucho más precisas que manchas de colores para todo aquello que queda afuera del lugar en donde el ojo hace foco. Así, la información que ingresa por el sistema visual sólo permitiría identificar un par de retratos en los que se hace foco, mientras que de todo el resto sólo podría recuperar manchas borrosas. Sin embargo, lo que el relato heterofenomenológico reportaría es la visión de toda una pared repleta de Marilyns, no una imagen con una Marilyn en el centro rodeada de manchas similares a Marilyn. ¿Cómo sucede esto? "¿Es acaso posible que el cerebro tome una de sus visiones foveales de alta resolución de Marilyn y la reproduzca, como si hiciera fotocopias, por la proyección interna del plano de la pared?" (Dennett, 1991a, p. 365). La respuesta de Dennett es que quizá esta hipótesis es posible, pero es poco plausible: "dudo que el cerebro se tome todas las molestias de llevar a cabo esta repleción. Habiendo identificado una Marilyn, y al no recibir información de que las demás manchas no son Marilyns, llega a la conclusión de que éstas son también Marilyns, y asocia a toda esa región una etiqueta que indica «más Marilyns», sin ocuparse de reproducir todas las Marilyns" (Dennett, 1991a, p. 366). Así, Dennett considera que a partir de esos procesos de "etiquetado" de cierta información, el cerebro va componiendo las imágenes, pero en ningún momento necesita "volver a pintar" el cuadro completo dentro de la cabeza. Según el filósofo, una vez que se ha

producido un análisis, el mismo no tiene por qué volver a presentarse en una recopilación de los resultados obtenidos ¿Para quién se estarían realizando estas correcciones? La idea de la corrección sería subsidiaria de la suposición de que hay algo así como un espectador de un teatro. Según la propuesta que estoy presentando, en el cerebro habría edición, pero no proyección de los resultados de esa edición. En el marco de la concepción del *teatro cartesiano* cabía pensar a la “proyección” como el flujo de conciencia real, y al recuerdo de lo proyectado como una apariencia sobre ese flujo de la conciencia. Desde las Versiones Múltiples, lo que se va memorizando son los actos simultáneos de edición e interpretación que se llevan a cabo en la percepción misma, por lo que no se puede trazar esa distinción entre el flujo de la conciencia real y el aparente. Puestos a reflexionar encontramos que, al mismo tiempo, la distinción en consideración (entre lo real y lo aparente) no parece ser aplicable al campo que estamos tratando: el de la subjetividad. Dennett califica de metafísicamente dudosa a la categoría surgida de aplicar dicha distinción en el ámbito de lo subjetivo; tal cosa nos llevaría a distinguir lo “objetivamente subjetivo”, de aquello que a alguien *sólo* “le parece que le parece” (Dennett, 1991a, p. 146). También sería extraño que alguien pueda hacer una descripción fenomenológica en la que diga que juzga algo de tal forma *como resultado* de que le pareció que era así; dado que, como ya observó Hume, no tenemos acceso a este tipo de hechos sobre la causalidad. En términos dennettianos es lo mismo que afirmar que la *heterofenomenología* otorga autoridad máxima a quien relata acerca de lo que él experimenta, pero no acerca de los procesos internos que causarían esa experiencia.

Los estímulos visuales evocan cadenas paralelas de acontecimientos en el cerebro, que van produciendo distintas discriminaciones o identificaciones ¿En qué lugar se reúne toda esta información? En ninguna parte. Algunos de estos estados portadores de contenido se desvanecen y otros son retomados por alguna de las variedades del conjunto perceptivo (testimonios orales, disponibilidad semántica, estados emocionales, inclinaciones comportamentales, etc.).

En definitiva, hay un “proceso de fijación de contenido” que genera versiones o borradores de fragmentos narrativos. Sin embargo, “(E)sta madeja de contenidos se asemeja a una narración únicamente por su multiplicidad; en cualquier intervalo de tiempo hay múltiples versiones de fragmentos narrativos en varios estadios de revisión, en varios puntos del cerebro” (Dennett, 1991a, p. 149). Sin embargo, a este

proceso se agrega el hecho de que, además, realizamos una especie de selección o identificación de determinados borradores y una integración de los mismos. En palabras de Dennett, esto ocurre porque en determinados intervalos de tiempo se lleva a cabo un “sondeo” de esa madeja de contenidos, que produce una narración distinta, dependiendo de cuál sea el momento en el que se realiza el sondeo. “Y estas son verdaderas narraciones: versiones únicas de una porción del ‘flujo de la conciencia’” (Dennett, 1991a, p. 149). Los distintos momentos en los que decidimos “sondear” esta madeja de contenidos, dan lugar a distintas narraciones. Al mismo tiempo, cada narración se ve sometida a una revisión continua, por lo que no hay una narración canónica (lo objetivamente subjetivo, “El Flujo de la Conciencia”).

Así, el sondeo que se puede realizar con respecto a lo que se ha experimentado es un acto que precipita narraciones. Esto lo puede hacer el mismo sujeto o un observador externo. Tanto los procesos de autorreconocimiento o las autoinvestigaciones -por parte del propio sujeto-, como los actos de habla que expresan interpretaciones de los contenidos de la conciencia -por parte de observadores externos-, fijan interpretaciones sobre lo que experimentó el sujeto. Se trata de juicios sobre lo que le parece al sujeto, juicios que el sujeto puede entender, recordar y en torno a los cuales puede actuar. La ventaja del Modelo de las Versiones Múltiples es que no necesita suponer que, además de tales juicios, haya habido una presentación o proyección de los elementos disponibles.

En resumen, aunque la conciencia se nos presenta como un sistema de procesamiento de información que funciona de manera lineal y está controlada por una entidad única, el cerebro, en realidad es un sistema de procesamiento paralelo que no cuenta con un sistema de control en el que después se reúna toda la información para ser procesada (nuevamente) de forma lineal. Junto con la idea del *teatro cartesiano*, se ha rebatido la del espectador del mismo, que no era más que una de las reencarnaciones de una concepción clásica sobre el “yo”. Según este nuevo modelo, ya no hay un conceptualizador central en el cerebro, no hay un capitán del barco, por así decir. ¿Cómo es que se explica entonces, el comportamiento que efectivamente manifestamos, dado que somos criaturas dotadas cognitivamente de forma tal que no sólo sabemos hacer que “el barco no se hunda”, sino que además hacemos planes a largo plazo, prevemos peligros, corregimos errores y llevamos

adelante grandes proyectos? Esa es una de las preguntas que intenta responder Dennett con su teoría del “yo” narrativo.

2.El problema filosófico del “yo” y la propuesta narrativista de Dennett

El problema filosófico del “yo”

Si bien es cierto que la pregunta acerca de la identidad personal tiene una larga trayectoria en la filosofía occidental, resulta dificultoso hablar de una historia lineal en la formulación y reformulación de la misma; más bien podríamos pensar en un paraguas de problemas que comparten algunos rasgos o preocupaciones en común. Este conjunto de interrogantes con rasgos en común no puede ser comprendido por completo si no se tiene en cuenta que en muchas de sus formulaciones el concepto en cuestión era el de *alma*, que de algún modo es el antecesor religioso del “yo” y con cual comparte varias características. Si bien la concepción secular del mundo, que empezó a ganar cada vez más terreno en el campo de batalla de la filosofía a partir del siglo XVII en adelante, trae problemas más inmediatos a la idea de alma que a la de “yo”, este último concepto hereda algunos de esos problemas, se deshace de otros y cosecha los propios.

Lowe (2000) ha señalado que la pregunta acerca de qué es el “yo” podría confundirse con la pregunta por su función lingüística, que se responde indicando, simplemente, que ‘Yo’ es la palabra (del español, en este caso) que todo el mundo emplea para referirse a sí mismo. Sin embargo, el interés filosófico por el significado de dicha palabra surge, según Lowe, de la dificultad para especificar, de una manera que resulte iluminadora, lo que es ‘referirse a sí mismo’ del modo particular en el que esto se da cuando se usa comprensivamente el término ‘Yo’. La variedad de teorías que se han formulado a este respecto comparten preocupaciones como las siguientes: ¿cómo es que el “yo”, y sus habilidades, perduran o permanecen?, o ¿qué es lo que nos importa que se mantenga en una situación determinada para decir que hemos sobrevivido? Una pregunta vinculada a estas preocupaciones es la de qué tipo de cosa es el “yo” e, incluso, si es una cosa. Otro interrogante, vinculado con el anterior, pero diferente, es el de por qué una persona o un “yo”, es (o no es) la misma persona (o el mismo “yo”) que otra a lo largo del tiempo. Lowe (2000) presenta este último interrogante bajo la siguiente formulación: “¿de qué modo han de estar relacionadas

una con otra la persona P_1 y la persona P_2 para que P_1 sea *idéntica* a P_2 ?” (p. 235). Dicha formulación permitiría entender el interrogante tanto en su aspecto ‘diacrónico’, esto es, haciendo referencia a la persistencia *a través* del tiempo, según el cual sería equivalente a la última pregunta mencionada anteriormente, como en su aspecto ‘sincrónico’, es decir, haciendo referencia a la persistencia *en* el mismo momento de tiempo.

Más adelante señalaremos que se puede realizar una distinción entre las teorías del “yo” y las de la Identidad Personal, problemas que hasta aquí hemos tratado de manera indistinta, aunque la misma no resulta muy relevante para este trabajo. En una primera parte de este apartado intentaré señalar algunos aspectos significativos del origen histórico de las preocupaciones en torno al “yo” y el alma, para, luego, realizar una breve exposición del estado actual de la discusión en torno a los problemas del “yo” y la identidad personal.

¿Es posible determinar un origen histórico de este conjunto de preocupaciones? Las tumbas que dejaron los Neandertales, rodeadas de piedras regulares dispuestas cuidadosamente de manera simétrica, son un indicador de que ya estos habitantes tempranos de nuestro planeta se relacionaban con la muerte de un modo diferente al que lo hacen el resto de los animales, lo que daría cuenta de una preocupación por la perduración o no en la existencia de aquellos cuyos cuerpos han dejado de vivir.¹² Sin embargo, la preocupación por la vida (o por algún modo de existencia de las personas) después de la muerte, puede implicar una diversidad de ideas muy diferentes entre sí, en muchos casos inconmensurables, y que parece muy difícil mapear de modo exhaustivo. En muchas culturas tradicionales es corriente encontrar la idea de una presencia de los ancestros ya fallecidos, que en ocasiones particulares pueden acompañar de distintos modos la existencia de los vivos. Sin embargo, en estos casos no está presente ningún tipo de anhelo por convertirse en ancestros que sobrevivan a la muerte, ni la idea o la preocupación de que a todos los miembros de la comunidad les aguarda ese destino. En cambio, en las culturas occidentales industrializadas empezamos a encontrar un interés personal por la

¹² La mayoría de los antropólogos que han investigado estas tumbas, las consideran un indicador de la creencia en la existencia de la persona después de la muerte del cuerpo. Se puede pensar que estas creencias fueron motivadas por visiones o sueños, en donde es frecuente experimentar el encuentro con alguien que ya murió (Martin y Barresi, 2006, pp. 290, 291).

propia supervivencia después de la muerte del cuerpo. El *Fedón* de Platón es una de las primeras señales de esta inquietud.¹³ Si bien en el caso de Platón hay acuerdo entre muchos intérpretes en que no podemos hablar de una ansiedad por experimentar esa vida después de la muerte, sí hay una preocupación por “practicar la muerte” en vida minimizando las distracciones del cuerpo, esto es, vivir una vida más pura, pero no por cobrar la recompensa después, sino porque eso hará mejor la vida misma. Esta postura presenta algunas diferencias respecto de la idea de que, de no haber un precio que pagar después de la muerte, lo más sensato sería aprovechar la vida en la tierra entregándonos al hedonismo (idea que se puede encontrar en el cristianismo, como se puede ver en San Pablo 1 Corintios 15:17¹⁴). En el cuerpo de creencias del cristianismo sí es posible encontrar, a diferencia del pensamiento griego, una ansiedad por esta elevada vida después de la muerte, que tiene como contrapartida una invitación a habitar el mundo en la tierra según un principio de posponer la buena vida. Si bien estos comentarios son sólo un esbozo de la compleja e intrincada historia de las concepciones occidentales sobre la perduración del “yo”, sí podemos señalar dos puntos acerca de ella. Por una parte, que cada una de estas variadas concepciones tiene fuertes implicancias éticas referentes a cómo vivir. Por otra, que sí es posible identificar un cambio en el foco de atención: mientras que griegos y romanos se preocuparon por el vivir bien en el presente, el cristianismo vira esa preocupación hacia el futuro. Aunque este trabajo no pretende dialogar con la historia de la filosofía occidental, ni mucho menos con concepciones religiosas del “yo”, resulta interesante mencionar estos aspectos del cristianismo porque han tenido profundas consecuencias para las teorías filosóficas del “yo” y de la identidad personal.

¹³ El *Fedón* o, como se le tituló después, “Sobre la inmortalidad del alma”, está ambientado en las últimas horas de vida de Sócrates. En una reunión con la mayoría de sus amigos, Sócrates discute el tema del alma humana, intentando responder a preguntas como cuál es su relación con el cuerpo, y qué pasa con ella después de la muerte, como puede verse en el siguiente pasaje: “En verdad, Simmias y Cebes, si no creyese encontrar en el otro mundo dioses tan buenos y tan sabios y hombres mejores que los que dejo en este, sería un necio, si no me manifestara pesaroso de morir. Pero sabed que espero reunirme allí con hombres justos. Puedo quizá hacerme ilusiones respecto de esto; pero en cuanto a encontrar allí dioses que son muy buenos dueños, yo lo aseguro en cuanto pueden asegurarse cosas de esta naturaleza” (Platón, *Fedón*, 63b6).

¹⁴ En los versículos 12-19, Pablo, en respuesta a algunos que expresaban dudas en la congregación de Corinto, a la que está dirigida la carta, aduce la importancia fundamental de la resurrección como una doctrina cristiana. En su argumento, reprende a la Iglesia de Corinto diciendo que si Jesús no resucitó después de la crucifixión, entonces no tiene sentido la fe del cristianismo.

En su libro *The rise and fall of Soul and Self* Raymond Martin y John Barressi (2006) reconstruyen lo que dan en llamar una historia intelectual de la identidad personal. Su reconstrucción historiográfica propone que en la filosofía occidental recién se podría hablar de los problemas filosóficos del “yo” y de la identidad personal a partir de la Patrística, que comienza en el siglo II d.C. Si bien hay discusiones más tempranas relativas a este tema, que son, de hecho, retomadas por los padres de la iglesia, este es el primer momento en el que se constituye como problema filosófico en el sentido de que quienes lo abordan intentan dialogar con lo que se ha dicho antes. En aquel contexto, la discusión se insertaba en el intento de desarrollar una teología religiosa adecuada para comprender la permanencia del “yo” en una vida después de la muerte; es decir, buscaban un modo de explicar algo que ya daban por cierto. Sin embargo, esto no significa que no podamos recuperar algunos aportes anteriores a este momento. En lo que sigue, identifico algunos hitos importantes en la historia de los problemas en torno al “yo” y la identidad personal, de acuerdo a la reconstrucción propuesta por M & B (2006).

Según los autores, en esta compleja y rica historia intelectual se pueden identificar tres posturas que son recuperadas y reformuladas en distintos momentos, pero que tuvieron su origen en la antigua Grecia. En este periodo podríamos encontrar tres respuestas diferentes que se habrían dado al problema de la identidad personal: la de Platón, la de Aristóteles y la del atomismo materialista. Haciendo un bosquejo más que simplificado podríamos decir que lo distintivo de la propuesta platonista radica en explicar la permanencia y la inmortalidad del “yo” o del alma por su pertenencia a un reino de cosas inmutables. En cambio, en Aristóteles habría una dimensión inmutable en cada objeto material, inclusive en el cuerpo humano, que permite la persistencia del “yo” durante la vida (aunque no necesariamente la inmortalidad). Por último, los atomistas materialistas se enfrentan con un mundo en el que los objetos materiales y sus cambios son el resultado de asociaciones y separaciones entre los átomos: pequeñas unidades materiales e inmutables. Algunas configuraciones medianamente duraderas de los átomos se vuelven identificables y, por ende, susceptibles de ser conocidas. Las personas serían, según la tercera visión, configuraciones materiales de este tipo.

La idea de alma inmaterial, tomada en la Alta Edad Media del pensamiento platonista, es retomada por el cristianismo y todo el pensamiento teológico de

nuestro pasado no tan lejano, como un postulado teórico respetable, que sólo empieza a tener fuertes problemas por la gran contradicción entre el dualismo ontológico de fondo y las nuevas concepciones físicas que iban trayendo los desarrollos científicos.¹⁵

El punto de inflexión en el que podemos ubicar el vaticinio de este cambio de concepción es la emergencia del nuevo abordaje científico que se da en el siglo XVII, que acarrea una nueva teoría física del mundo: el mecanicismo corpuscular. El surgimiento de la ciencia moderna ocasionó un movimiento del terreno que no sólo hizo caer el prestigio de determinadas posturas, sino que además cambió el marco de la pregunta. Si bien lo que nos interesa aquí no es la idea de alma inmaterial, sino lo que puede considerarse su descendiente laico, el “yo”, el cambio de concepción acarreado con el advenimiento de la ciencia moderna es un momento de mucha relevancia para el problema filosófico en cuestión. Dicho cambio no solamente afectó a la noción antecesora, de la cual deben permanecer muchos aspectos -más o menos ocultos- en la heredera, sino que además puso en problemas la noción misma de “yo”, al menos en dos aspectos fundamentales: el estatus del “yo” como una entidad real (más que como una ficción conveniente) y la unidad del “yo” y su capacidad de explicar otras unidades, como la alegada unidad de la conciencia. El “yo” ya no es un concepto al que se puede apelar para explicar otros fenómenos, sino que es más bien algo que debe ser explicado. La nueva cosmovisión materialista no sólo puso en problemas la perdurabilidad del “yo” después de la muerte del cuerpo, sino que dejó de dar por supuesta la perdurabilidad en el día tras día, momento tras momento, (que antes era garantizada porque más allá de los cambios físicos y psicológicos, lo que persistía era el alma inmaterial). Lo que estamos intentando decir no es que se comenzó a poner en tela de juicio las creencias *folk* en su uso cotidiano, ni que cambió el modo en que nos comportamos o nos representamos nuestra identidad y la

¹⁵ Por otra parte, resulta interesante señalar que Descartes es el primero en usar la palabra en latín para mente [*mens*] como sinónimo de alma [*anima*] (ver Burke, 1997). En sus reflexiones, se entiende que cada uno de nosotros es una mente, no extensa, y que la esencia de la mente es la actividad de pensar. Desde esta perspectiva, ni los animales ni las plantas tendrían un alma, la que, por otra parte, tampoco es entendida como un requisito para la vida. Como intentamos sostener, en la actualidad nuestra intuición es heredera de esta conexión entre la identidad personal y la mente, que se ve plasmada en la filosofía de Descartes; lo que constituye otro motivo por el cual desde perspectivas ontológicamente deflacionadas de lo mental, como la idea de la intencionalidad derivada de Dennett, resulta particularmente desafiante plantear el problema de la identidad personal.

de otros humanos en el día tras día, sino que ella atenta contra la noción de “yo” como una noción científicamente explicativa y/o teóricamente justificada.¹⁶

En el contexto actual, en el que la ciencia física moderna constituye el telón de fondo contra el cual debe desarrollarse una explicación del “yo”, sería de esperar que, dentro de las tres opciones presentadas por M & B, sean distintas versiones de materialismo atomista las que han tenido más peso e importancia. Sin embargo, sin intentar un repaso exhaustivo de esas versiones, podemos adelantar que los pocos conocimientos que tenemos sobre el ordenamiento de la materia de nuestros cuerpos dan por tierra con la posibilidad de pensar que podemos asentar la identidad personal en el ensamblaje de determinados átomos. Sabemos, por ejemplo, que el intercambio energético de nuestro cuerpo con el medio implica una incorporación y pérdida de átomos constante, y que la muerte y nacimiento de células durante nuestra vida es tal que, por ejemplo, una niña de ocho años tiene la mayoría de células diferentes a las que tenía con respecto a cuando tenía cinco años. En consecuencia, la explicación de la permanencia del “yo” no puede residir en la permanencia de un determinado conjunto de materia.

Sin embargo, la última objeción del párrafo anterior no basta para rebatir otras formas de identificación del “yo” con el cuerpo. Para dar cuenta de ello, retomaré una distinción más reciente entre teorías partidarias de criterios *físicos* y teorías partidarias de criterios *psicológicos* para pensar la persistencia de la identidad personal. Schechtman (1996), por ejemplo, ubica a Williams en el primer grupo y a Shoemaker, Lewis, Parfit y Perry, en el segundo. Con la misma intuición, Lowe (2000) distingue entre criterios de continuidad *corporales* y *psicológicos*. Según la descripción de Lowe, el criterio corporal nos permite sostener que la persona P_1 es la misma que la persona P_2 (en distintos o en el mismo momento) si tienen el mismo cuerpo. Como dijimos, esta manera más amplia de formular el criterio permitiría responder a la objeción anterior. Ante la constatación de que una persona sobrevive a la renovación de átomos, e incluso, podríamos agregar, a la pérdida de ciertas partes de su cuerpo, el teórico del criterio corporal responder que el cuerpo mismo también lo hace.

¹⁶ Con esta aclaración no se pretende negar que las discusiones filosóficas, religiosas y científicas, muchas veces modifican de manera procesual las representaciones culturales y las creencias *folk*. Aún así, es posible realizar la distinción y entender que en ciertas circunstancias unas se verán amenazadas de manera directa mientras que las otras no.

Ahora bien, es posible ofrecer otras objeciones en contra de esta posición, mostrando ejemplos en los que parece que hay más de una persona en el mismo cuerpo, como ocurre en los casos de Trastorno de Personalidad Múltiple¹⁷, o señalando que una persona podría cambiar de cuerpo con otra. En palabras de Lowe: “Esta última posibilidad la sugiere el procedimiento, al parecer factible pero hasta ahora no realizado, de trasplante total del cerebro, en la cual el cerebro de una persona se trasplanta al cráneo de otra y a la inversa” (Lowe, 2000, p. 236). Frente a esto, un teórico del criterio corporal podría decir que la continuidad personal no está garantizada por el cuerpo entero, sino sólo por el cerebro.¹⁸ Sin embargo, nuestras intuiciones parecen apuntar hacia un lado bastante diferente, por lo que Lowe dice que quien realiza tal reformulación en realidad está señalando que lo importante son los recuerdos y el comportamiento (criterio ‘psicológico’), no el órgano cerebral. Es fácil advertir que lo desesperante de *La metamorfosis* (Kafka, 1915) es que cuando Samsa se convierte en el desagradable insecto, sigue siendo Samsa. En películas como *Freaky Friday* (Waters M, 2003), cuando madre e hija se despiertan con sus cuerpos intercambiados, el aprendizaje que adquieren de vivir la vida que vive el cuerpo de la otra lo adquiere cada una de ellas en su “versión original”, que continúa siendo la misma persona sólo que con nuevas experiencias. Más aún, si hubieran permanecido con el cuerpo cambiado, hubiéramos considerado que simplemente eran las mismas personas condenadas a vivir en nuevos cuerpos. Si nuestras intuiciones señalaran que la identidad personal radica en que conservamos nuestros cuerpos, o alguna parte de ellos, ninguna de estas historias tendría el efecto que tiene.

Lo desarrollado anteriormente parece conducirnos a la posición que Lowe contrapuso al criterio ‘corporal’ de la permanencia de la identidad personal: esto es, el criterio ‘psicológico’. Según éste, la persona P_1 es idéntica a P_2 si, y sólo si, comparten ciertos aspectos vitales de la personalidad humana. Es menester aclarar que esta explicación tampoco está exenta de importantes críticas. Según Lowe, la

¹⁷ Hoy conocido como Trastorno de Identidad Disociativo, es descrito en el DSM IV (APA, 2000) como el trastorno consistente en la existencia de dos o más identidades o personalidades en un individuo, cada una con su propio patrón de percepciones y acciones en relación con el ambiente. A pesar de la reciente modificación del término, en el resto del trabajo mantendremos el nombre de “Trastorno de Personalidad Múltiple” que es el que emplea Dennett cuando se aboca a su estudio (Dennett, 1989, 1991, Humphrey y Dennett, 1989). El reconocimiento y caracterización del trastorno sigue siendo objeto de fuertes controversias.

¹⁸ Ver también “¿Donde estoy?” de Dennett, en Hofstadter y Dennett (1985)

más difícil de rebatir es la de que podría haber dos cerebros distintos con los *mismos* aspectos de la personalidad humana. Esta objeción se puede formular proponiendo un experimento mental acerca de una *fisión*¹⁹ personal. Suponiendo que todos los aspectos de la personalidad humana están completamente asentados en cada uno de los dos hemisferios del cerebro, si se los separara y se los trasplantara a dos cráneos de cuerpos humanos distintos a los que previamente se les ha extraído el cerebro, pareciera que nos veríamos confrontados con dos personas diferentes. Sin embargo, estas personas tendrían las mismas características psicológicas, por lo que según el criterio en cuestión deberíamos decir que no son dos personas, sino que son una y la misma, lo que resulta profundamente anti-intuitivo. Si bien el criterio psicológico no termina de resultar satisfactorio, parece acercarse un poco más que el criterio corporal en la identificación de los rasgos que nos resultan relevantes para atribuir persistencia personal. Así, no es necesario comprometerse filosóficamente con la existencia de la *res cogitans* como una sustancia separada para que el criterio corporal resulte claramente insatisfactorio; más bien parece que nuestra idea de “yo” está asociada necesariamente a la de conciencia o a la de mente.

M & B (2006) también señalan la importancia de los aportes de cierta perspectiva filosófica “social” al problema. En momentos tempranos, la idea de que el “yo” pudiera identificarse con el organismo parecía un problema para el dualismo, mientras que actualmente las teorías seculares casi nunca son dualistas. Aunque no está muy claro qué debemos entender por materialismo, parece haber acuerdo en que debemos adherir a alguna variante del mismo. Sin embargo, se suele señalar un tipo de razón diferente por la cual el “yo” sigue sin poder identificarse con el organismo y, por ende, sigue sin quedar unificado por el organismo; el problema ya no sería que el “yo” es inmaterial, sino que es social (por lo que la unidad biológica ya no es suficiente para brindar una explicación). Estos aportes son realizados por filósofos como Hegel, James, Lacan, Foucault, Derrida, y los constructivistas sociales. Si bien Dennett no dialoga directamente con todos ellos, podemos decir que en alguna medida hace justicia a muchas de sus observaciones²⁰.

¹⁹ El término “fisión” proviene de la química, donde se le llama “fisión nuclear” a la reacción en la cual un átomo es dividido en una o más partes.

²⁰ De hecho, Carr (1998) señala, desde una perspectiva crítica, que los aspectos valorables de la propuesta narrativista de Dennett no implican innovaciones con respecto a aportes de Sartre o de Husserl.

Considero que la propuesta de Dennett no solamente permite incorporar ciertos aspectos relevantes que los partidarios del criterio psicológico habrían reivindicado al rechazar el criterio físico, como los recuerdos y el temperamento personales, sino también presentar una teoría que incorpora, a su vez, algunos aspectos sociales. En la teoría de Dennett estos están vinculados, como intentaré desarrollar más adelante, con la pertenencia a una cultura y a un lenguaje natural y, más específicamente, al rol que las narraciones tienen en la conformación de la identidad.

Así, desde una perspectiva naturalista, mostrando la posibilidad de una interacción fructífera con los estudios evolucionistas, las ciencias de la conciencia y la psicología cognitiva, Dennett ofrece una explicación del “yo” que permite responder a la pregunta ontológica sin divorciarse de la cosmovisión que nos brinda la física contemporánea, y dar cuenta, al mismo tiempo, de esta dimensión social que pondría en tela de juicio una concepción sustancialista del fenómeno en cuestión que, en determinados momentos, parece irrenunciable, aunque incompatible con otros aspectos del mismo.

Analíticamente, se podría distinguir entre las teorías del “yo” y las teorías de la Identidad personal. Las primeras se encargan de decir qué tipo de cosa es el “yo” e, incluso, si es una cosa. Las segundas, por otra parte, son teorías de la identidad a lo largo del tiempo, esto es, teorías que explican por qué una persona o un “yo” es (o no es) la misma persona (o el mismo “yo”) que otra a lo largo del tiempo. Ahora bien, la distinción no tiene mucha importancia para la investigación realizada en este trabajo, en particular por el hecho de que en el pensamiento de Daniel Dennett las respuestas a ambas problemáticas se ven profundamente vinculadas. Si bien el modo de abordar la problemática está articulado en torno a una pregunta ontológica, que apunta quizá de un modo más directo a la primera de ellas, entiendo que su respuesta implica a su vez una respuesta a la segunda.

El “yo” según Dennett

Uno de los motivos por los cuales consideramos no problemático abstenernos de realizar una reconstrucción de la larga e intrincada historia del problema filosófico del “yo” es porque el mismo Dennett construye su teoría sin elaborar un diálogo explícito con la tradición con respecto a este problema. Su preocupación

principal podría reconstruirse, según la opinión de Menary (2008), como un intento de brindar una explicación del “yo” que no precise la postulación de una sustancia separada (cartesianismo), ni su identificación con el cuerpo o el cerebro (que sería una especie de reduccionismo fisicalista o, como le llamamos en el apartado anterior, de criterio ‘corporal’). La ciencia nos ha enseñado que el alma no existe, “(P)ero no todas las ideas de que somos algo diferenciado en nuestro cuerpo puramente físico son tan vulnerables al ridículo y la refutación. Algunas versiones florecen de hecho en el jardín mismo de la ciencia” (Dennett, Hofstadter, 1981, p. 13).

En conjunto con Hofstadter, Dennett (1981) elabora un primer libro - traducido al castellano con el título *El ojo de la mente*- dedicado a retomar preguntas e intuiciones de nuestro sentido común y nuestros planteos filosóficos más bien espontáneos, a partir de distintos textos, muchos de los cuales son de carácter literario.²¹ De este modo, se vislumbran algunas pinceladas que empiezan a marcar el carácter del problema para Dennett, pero en ese libro todavía no ensaya ninguna respuesta concreta. Si bien en sus primeros textos ya encontramos algunas preocupaciones vinculadas con el problema en cuestión, como cuáles son las condiciones necesarias para considerar a alguien una persona (1973, 1976) o cómo entender el llamado “Trastorno de Personalidad Múltiple” (Dennett y Humphrey, 1989), recién encontramos una exposición sistemática de sus ideas sobre el “yo” en un artículo de 1988, titulado “Why Everyone is a Novelist”, en cuya oración inicial plantea el problema: “¿Qué es un yo?” (Dennett, 1988, p. 16, mi traducción).²² Esta pregunta encuentra como respuesta una primera formulación de la teoría dennettiana del “yo” como “centro de gravedad narrativa”. Al año siguiente, es publicado “The Origins of Selves”, dedicado a intentar mostrar al “yo” como algo que se origina a través de un proceso, como todos los fenómenos que tienen lugar en los organismos biológicos. Aunque es inevitable ir tejiendo el recorrido de la aparición y reformulación de las ideas de Dennett en torno al “yo” a partir de una exploración de estos textos -que constituyen sólo un pequeño porcentaje de su vastísima bibliografía-, resulta ocioso para nuestros fines reconstruir y explicitar lo anecdótico

²¹ Entre los autores de los textos recopilados encontramos escritores, filósofos y científicos; como Jorge Luis Borges, Justin Leiber, John Searle, Thomas Nagel, Richard Dawkins, Harold Morowitz, y los mismos Dennett y Hofstadter, entre otros.

²² Una versión posterior de este artículo es publicada, con escasas modificaciones, bajo el nombre de “The Self as the Center of Narrative Gravity”, en 1992. En general, citaré la versión de 1988, a menos de que resulte más pertinente referirse a la de 1992.

de las recurrencias y transformaciones en el desarrollo de estas ideas, dado que es evidente que la más completa y desarrollada versión de las mismas se puede encontrar en el capítulo destinado a “La realidad de los Yos” en *La conciencia explicada* (1991a).²³

En el libro en cuestión, Dennett sostiene que:

Un “yo”, de acuerdo con mi teoría, no es un viejo punto matemático, sino una abstracción que se define por la multitud de atribuciones e interpretaciones (incluidas las autoatribuciones y las autorepresentaciones) que han compuesto la biografía del cuerpo viviente de la cual es su centro de gravedad narrativa. (Dennett, 1991a, p. 437).

Para explicar la metáfora conceptual del “centro de gravedad narrativo”, Dennett (1988) realiza dos analogías, una con el centro de gravedad de un objeto, y otra con el personaje de una ficción, que serán desarrolladas en el capítulo siguiente. Sin embargo, podemos adelantar que dicha metáfora hace referencia a que el “yo” es aquello en torno a lo cual giran todas las narraciones que tratan sobre nosotros. Dichas narraciones son producidas, en gran medida, por nosotros mismos. Desde chicos, aprendemos a relatar nuestra vida, incluyéndonos a nosotros mismos en dichas narraciones, como personajes fundamentales de nuestra historia.

Sin embargo, Dennett aclara que nuestro “yo” también está conformado por las atribuciones de las narraciones de otros. Si bien no lo especifica, podemos suponer que no se trata de las narraciones sobre nosotros que produzca *cualquier* otro. En primer lugar, aquellas que alguien nunca llega a conocer no pueden formar parte del *corpus* de narraciones que constituyen su “yo”. Incluso si las mismas llegaran a ejercer efectos sobre la vida de esa persona (por ejemplo, que alguien sea contratado en un trabajo por buenas recomendaciones), la narración de la que no toma conocimiento sólo modifica su “yo” de manera indirecta y dichas modificaciones son incorporadas a su “yo” a partir de las nuevas narraciones que surjan a partir de los hechos de los que sí tiene conocimiento (la adquisición del nuevo empleo, la eventual noticia de que alguien lo ha recomendado, etc.) Por otra parte, habrá muchas narraciones de terceros de las que la persona toma conocimiento y, aún así, no acepta como legítimas. Por ejemplo, la que produce una

²³ Esta propuesta de lectura es sostenida por Carr (1998), Corcuera (2013) y Sánchez (2014).

persona adulta diciéndole a su hija adolescente que lo único que ha hecho durante al año ha sido perder el tiempo, o la de una profesora diciéndole a un alumno que sus desempeños escolares demuestran que lo suyo son las matemáticas. Incluso si muchas de éstas pueden afectar a la persona sobre la que tratan, pienso que de allí no se sigue que definan su “yo”, dado que ella puede no aceptarlas. Ante casos como estos, las personas podrían reconvertir sus narraciones, hacer más hincapié en otros hechos, resignificar cosas como “perder el tiempo”, y revalorizar otras habilidades personales, concluyendo, por ejemplo, que su desempeño escolar muestra que lo mejor que puede hacer es dedicarse a las artes plásticas. En definitiva, podemos decir que las narraciones que constituyen ese *corpus* del cual nuestro “yo” es el centro, son aquellas que hablan sobre nosotros y que pueden ser: o bien elaboradas por otros, en la medida en que nosotros las aceptemos y nos las apropiamos, o bien, elaboradas por nosotros (en algunos casos nutriéndonos de cosas que los otros han dicho de nosotros a las que aceptamos o a las que nos oponemos).

Una de las dificultades para interpretar su propuesta es que Dennett no ha realizado ninguna referencia respecto de cuál es la noción de narración que está empleando. Menary (2008) sostiene que una narración, en el sentido más básico, requiere sólo una secuencia de eventos que estén relacionados de algún modo. Lamarque (2004) señala que en las narraciones se representan, al menos dos, eventos, guardando algún tipo de relación entre sí. Desde su perspectiva, en las narraciones hay, crucialmente, una dimensión temporal. Es conveniente señalar que dicha representación es de carácter lingüístico.²⁴ Podemos tomar a estas características generales como una base segura para pensar lo narrativo, evitando atribuirle más de lo que demanda la concepción de Dennett. En un contexto temático diferente (una discusión sobre qué modelo de la conciencia permitiría explicar algunos experimentos visuales) Dennett señala que es importante tener en cuenta que muchas veces la ordenación temporal del “representante” [*representing*] puede no ser igual a la ordenación temporal de lo que se está representando. Según él, esta distinción nos permitiría trazar una categoría metafísica fundamental: “cuando una porción del mundo pasa, de acuerdo con este método, a componer un ovillo de

²⁴ Podría pensarse que hay narraciones que no cuentan con elementos lingüísticos en un sentido estricto, como, por ejemplo, la sucesión de hechos articulada en una tira cómica que no posee diálogos. Dejaremos de lado esa posibilidad en este trabajo dado que, como explicaremos en capítulos ulteriores, la conformación de las narraciones en la teoría de Dennett es específicamente lingüística.

narraciones, dicha porción del mundo es un observador” (Dennett, 1991a, p. 150). Así, podríamos pensar que, para Dennett, una de las principales características de la construcción narrativa es la de poder realizar un tejido propio que vincula algunos de los eventos que suceden en el mundo (dejando afuera a la gran mayoría de ellos) y cuya reconstrucción conlleva una relativa independencia respecto de la sucesión temporal efectiva de los mismos.

Muchos autores han remarcado el rol central de las construcciones narrativas para las relaciones intersubjetivas (por ejemplo, Hutto, 2007, 2008), señalando que la construcción de relatos sobre eventos y acciones, propios y de los demás, sería una práctica esencial para poder entendernos los unos a los otros. En el capítulo anterior señalé que Dennett mismo recupera una parte de nuestra psicología *folk*, la estrategia intencional, que consiste en atribuir creencias y deseos a las otras personas para poder predecir su comportamiento en función de los que pensamos que son sus objetivos. Dichas atribuciones pueden ser pensadas como pequeños relatos que, seguramente, forman parte de las narraciones que construimos. Sin embargo, lo que los otros van a hacer conforma sólo una pequeña porción de aquello que nos interesa de los demás en las relaciones intersubjetivas, por lo que no parece adecuado pensar que la totalidad de las narraciones que constituyen el “yo” está compuesta exclusivamente del tipo de relatos que construimos desde la actitud intencional.

En un texto temprano, Dennett (1973), propuso distinguir la actitud intencional de la *actitud personal*: mientras que las actitudes física, de diseño e intencional no descansan las unas en las otras (en el sentido de que puedo realizar predicciones posicionándome en cualquiera de ellas sin haberlo hecho antes en otra), la actitud personal requiere haber adoptado previamente la intencional y le agrega a la misma la atribución de responsabilidad moral. A diferencia de la intencional, la misma no podría aplicarse a animales ni a robots. Luego, Dennett (1976) precisará que las dos condiciones adicionales para tratar a alguien como persona, y no meramente como un sistema intencional, son: (i) que sea capaz de reciprocidad, es decir que sea capaz de tratar a otros como personas y (ii) que sea capaz de comunicación verbal. Así, los relatos que construimos en nuestras relaciones intersubjetivas no sólo se constriñen a la predicción de comportamientos, sino que además incluyen una dimensión moral. Sin embargo, considero apropiado entender que las narraciones egoicas que constituyen el “yo” dennettiano no se limitan a los

relatos que se pueden formular desde la actitud intencional y desde la actitud personal, sino que cuentan con una riqueza adicional. Dicha riqueza estaría dada por añadidos históricos, afectivos, estéticos, culturales, geográficos, que conforman un “yo” como algo que tiene un punto de vista particular y diferente al de los otros. Es posible pensar que ese *corpus* narrativo no sólo relata acciones, sino que da forma y coherencia a diferentes secuencias de nuestras vidas, proveyendo un marco explicativo para comprender nuestro modo de movernos en el mundo. Así, con respecto a estos conjuntos de narraciones no sólo evaluamos la racionalidad, o la aptitud moral, sino que nos relacionamos con ellos en virtud de que, además, despiertan diferentes sentimientos estéticos, producen mayor o menor interés, despiertan distintos grados de empatía, etc.

Así pues, para Dennett, cada uno se vuelve un “yo” a medida que va construyendo, en conversaciones o escritos, con otros y en soledad, aquello que constituye su autobiografía. El “yo” no es simplemente una cosa que se modifica al cambiar sus propiedades y experiencias; en ese punto no se diferenciaría de los objetos en general. Su característica distintiva es que se modifica por los pensamientos y experiencias, y en virtud del significado que se les otorga; y el modo en que esto se hace es a partir de contar historias, por lo que tal característica (la de modificarse en virtud del sentido que se les otorga) es exclusiva de los fenómenos narrativos como tales. Con esto empezamos a ver que el “yo” es más bien un artefacto: el producto de un cierto tipo específico de actividad humana.

Por otra parte, si bien es cierto que en la constitución del “yo” también juegan un lugar importante las narraciones de otros, una gran cantidad de las narraciones sobre nuestra vida son creadas por nosotros mismos. Es por esto que una de las características propias de la propuesta narrativista, de las que Dennett ha sido uno de los precursores, es que en alguna medida nosotros somos nuestros propios autores, somos Frankenstein, el monstruo, y Frankenstein, el científico, al mismo tiempo:

Así pues, construimos una historia que nos define a nosotros mismos, organizada alrededor de una especie de punto luminoso de autorrepresentación. El punto no es un Yo, por supuesto; es la representación de un Yo ... Lo que hace que un punto sea un punto-yo... no es la forma en que se aparece, sino aquello para lo que se lo utiliza. (Dennett, 1991a, p. 439)

¿Y qué es aquello que representa? Es nada más ni nada menos que un centro de gravedad narrativo. El próximo capítulo estará destinado a examinar el significado de esta expresión metafórica.

3. Centros de gravedad narrativos: ¿De qué están hechos?

Hay un mundo sólido y compacto alrededor nuestro, relleno de cosas menos sólidas, pero no menos reales; este mundo nos es dado desde los días tempranos de la infancia, y no hay dudas de su existencia (Reichenbach, 1938, p. 91, mi traducción).

En la perspectiva de Dennett, el “yo” es, *nada más ni nada menos* que el “centro de gravedad” de las narraciones que nosotros y los otros contamos sobre nosotros mismos. Uno de los mayores obstáculos para valorar esta propuesta, radica en comprender qué tipo de cosa es, para Dennett, el “yo”, y si debemos entenderlo como algo similar a otras cosas, o como a algo que en realidad no existe. Dado que un centro de gravedad es una abstracción, creo que el mejor modo de realizar una reconstrucción interpretativa de su noción de “yo” es apelando a la distinción elaborada por Reichenbach (1938) entre tres clases de entidades: *Concreta*, *Abstracta* e *Illata*. Pretendo defender que los centros de gravedad narrativos pueden ser entendidos como *Abstracta* en el sentido de Reichenbach, apoyándome en que: i) Dennett emplea la distinción de Reichenbach en otras partes de su obra (Dennett, 1987, 1991b); ii) que traza explícitamente la analogía del “yo” con los centros de gravedad que son, justamente, uno de sus ejemplos de *Abstracta*; iii) que este modo de entender al “yo” permite dar sentido y entender mejor ciertos aspectos de la teoría narrativista dennettiana.

En la primera parte de este capítulo, intentaré reconstruir la distinción trazada por Reichenbach entre estas tres clases de entidades en *Experiencia y predicción: Un análisis de los fundamentos y la estructura del conocimiento* (1938). Por otra parte, si bien Dennett no explicita que hay que interpretar a los centros de gravedad narrativos como *Abstracta* reichenbachianos, sí lo hace en el caso de los estados mentales. Por ende, en una segunda instancia, intentaré dilucidar en qué consiste que los estados intencionales sean entidades abstractas en este sentido (y, a partir de esto, cómo entender la distinción entre las clases de psicología intencional que plantea Dennett). Luego, intentaré reconstruir qué serían los centros de gravedad narrativos a partir de lo elaborado en las dos instancias anteriores e incorporando la

analogía que establece Dennett entre estos y los personajes ficticios literarios. Finalmente, defenderé que la distinción de Reichenbach permite entender mejor la distinción propuesta por Dennett entre “tres clases de psicología intencional”: psicología *folk*, teoría de los sistemas intencionales y psicología cognitiva sub-personal; en función de lo cual arriesgaré una hipótesis sobre a cuál de estos tipos de psicología intencional pertenece el concepto de ‘Yo’ que Dennett intenta explicar.

Clases de cosas

En primer lugar, en la propuesta ontológica que Reichenbach elabora en *Experiencia y predicción: Un análisis de los fundamentos y la estructura del conocimiento* (1938) los *Concreta* son los ladrillos que constituyen el mundo. Desde esta perspectiva, un *Concreta* es, por ejemplo, una mesa; no la impresión causada por mi percepción de la mesa, sino la mesa como objeto físico en el mundo. Para este filósofo, las impresiones que tenemos del mundo no son directamente observadas: más bien, son inferencias que realizamos a partir de nuestra experiencia de las cosas concretas con las que tenemos trato (o experiencia) cotidiano. Se asume, desde una posición claramente realista, que son las cosas físicas las que causan las impresiones. Ahora bien, ¿qué lugar ocupan, en este mundo de cosas sólidas, las Naciones, los goles de un partido de fútbol, los centros de gravedad? Estas entidades abstractas son inferidas a partir de los *Concreta*. En la discusión sobre cómo entender la existencia de los *Abstracta*, Reichenbach recupera los argumentos de nominalistas y realistas en torno a la existencia de los universales. Mientras que los nominalistas aciertan en la idea de que las abstracciones pueden ser reducidas a cosas concretas, se equivocan al pensar que de ello se sigue la negación de la existencia de las primeras. En términos de la lógica moderna, se puede decir que las afirmaciones que contienen términos abstractos pueden ser traducidas a afirmaciones que sólo contengan términos concretos y mantener el mismo valor de verdad.²⁵ En algún sentido, la pregunta por si se puede aplicar la categoría de “existente” para términos lógicos es una cuestión de convención, por eso la pregunta es, para el filósofo alemán, acerca de un pseudo-problema. Por otra parte, si bien la postura realista tiene algo de razón al entender que los términos abstractos pueden ser entendidos como existentes, se equivoca al ubicarlos en una esfera especial. Este error sería el resultante de una

²⁵ Es a partir de esto que Reichenbach señala que esta teoría está conectada con una teoría verificacionista del significado. (Cfr. Reichenbach, 1938, cap: La existencia de Abstracta)

mala comprensión de algo que es propio de la relación entre *Abstracta* y *Concreta*: no se pueden “sumar” unos con otros, no se los puede tratar como cosas del mismo nivel; por ejemplo, no puedo estar enumerando las coordenadas del espacio que son ocupadas por las partículas que ocupa un cuerpo y agregar a ellas las ocupadas por su centro de gravedad. Para Reichenbach, esta distinción entre dimensiones es simplemente una cuestión de reglas del lenguaje, basada en la existencia de términos que refieren a entidades de distinto tipo, pero que habitan el mismo mundo, de donde no se debería suponer un salto ontológico que nos lleve a postular la existencia de otra esfera de lo real. Dado que -como dijimos- los *Abstracta* pueden ser “reducidos” a la existencia de *Concreta*, se los puede concebir como compuestos de *Concreta*. Reichenbach propone entenderlos como compuestos “reducibles” [*reducible complexe*] señalando que, si bien no hay una definición precisa de cómo debe entenderse la relación entre el todo y las partes en tales compuestos, se puede decir que se trata de una relación de reducción, definida por la equivalencia.

A partir de la existencia de los *concreta* es posible inferir la existencia de otras entidades que no son abstracciones, pero tampoco son cosas concretas, en la medida en que no tenemos experiencia inmediata de su existencia. Se trata de los *Illata*, que abarcan fenómenos como la electricidad, las ondas de radio, los átomos. A diferencia de los *Abstracta*, estas entidades no mantienen una relación de reductibilidad con los *Concreta*, sino que entre ellos media una inferencia probabilística: Nuestras observaciones sobre las cosas concretas confieren sólo cierta probabilidad a la existencia de los *Illata*. Serían también un tipo de entidad compleja, pero en lugar de ser un compuesto “reducible” [*reducible complexe*], es un compuesto “proyectable” [*projective complexe*].

Como se puede ver, el criterio en el que se basa esta clasificación no se monta sobre un análisis de cómo se dan las relaciones físicas entre las cosas del mundo, sino en función del modo en el que nosotros las conocemos. Así, ya que la distinción está dada por nuestra situación particular y contingente en el mundo, no debe ser entendida como una distinción *a priori* o necesaria. Además, estos límites entre distintos tipos de entidades no están profundamente marcados y pueden cambiar en virtud de futuros desarrollos tecnológicos, procesos psicológicos, etc.

En función de lo expuesto, podemos señalar un aspecto por el cual la clasificación de Reichenbach puede resultar particularmente atractiva para Dennett.

En la introducción de *El Ojo de la Mente* (Hofstadter, Dennett, 1981), los autores mencionan algunas cosas de las que es difícil hablar o decir que existan del mismo modo que existen los objetos físicos. Se trata de cosas como voces, cortes de pelo, agujeros, sinfonías, marchas patrióticas, lenguas muertas, juegos. Lo llamativo de ellas es que, por una parte, no se pueden identificar con un objeto físico con masa o composición química, pero, por otra parte, tampoco son puramente abstractos (como el símbolo π). Antes bien, tienen un lugar de nacimiento y una historia; pueden cambiar y pueden moverse. La definición de *Abstracta* de Reichenbach, permite explicar estas características, ya que los mismos pasan a ser entendidos como algo que depende de los *Concreta*; como una forma de abreviar o simplificar relaciones que se dan entre ellos.²⁶

Los objetos intencionales como *Abstracta*

Las referencias que encontramos acerca de los *Abstracta* por parte de Dennett están casi en su totalidad dedicadas a explicar su concepción sobre las creencias y los deseos (Dennett, 1987, 1991b), y sólo de manera genérica para hablar del “yo” (Dennett 1989, 1992). Por ende, nuestro segundo paso será examinar qué significa para el filósofo afirmar que los objetos intencionales sean abstracciones.

Para establecer qué tipo de entidad serían las creencias y deseos, Dennett evalúa qué tipo de entidad *deberían ser* en función del rol que cumplen en el sistema de atribuciones de la psicología *folk*. Es menester aclarar que esto no resulta equivalente a evaluar cómo serían efectivamente clasificadas por la psicología *folk*, cuestión que será explicitada más adelante. Entender las creencias y deseos atribuidos por la psicología popular como *Illata* implicaría concebir que se trata de entidades postuladas, frente a las cuales nuestra observación de cosas concretas sólo confiere cierta probabilidad a su existencia. Por ejemplo, sería pensar que cuando alguien atribuye creencias y deseos, está postulando en aquella persona a a quien se los atribuye, la existencia de “estados destacados interpuestos de un sistema causante de conducta interno” (Dennett, 1987, p. 58). Dennett considera que ese no es el modo correcto de entender los estados intencionales en función del rol que desempeñan en el sistema de atribuciones de la psicología *folk*. Como explicamos en el capítulo 1, una creencia es la entidad que resulta fructífero postular para realizar una interpretación

²⁶ Dejamos de lado deliberadamente la discusión sobre el número π , que depende de cómo se entienda el estatuto de las herramientas conceptuales matemáticas.

predictiva adecuada del comportamiento de un agente. Por ende, considera que dicho rol es más parecido al que cumple un concepto como el de centro de gravedad, y las predicciones que realiza son producidas por cálculos como los que se hacen con un paralelogramo de fuerzas. Los conceptos de centro de gravedad y de paralelogramo de fuerzas son dos de los ejemplos más claros que Dennett ofrece para entender qué son las *Abstracta*. Del mismo modo que los cuerpos tienen *realmente* centros de gravedad, y que la tierra tiene *realmente* un Ecuador, las personas tendrían *realmente* creencias y deseos.

Como esbozamos en el capítulo 1, esta forma de entender las creencias y deseos, aunque configura una suerte de realismo moderado, contrasta con la versión realista más “dura” de Fodor, entre otros; según la cual compartir una creencia es compartir un estado interno estructuralmente similar. Las creencias y deseos sí serían considerados en este caso como estados internos reales con una clara incidencia causal. Las mismas estarían incluidas en ciertas leyes basadas en regularidades causales obtenidas inductivamente. El argumento de Fodor parte también de una evaluación respecto de cómo son considerados estos fenómenos por la psicología *folk*. Por ejemplo, señala que nuestras atribuciones no son el resultado de un cálculo idealizado y racional²⁷ - por lo que no podríamos decir que se desprenden de la actitud intencional-, sino más bien de un conocimiento obtenido a partir de ciertas regularidades empíricas. Ahora bien, aunque Dennett le concede a Fodor que en la cotidianeidad usamos muchas veces generalizaciones empíricas a la hora de atribuir creencias y deseos, esto no es un problema para su teoría de los sistemas intencionales, dado que la misma no propone captar de la manera más exacta posible cómo realizamos efectivamente estas atribuciones. A diferencia de esto, Dennett piensa que para formular su teoría, lo importante es dilucidar los aspectos de la psicología *folk* que pueden ser incorporados a la ciencia: esto es, cómo funciona en el mejor de los casos. Por ende, argumentar en su contra no consistiría en demostrar que la psicología popular procede de otro modo, sino en demostrar que debería hacerlo. De hecho, establece que el concepto de creencia que se encuentra en la comprensión común (en la psicología *folk*) estaría a mitad de camino entre *Illata* y *Abstracta*. Así, con respecto al hecho de que desde este conocimiento popular se

²⁷ Ejemplos de este tipo de actitudes serían pensar que alguien va a privilegiar la interpretación que lo deja mejor parado, o que alguien va a querer tener un producto si piensa que otra gente agradable lo tiene.

piense que al atribuir una creencia se está postulando un objeto estructurado de manera similar en cada cabeza, podemos decir que ese es simplemente un paso injustificado, una parte de la teoría que debemos dejar de lado.

Sin embargo, es preciso aclarar que Dennett diferencia su posición filosófica (su “ismo”),- de acuerdo con el cual creencias y deseos son abstracciones útiles- del *ficcionalismo*:

Algunos instrumentalistas han respaldado el ficcionalismo, el punto de vista de que ciertas afirmaciones teóricas son falsedades útiles, y otros han sostenido que las afirmaciones teóricas en cuestión no eran ni verdaderas ni falsas sino meros instrumentos de cálculo. No defiendo ninguna de estas variedades de instrumentalismo; como lo dije cuando usé el término por primera vez antes: “la gente tiene realmente creencias y deseos en mi versión de la psicología popular, tal como tiene verdaderamente centros de gravedad” (Dennett, 1987, p. 75).

Ahora bien, ¿en qué sentido Dennett puede decir que tenemos *realmente* creencias y deseos? Como ya dijimos, los *Abstracta* son constructos teóricos útiles y su utilidad no es azarosa, sino que se explica por el tipo de relación que guardan con los objetos concretos del mundo físico. En *Real Patterns*, Dennett (1991b) señala que la pregunta acerca de si los objetos abstractos son reales - y si se debe asumir una posición realista frente a ellos- puede verse de dos maneras, que podemos distinguir llamándolas el enfoque metafísico y el enfoque científico. El enfoque metafísico trata de manera general la realidad o la existencia de los objetos abstractos y no los distingue en términos de su utilidad científica. Dennett defiende que es, en cambio, el enfoque científico el que debemos adoptar. Desde esta perspectiva, podemos decir que los *Abstracta* “son *reales* porque son (de alguna manera) “buenos” objetos abstractos” (Dennett, 1991b, p. 29, la cursiva es mía). Los buenos objetos abstractos son aquellos que resultan útiles epistémicamente, es decir, que sirven como representaciones inteligibles de estados de cosas reales. Por ende, dicha utilidad no es azarosa, sino que se debe a que efectivamente representan, de modo condensado o simplificado, un estado de cosas en el mundo. En palabras de Dennett (1991b): “Si llegamos a distinguirlos como “reales” (contrastándolos quizás con esos objetos abstractos que son “irreales”) es porque pensamos que sirven como representaciones inteligibles de fuerzas reales, de características naturales y de cosas semejantes” (Dennett, 1991b, pp. 28, 29). La explicación que brinda Dennett da lugar a la

interpretación de que los objetos abstractos que consideramos reales desde el enfoque científico constituye un subgrupo de aquellos que consideraríamos reales desde el enfoque metafísico (un ejemplo de estos últimos es “el centro de las medias perdidas de Dennett” (Dennett, 1991b, p. 28), esto sería el punto de la esfera más pequeña que se puede inscribir abarcando todas las medias que Dennett ha perdido en su vida). Serían éstas las condiciones que reúne para Dennett un “buen” objeto abstracto: utilidad epistémica, que se debe a que representa un orden de cosas en el mundo.

Las creencias y los deseos son, entonces, constructos teóricos reales, en el sentido de que son abstracciones sumamente útiles. Su utilidad reside en que hacen referencia, de un modo simplificado, a ciertos estados de cosas que se dan efectivamente en el mundo (por ejemplo, que una persona esté dispuesta a conducirse de cierto modo, bajo determinadas circunstancias). Más específicamente, podemos decir que dan cuenta de ciertas capacidades con las que contamos.

En *La Actitud Intencional*, Dennett (1987) plantea la pregunta de por qué la actitud intencional funciona. Al respecto, señala que puede haber dos modos de entender esta pregunta, uno en el que resulta fácil de responder, y otro en el que es difícil. La respuesta “difícil”, que aquí no nos interesa, intentaría explicar cómo funcionan los mecanismos de los que están dotadas las mentes de los seres humanos de modo tal que nos hacen ser racionales, creer lo que debemos creer, y desear lo que debemos desear (atributos por los cuales la actitud intencional funciona cuando se aplica a personas). La respuesta “fácil” es que la evolución ha diseñado a los humanos para poseer estos atributos: “El hecho de que seamos los productos de un proceso evolutivo largo y exigente garantiza el hecho de que usar la estrategia intencional en nosotros sea una apuesta segura” (Dennett, 1987, p. 42). Cuando hablamos de las creencias y deseos como *Abstracta*, hacemos referencia a determinados modelos o patrones a los que da lugar esta capacidad que hemos adquirido evolutivamente, independientemente de cuáles sean las respuestas que podamos encontrar en un futuro para la pregunta difícil.

Los centros de gravedad narrativos como *Abstracta*

A partir de lo desarrollado anteriormente, podemos decir que la respuesta que Dennett brinda a la pregunta de qué tipo de cosa es el “yo” se desarrolla dentro de un

marco ontológico de acuerdo con el cual existen algunas cosas en el mundo que no están simplemente construidas con los “ladrillos” de la física y que no por eso son fantasmagóricas o misteriosas. En lugar de intentar eliminarlas, Dennett sostiene que pueden ser explicadas y que son imprescindibles desde un punto de vista teórico. Estas son las entidades que propone entender como *Abstracta*, etiqueta bajo la cual pienso que debería ubicarse al “yo” dennettiano. Uno de los principales ejemplos de *Abstracta* señalado por Dennett (1987, 1991b) es el de los centros de gravedad en la física y es precisamente apelando a una analogía con ellos que esboza su primera explicación del “yo” (1988). La analogía del “yo” como centro de gravedad narrativo con el centro de gravedad de un objeto se mantiene en pie en algunos aspectos, mientras que en otros encontramos diferencias. Explicitar algunas de las características de los centros de gravedad de los objetos, e intentar establecer los puntos de la analogía puede ser útil para comprender mejor a nuestro mucho menos familiar *Abstracta* en cuestión: el “yo”.

Como sabemos, los centros de gravedad de los objetos no son identificados con ninguna de las partículas materiales con las que la física contemporánea amuebla el universo. La única propiedad física que poseen es la de tener una ubicación espacio-temporal, careciendo de color, masa, y de todas las demás. En ese sentido decimos que es una ficción teórica (es decir, una ficción en el marco de una teoría); pero se trata de una ficción robusta y bien definida, que posee un rol determinado dentro de la física y que nos sirve para hacer predicciones, incluso en nuestra vida cotidiana. Así, es una construcción sumamente útil, y la razón de esa utilidad es que refleja un determinado modo en el que los *Concreta* se relacionan efectivamente entre sí. Si alguien que estuviera investigando un objeto complejo que se encuentra en movimiento constante en una superficie, pudiese determinar precisamente dónde se encuentra el centro de gravedad de dicho objeto en movimiento, y descubriera que su ubicación coincide exactamente con la de un átomo o una partícula subatómica, podría verse tentado a identificarlos; sin embargo, ello supondría caer en un error categorial. En ese sentido no hay que olvidar que las abstracciones son sólo objetos ficticios, “pero cuando digo que es un objeto ficticio no lo hago menospreciándolo, es un maravilloso objeto ficticio” (Dennett 1988, p. 16, mi traducción). Lo mismo sucede con el “yo”. Por un lado, caeríamos en un error categorial al identificarlo con una parte del cerebro, dado que se trata de una abstracción. Por otro lado, eso no quita que sea un “buen” objeto ficticio.

El centro de gravedad de un objeto físico está dado por la *resultante* de todas las fuerzas de gravedad que actúan sobre las porciones materiales del objeto. Esta fuerza *resultante* se calcula por medio de una suma vectorial, es decir de una suma en la que no sólo intervienen magnitudes, sino también direcciones y sentidos. Se puede pensar, análogamente, que la fuerza que produce determinada narración acerca de alguien (con su “magnitud”, su “dirección” y su “sentido” propios), al ser sumada con otra narración sobre él, puede anularse, desplazarse, acortarse o alargarse, “ubicando” al “yo” que es protagonista de las mismas en un lugar u otro con respecto al todo constituido por las cosas que se dicen actualmente y que se han dicho con anterioridad sobre él mismo -y tiene cierto peso según su valoración-. Los términos espaciales deben ser entendidos de modo estrictamente metafórico para el caso del “yo”, dado que una de las diferencias entre éste y los centros de gravedad es que no posee una localización precisa, antes bien, su posición espacio temporal está groseramente definida (Dennett, 1988). Sin embargo, la analogía es útil para señalar que, del mismo modo que el centro de gravedad como fuerza resultante es simplemente el resultado de la sumatoria de las fuerzas que actúan sobre las distintos porciones materiales del objeto, el “yo”, que sufre los “desplazamientos” de las narraciones que lo constituyen, no es algo previo a las “fuerzas” de las narraciones, sino la “fuerza” misma que, se podría decir, surge con la primera narración acerca de alguien y que va sufriendo modificaciones producto de la “adición” de nuevas “fuerzas” dadas por nuevas narraciones.

El hecho de que nadie haya visto nunca un centro de gravedad no lo hace un concepto menos robusto. Del mismo modo, la robustez aparente del “yo” – abstracción del mismo tipo y de mayor complejidad- no entra en conflicto con el hecho de que no lo podamos ver, fenómeno bien señalado por Hume en el *Tratado de la Naturaleza Humana* (1739). Si bien estas dos maravillosas y útiles abstracciones poseen algunas características en común, la relativa al “yo” constituye un concepto mucho más enrevesado, que intentaremos dilucidar en lo que resta de este capítulo.

Así, en función del concepto de *Abstracta* de Reichenbach y del modo en que Dennett lo aplica para entender a los objetos intencionales, podemos afirmar que los “yoes” no tienen una existencia separada, un cuerpo hecho de “sustancia mental”, ni se pueden identificar con una parte del cuerpo o del cerebro. Antes bien, son una

forma abstracta de ordenar los *Concreta*. Con la aparición de un “yo” no comienza a existir de un momento al otro una cosa nueva en el mundo además las narraciones de las cuales ellos son el centro de gravedad narrativo. Sin embargo, para entender de qué modo una abstracción tan compleja como la del “yo” tiene una relación de equivalencia con objetos concretos en el mundo, la analogía con los centros de gravedad de los objetos resulta muy pobre. Por ese motivo, reconstruiré un par de aspectos en común que el “yo” guarda con otro fenómeno con el que Dennett (1988, 1992) hace una analogía: el personaje de una novela. Sin embargo, antes de desarrollar dicha analogía realizaré una aclaración para evitar una interpretación errónea de los alcances de la misma.

En sus artículos tempranos sobre la identidad personal, Dennett (1988, 1989, 1992), habla del “yo” como de una ‘ficción’. Podría pensarse que, con esto, Dennett se refiere al “yo” como un invento arbitrario. Sin embargo, como señalamos al comienzo del apartado, en los mismos textos Dennett señala que los centros de gravedad de los objetos -ejemplo de *Abstracta* por excelencia- pueden ser entendidos como “objetos ficticios”. Así, el hecho de que diga que el “yo” es una ficción no tira abajo la idea de que podemos entenderlo como un *Abstracta*. Por ende, dado que con respecto a las creencias y los deseos Dennett se posiciona diferenciándose del instrumentalismo y del ficcionalismo, debemos considerar que su teoría de la identidad personal tampoco constituye una versión de los mismos. En esta línea, es preciso señalar que la ficción que constituye al “yo” no debe ser entendida como una ficción en el sentido de un invento arbitrario, o una mero instrumento de cálculo, sin valor de verdad. Para justificarlo, podemos apelar a la categoría elaborada por Dennett (1991b) de “buenos” objetos abstractos, que son, en ese sentido, “reales”. Como dijimos, se trata de abstracciones epistémicamente útiles que, a su vez, deben su utilidad al hecho de que captan determinados hechos y relaciones efectivas de los objetos concretos. En función lo expuesto podemos sostener que cuando Dennett habla del “yo” como una ficción debemos entender que se trata de un objeto ficticio “real”, en el mismo sentido en el que lo son los buenos objetos abstractos, como los centros de gravedad y las creencias.

Ahora bien, Dennett dice que el “yo” puede pensarse como un personaje literario: “Agarremos *Moby Dick* y lo abramos en la primera página. Dice “Llámenme Ishmael” ¿A quién llamamos Ishmael? ¿Llamamos Ishmael a Melville? No, llamamos

Ishmael a Ishmael, el personaje ficticio que ha creado Melville” (Dennett, 1988, p. 17). Algo similar ocurre cuando relatamos historias, y atribuimos los hechos narrados a nosotros mismos o a los demás: contribuimos al conjunto de narraciones que giran en torno a un “yo”.

En esta analogía también podemos establecer aspectos de continuidad y otros de diferencia. Cuando leemos una obra literaria, aprendemos de sus personajes a través de lo que la misma nos va relatando. Algunos datos los podemos inferir. Lo que se puede saber del personaje (que se identifica con lo que el personaje es, dado que no podemos hablar de aspectos de un personaje ficticio que nadie conoce) depende del texto, pero también de la cultura en la que ese texto fue producido, y que debe tomarse como marco de referencia al leerlo. Pero, más allá de esa extrapolación, los mundos ficticios son simplemente indeterminados: respecto de un sinnúmero de aspectos o rasgos posibles, simplemente no se puede determinar si existen o no, ni, por la tanto, afirmar o negar su existencia. Hay cosas que podemos afirmar sobre esos personajes sin que se nos las hayan dicho explícitamente, pero hay preguntas sobre ellos que simplemente no tienen una respuesta ni afirmativa ni negativa; ¿acaso tiene Sherlock Holmes un lunar en su hombro izquierdo? Si en el texto no se ha hecho referencia a ello, no podemos afirmar ni negar tal cosa. Con los hombres reales, ocurre que la pregunta sobre si alguien tiene o no un lunar en su hombro izquierdo, tiene una respuesta determinada, afirmativa o negativa. Incluso aunque no lo podamos saber nunca, la afirmación de que Aristóteles tenía un lunar en su hombro izquierdo es, o bien verdadera, o bien falsa. Sin embargo, hay un punto en el que nuestras ficciones útiles no se alejan tanto de los personajes ficticios literarios. Aunque en las ficciones literarias podemos esperar que la cantidad de información disponible haga a los mundos inventados mucho más indeterminados que el mundo real, en los relatos de los “yoes” reales también podemos encontrar, en un grado menor, algún grado de indeterminación. Así, ambas entidades sólo poseen las propiedades que las teorías que los constituyen han definido para ellas. Respecto de los “yoes”, también habría casos en los que no estaríamos justificados en atribuir ni en no atribuir determinados hechos o estados mentales a un “yo” particular. Un ejemplo podría ser, quizá, el de alguien que nunca se planteó el objetivo de visitar Japón, y para quien no puede decirse ni que desea ni que no desea viajar a Japón, o el de alguien que desconoce la existencia de un hecho cualquiera, y para quien carece de sentido preguntarse qué opinión se ha formado de él. El hecho de que las

propiedades que se le puedan atribuir a la abstracción en cuestión, estén limitadas por los contenidos de las narraciones de las cuales ésta es su centro, no significa que las afirmaciones que involucran al “yo” no posean valores de verdad; excepto por esta restricción de contenidos, dichas afirmaciones son verdaderas o falsas.

En consonancia con lo anterior, hay una gran diferencia entre nosotros y los personajes ficticios en los referido a la relación que guardan con la consistencia en sus constructos. Si un personaje de una larga saga se encuentra, según el contenido de dos libros diferentes, en lugares distintos al mismo tiempo, vamos a interpretar esto como un despiste del escritor. En cambio, los únicos casos en los que podríamos encontrar inconsistencia en los relatos sobre alguien real -como por ejemplo una persona sosteniendo un día que es completamente atea y a los pocos días diciendo que es una católica convencida y siempre lo ha sido, sin tratarse de una mentira intencional- son los casos de Trastorno de Personalidad Múltiple, en los que, justamente, se trata de más de un “yo”. Así, los “yoes” como abstracción tienen más restricciones de consistencia que los personajes literarios.

Otra diferencia entre los “yoes” y los personajes literarios se puede establecer a partir de que estos últimos aparecen siempre como “hechos consumados” (Dennett, 1992). Como dijimos, hay preguntas que ya no tienen respuestas, dado que serían datos que no se pueden inferir del texto. Si en ninguna de las novelas de Doyle, Sherlock Holmes aparece comiendo un lemon pie, o algún contexto nos induce a pensar que lo está haciendo, las preguntas acerca de si alguna vez lo probó, o si le hubiera gustado, son preguntas que sencillamente carecen de respuestas. En cambio, si yo nunca en mi vida probé el lemon pie (y no soy alérgica a ninguno de sus ingredientes), afortunadamente estoy a tiempo de hacerlo y de responder a todas las preguntas que quieran hacerme sobre qué me parece. La diferencia en cuestión va más allá del hecho de que nosotros tenemos experiencias por delante. Dennett propone que pensemos en un personaje que aparece en varias novelas, obras de un escritor que todavía vive. Sería posible pedirle al autor que escriba una nueva novela, en donde responda a las dudas que tenemos acerca de ese personaje. Esto no es lo que habitualmente ocurre con los personajes de ficción. Sin embargo, es algo que en realidad hacemos a menudo con nuestros “yoes”: nos hacemos preguntas los unos a los otros y nos vamos construyendo o reconstruyendo a nosotros mismos, a nuestros “yoes”, en función de esas demandas. Si bien no podemos deshacer o anular las

partes de nuestro pasado que están determinadas, nuestros “yoes” están determinándose constantemente en función de cómo el mundo nos afecta. Incluso es posible que una persona se involucre hermenéuticamente con su pasado, con la interpretación de sí mismo y de sus memorias, lo que puede llevarlo a re-pensar y a re-escribir su pasado. Este proceso de reescribirnos a nosotros mismos haciéndonos cada vez más determinados, sería completamente mágico e incomprensible si el “yo” no fuera una abstracción. Nuevamente, el concepto de *Abstracta* de Reichenbach es sumamente fructífero dado que, como señalamos al comienzo, permite dar cuenta del hecho de que hay un tipo de abstracción que, a diferencia del número π , puede tener un nacimiento y una muerte y sufrir profundas modificaciones. Del mismo modo que un centro de gravedad de un objeto puede ser modificado por nuestra intervención sobre el mismo, nuestro “yo” como centro de gravedad puede ser alterado a partir de la emergencia de nuevas historias de las que él es protagonista.

En el capítulo 1 explicamos que, para Dennett las creencias, en la medida en que son abstracciones que ordenan de un modo particular los objetos concretos, pueden ser consideradas como patrones. Si consideramos que el “yo” tiene el mismo tipo de existencia que tienen los estados intencionales, se podría pensar que el “yo” de algún modo también es un patrón. La analogía con los personajes literarios puede ser útil para interpretar mejor en qué consistiría entender al “yo” como patrón. Si entendemos a un patrón como una regularidad distinguible, y a una narración como al resultado de una referencia lingüística (o visual) de una sucesión de hechos que le ocurren a un personaje, se puede pensar que en el marco de una estructura narrativa como la señalada siempre hay un primer patrón elemental: el personaje mismo. El “yo” es aquello a lo que le atribuimos las historias de vida de alguien. Los “yoes”, tendrían entonces el mismo carácter dual de los patrones: por una parte, dependen de las historias que los hacen ser personajes a los que les acontecen los hechos, es decir, de algún tipo de atribución narrativa, al tiempo que, por otra parte, su existencia tiene un componente objetivo que los distingue de los meros personajes ficticios.²⁸

²⁸ Recientemente me han señalado que en la lengua inglesa “center of gravity” es una expresión metafórica corriente, que está vinculada con el verbo “gravity” que significa acontecer, suceder. El centro de gravedad es aquello en torno a lo que suceden los eventos (por lo que muchas veces se lo utiliza para señalar que es lo importante de algo. Por ejemplo: “For this reason, the War College has determined that its center of gravity is its students” (extraído de la página: <http://www.linguee.com/english->

Clases de psicología intencional

La distinción reichenbachiana entre tipos de entidades desarrollada al comienzo del capítulo puede servirnos también para entender mejor una clasificación que realiza Dennett entre tres clases de psicología intencional (Dennett, 1987, cap: “Tres clases de psicología intencional”). La primera de ellas es la *Psicología Folk*, que, como ya explicamos, además de ser imprescindible para movernos en el mundo, puede entenderse como una “fuente de teorías”, en la medida en que mucho de lo que pregona puede ser entendido como buen candidato para ser incorporado al *corpus* científico, para lo cual deberá pasar por un examen crítico que elimine los aspectos mal fundados. A partir de una depuración de dicha psicología *folk* se obtiene la *Teoría de los Sistemas Intencionales*, como modelo generador de teorías (que, a diferencia de ella, que versa sobre una habilidad específicamente humana de interacción cotidiana, está pensada como herramienta que permite generar teorías en muy diversos campos, como la etología cognitiva, la economía, la teoría del juego, entre otras). A su vez, este segundo tipo de teoría psicológica puede ser muy útil para el desarrollo de una *Psicología Cognitiva Sub-personal*, que debería permitir explicar, en términos de procesos internos del sistema, por qué la teoría de los sistemas intencionales funciona para generar predicciones del comportamiento humano. La teoría de los sistemas intencionales es de carácter holista, normativa, idealizadora e instrumental, la mayoría de sus explicaciones son de tipo conceptual, tienen lugar en un nivel personal, y sus términos propios son los *Abstracta*. Por otra parte, la psicología cognitiva sub-personal, que puede entenderse como el resultado de un abordaje de la mente humana desde la actitud del diseño, es una teoría científica y de carácter empírico de la comprensión de los sistemas intencionales. La mayoría de sus explicaciones son causales y en el nivel sub-personal. Aquí sí encontramos *Illata* como entidades importantes para conformar las explicaciones.

Desde nuestra interpretación, lo examinado hasta aquí nos permite defender la hipótesis de que el fenómeno del “yo”, según la propuesta de Dennett, debe ser ubicado en el nivel de la psicología *folk*. Si investigamos en el nivel de la psicología

spanish/translation/its+center+of+gravity.html). Esta forma de abordar la metáfora arroja resultados coherentes con los que desarrollado hasta aquí, dado que el Yo sería aquello en torno a lo cual acontecen los eventos de las narraciones que construimos con otros sobre nuestra vida. Agradezco a Carolina Mahler este aporte.

cognitiva subpersonal, podemos encontrar una serie de dispositivos que dan lugar a algo así como una “ilusión de Usuario” (que explicaremos en el capítulo siguiente), pero ningún fenómeno físico último que podemos llamar con el nombre de “Yo”. Por ende, si descartamos la posibilidad de ubicarlo en el marco de la psicología cognitiva sub-personal, hay que considerarlo como un fenómeno propio de la psicología *folk*, o de la teoría de los sistemas intencionales. Ahora bien, con respecto a esta última opción, si destilamos el fenómeno con el que nos encontramos en la psicología *folk*, del modo en que lo hacemos con los fenómenos “mentalistas” para formular explicaciones en el nivel de la teoría de los sistemas intencionales, podemos obtener un concepto de agente y de portador de las creencias y deseos, pero, como intenté defender en el capítulo anterior, las narraciones que constituyen el “yo” exceden los tópicos de lo que se atribuye a un sistema intencional. Por ende, ubicar al “yo” en el nivel de la teoría de los sistemas intencionales, sería ofrecer una versión demasiado deflacionada de nuestro fenómeno de carácter narrativo; y, por lo tanto, no parece que podamos considerarlo equivalente a nuestro concepto de “yo”. Por otra parte, la teoría de los sistemas intencionales es un modelo generador de teorías, que también puede aplicarse a, por ejemplo, estudios etológicos o económicos, por lo que este fenómeno relativo al modo en que nos relacionemos los humanos en la constitución de nuestra identidad personal, no parece poder contemplarse de manera exhaustiva desde un modelo que sirve para generar teorías en campos tan diversos.

Ahora bien, el “yo”, con toda su riqueza y complejidad, puede ser abordado de manera íntegra desde la psicología *folk*, dado que no sólo incluye narraciones de tipos muy diferentes (intencionales, morales, históricas, estéticas, curiosas, raras), sino que además es en sí mismo un dispositivo “creado para” hacer posible esa interacción cotidiana. En algún punto, pues, la respuesta de Dennett a qué tipo de cosa es el “yo” propone un cambio en el eje de la pregunta: ya no se trata sólo de si hay algo en términos científicos a lo que nos referimos con este término, sino de que, desde que asumimos que es algo -que es lo que hacemos, y eso llevó a originar la pregunta-, hace todo lo que necesitamos de él. Lo que hace que el “yo” sea lo que es, es aquello para lo que se lo usa: es usado por los otros para entender a ciertos dispositivos como *alguien*, y por cada uno de nosotros, como un punto luminoso de autorrepresentación (Dennett, 1991a) (que es lo que permite que devenga en un *locus* de auto-control (Dennett, 2003) -cuestión a desarrollar en el capítulo 5-).

Se puede hacer una última aclaración, retomando que, tanto en el capítulo anterior, como en esta sección, dije que las narraciones que constituyen al “yo” pueden exceder, en tópicos y complejidad, a los relatos que construimos desde la *actitud personal* (Dennett, 1973, 1976). De aquí se seguiría que, en Dennett, el concepto de “Yo” difiere del de “persona”. Pienso que esta es una lectura adecuada, dado que el de persona es un concepto que se aplica de forma menos específica, en tanto que refiere al modo en que interpretamos a alguien como un sistema intencional capaz de tener atributos morales, es decir, que rescata un aspecto general de aquel a quien se atribuye. En cambio, en función de todo lo expuesto, podemos decir que el del “yo” es un concepto mucho más específico, que se le atribuye a cada persona en particular, en función de las miles de cosas que componen su historia. Por otra parte, podría pensarse que, en Dennett, todo aquel sistema frente al cual podemos adoptar la actitud personal es, al mismo tiempo, un “yo”, dado que las condiciones necesarias para ser considerado una persona (la capacidad de reciprocidad y de comunicación verbal) no parecen ser posibles sin que el agente haya ingresado a un ámbito de comunicación regido por el lenguaje natural y haya sido integrado a él como un punto de representación al que se le atribuyen las vivencias de un ser humano particular. En cambio, creo que podría pensarse que el concepto de “yo” sí podría aplicarse para algunos seres humanos a los que no se atribuye responsabilidad moral, por ejemplo, a un anciano con una discapacidad muy severa, que ya no es capaz de tener capacidades como la de la reciprocidad de tratar a los otros como personas, pero sigue pudiendo referirse a sí mismo como un alguien y, por ende, siendo un centro de gravedad narrativo. Sin embargo, una reflexión acabada acerca del vínculo entre estos dos conceptos en Dennett debería ser objeto de futuras investigaciones.

4. Cómo se constituye el “yo”

Para entender mejor la concepción narrativista de Dennett, intentaremos reconstruir su explicación acerca del origen y la constitución del “yo”. Al bosquejar una narración histórica que comienza con los organismos más primitivos²⁹ en los orígenes de la vida biológica y conduce hasta la aparición de “yoes” humanos, su explicación queda estrechamente vinculada con la de aquello que hace posible la existencia de estos últimos: la aparición de la conciencia. En el primer capítulo expliqué la base de la teoría del contenido de Dennett que, según las palabras finales de *La conciencia explicada* (Dennett, 1991a), es independiente y más fundamental que su teoría de la conciencia. Estos dos grandes tópicos - contenido y conciencia- se corresponden justamente con las dos partes de su libro temprano *Contenido y Conciencia* (1969), cada una de las cuales fue adquiriendo cada vez mayor sofisticación hasta merecer cada una un libro propio: *La Actitud Intencional* (1987) y *La Conciencia Explicada* (1991a), respectivamente. En la primera de las teorías se trata de manera uniforme el problema general del contenido (ya sea el atribuido a sistemas dotados de cerebros, computadoras o procesos evolutivos), estableciendo así una base sobre la que puede construirse una explicación de los procesos conscientes.

En el desarrollo de la explicación dennettiana de los orígenes de la conciencia, podremos ver en acción su espíritu de colaboración con las ciencias, dado que incorpora explicaciones brindadas por biólogos, neurocientíficos, psicólogos e informáticos. Uno de los ejes que orientan dicha investigación será una pregunta planteada en los artículos tempranos sobre el “yo” (Dennett, 1988, 1989, 1992): para qué nos sirve la práctica de hablar con nosotros mismos - que es una de las prácticas constitutivas de la creación narrativa de los “yoes”. Dicha pregunta terminará vinculándose a otra que también se formula en estos escritos tempranos sobre el “yo” (puntualmente en Dennett 1988, 1989). Si bien estos textos son anteriores a *La Conciencia Explicada* (Dennett, 1991a), en donde encontramos un desarrollo pleno

²⁹ El término “primitivo” no debe ser leído con connotaciones negativas. Simplemente se hace referencia a que el origen de sus rasgos se encuentran en mayor proximidad con los orígenes de la vida.

de su teoría de la conciencia, ya se encuentran algunas ideas germinales de lo que se desarrollará en este libro, dado que concibe al funcionamiento del cerebro bajo las características de que el procesamiento cerebral se da en paralelo y presenta cierto grado de inaccesibilidad entre sus subsistemas. Con respecto a esto, Dennett advierte que si tal concepción es acertada “podría llegar a proveernos la respuesta a una de las más enigmáticas preguntas sobre el pensamiento consciente: ¿qué tiene de bueno?” (Dennett, 1988, p. 12, mi traducción). Dennett aclara que esta es una pregunta que solicita una respuesta evolutiva.

La idea de que el interrogante en cuestión requiere una respuesta evolutiva conduce a una discusión vinculada con el estatus epistemológico de las explicaciones en evolución; específicamente, con la posibilidad de inferir relatos evolucionistas a partir de una observación de las ventajas adaptativas actuales que podemos atribuir a un rasgo determinado. Este es, justamente, uno de los principales blancos de ataque contra Dennett. En palabras de Brook y Roos: “Aunque es un poco difícil entender por qué Dennett terminó siendo acusado de ser un Ultra-darwinista (lo que sea que signifique eso) y el perrito faldero de Dawkins” (2002, p. 9). Es cierto que su modo de entender la evolución se inscribe dentro de un paradigma más bien clásico, como lo dejan ver sus frecuentes alusiones a Dawkins. El pensamiento de muchos de los autores inscriptos en este paradigma ha recibido severas críticas en los últimos años, provenientes de distintos campos y perspectivas. Gould ha elaborado algunas de las más importantes. Entre ellas, su crítica al “gradualismo”, considerándolo como uno de los aspectos que la Nueva Síntesis Moderna debería revisar. Pero las críticas que vienen al caso en este contexto son las que Gould formula, conjuntamente con Lewontin (1979, 1983), en contra de lo que llaman el paradigma “pan-adaptacionista”, y en la propuesta, realizada conjuntamente con Vrba (1982), sobre el reconocimiento de las “exaptaciones”. Dennett se hizo eco de dicha discusión, según sus propias palabras, gracias a que determinados críticos que “afirmaban que mi posición era una versión del adaptacionismo, atrajeron mi atención hacia el trabajo de aquellos” (Dennett, 1987, p. 231). Antes de realizar una reconstrucción de la explicación dennettiana del origen de la conciencia, haremos una breve referencia a dichas objeciones para poder valorar su relevancia.

Explicaciones evolucionistas: Algunos reparos

Gould es uno de los más destacados críticos de los intentos de elaborar explicaciones de las características actuales de distintas especies –la humana, en el caso que nos convoca- a partir de enfoques evolucionistas. Dejando de lado su desarrollo de la teoría del equilibrio puntuado -que se opone a la idea de que todo cambio evolucionista se desarrolla de manera gradual-, a partir del conjunto de sus otras críticas es posible reconstruir dos núcleos problemáticos diferentes que son de algún modo complementarios. Uno de ellos se centra en el reconocimiento del fenómeno que Gould y Vrba proponen llamar *exaptación* (1982). El otro, vinculado al anterior, es la crítica dirigida por Gould y Lewontin hacia lo que ellos denominan “pan-adaptacionismo” o “paradigma panglossiano” (1979, 1983). Ambos errores serían el resultado de la falsa concepción que los defensores del paradigma clásico tienen sobre la omnipotencia de la Selección Natural.

El famoso artículo de Gould y Vrba (1982) comienza señalando que en las investigaciones y debates biológicos actuales conviven dos significados diferentes de *adaptación*. Uno relativo al origen, según el cual una característica es una adaptación si fue creada [*built*] por la selección natural para la función que desempeña ahora, y otro relativo a la utilidad actual, en el que se entiende a la adaptación como cualquier característica que mejora el *fitness*³⁰. En la propuesta de los autores, sólo la primera de estas definiciones recoge de forma adecuada lo que es una adaptación, mientras que la segunda constituye un grupo más amplio, que debería ser llamado *aptación*. El conjunto de las aptaciones estaría compuesto por los subconjuntos excluyentes de adaptaciones y exaptaciones. Las exaptaciones, a su vez, se definirían como “caracteres evolucionados para otros usos (o para ninguna función en absoluto) y luego “cooptados” para su rol actual” (p. 6).

¿Qué se pone en juego en esta distinción entre adaptaciones y exaptaciones? Se puede pensar este planteo a través del testimonio de Williams, según el cual sería importante no confundir las verdaderas adaptaciones y sus funciones con “los efectos meramente fortuitos” (1966, p.5). ¿Qué haría a las exaptaciones más azarosas que las adaptaciones? Cuando hablamos de adaptaciones se entiende que la aleatoriedad

³⁰ El término *fitness* es difícil de traducir al castellano. Se trata de la “aptitud”, “ajuste” o “encaje” de algo en el medio. Su acepción técnica hace referencia a la capacidad de un organismo para sobrevivir más tiempo, dejar más descendencia y, con ella, mayor herencia genética en las generaciones siguientes.

cumple un rol central en la generación de la variabilidad, pero no en la perpetuación de la misma: los caracteres que perduran son los elegidos por la Selección Natural por la función que desempeñan. La exaptación añadiría un segundo momento azaroso: no sólo el que se da a partir del encuentro entre la generación de la variación y los motivos de su perpetuación, sino el que aparece entre la perpetuación de determinada característica y su cooptación para un nuevo rol. Una vez señalado esto, podemos profundizar en una distinción ulterior. Las fuentes de la exaptación son tanto viejas adaptaciones como no-aptaciones [*nonaptations*], es decir, características que no cumplen ninguna función adaptativa. Este último caso resulta más innovador con respecto al marco conceptual tradicional que se está criticando. De hecho los mismos autores señalan que “(S)í todas las exaptaciones hubieran comenzado como adaptaciones para otra función en sus ancestros, nosotros no hubiésemos escrito este artículo” (Gould y Vrba, 1982, p.12). La simple incorporación de las exaptaciones que previamente fueron adaptaciones se podría dar en el marco de una concepción que, aunque contemple la posibilidad de que una característica sea cooptada para un nuevo uso, siga entendiendo que todas ellas han permanecido por la acción de la Selección Natural. Lo que se sigue de la noción de exaptación tal como la entienden Gould y Vrba, en cambio, es que, además de actuar sobre las mutaciones aleatorias, la Selección Natural actúa sobre el pool de posibilidades constituido tanto por viejas adaptaciones como por no-aptaciones. Más allá de las nuevas terminologías, la gran novedad que, desde mi punto de vista, estaría introduciendo esta forma de relatar los procesos evolutivos, sería la incorporación de la posibilidad de que una nueva característica se extienda en una especie y perdure en el tiempo, sin que el proceso de la Selección Natural haya tenido responsabilidad previa en ello.³¹

³¹ Sin embargo, hay algo extraño en el concepto de no-aptación. Los autores mismos señalan la ironía de definir por la negativa a aquellas características que no tienen *fitness* actual; en tanto esto puede ser visto como un efecto no deseado de nuestra falsa concepción de la primacía de la adaptación. Dennett (1995) da cuenta de forma mucho más aguda de lo trivial que puede llegar a resultar hablar de las características de la biósfera que existen sin razón alguna: “Por ejemplo, la propiedad del elefante de tener más extremidades que ojos, la flotabilidad de las margaritas” (p. 452). Como respuesta a Dennett, podríamos tomar uno de los ejemplos de exaptaciones que Gould y Vrba incluyen en su artículo. Se trata del caso de las hienas, cuyos órganos sexuales femeninos y masculinos son indistinguibles, lo que permite la realización de un ritual de reconocimiento particular, dependiente de dicha indistinción. Dado que el ritual de reconocimiento es muy importante para la supervivencia de una hiena -el tipo de vida que lleva contiene etapas de soledad y etapas de incorporación en manadas- muchos biólogos entendieron al mismo, y la mimesis que lo hace posible, como

Por otra parte, la fuerte crítica que Gould y Lewontin dirigen a las explicaciones evolucionistas tiene múltiples aristas. El blanco de ataque central es el mecanismo explicativo utilizado por muchos teóricos evolucionistas, según el cual todos los caracteres de un organismo son adaptativos. Sostener esto último es, una vez más, sostener que la Selección Natural es omnipotente y, según los autores, esto constituye una reencarnación del pensamiento del Doctor Pangloss, el personaje creado por Voltaire (1759) para ridiculizar la idea leibniziana de que vivimos en el mejor de los mundos posibles. Dicha concepción conduce a “atomizar” a los organismos que se estudian, intentando buscar la utilidad adaptativa de cada parte tomada aisladamente. En definitiva, es la creencia según la cual todas las características de un organismo fueron creadas para aumentar la supervivencia. La diferencia que los críticos tendrían con respecto a este paradigma “pan-adaptacionista” es que sostienen que hay variaciones que han permanecido sin haber sido elegidas por la Selección Natural.

Los artículos de Gould y Lewontin concluyen con un llamado al pluralismo explicativo y un señalamiento de que sus aportes no buscan atacar el marco teórico darwiniano, sino pulirlo. Estas conclusiones contrastan llamativamente con las consideraciones finales del texto escrito entre Gould y Vrba (1982), en el que se defiende que la existencia de las exaptaciones constituye un buen motivo para abandonar las explicaciones evolucionistas sobre nuestro cerebro y nuestra mente:

Mucho de lo que nuestro cerebro hace ahora para mejorar nuestra supervivencia descansa en el dominio de la exaptación(...)¿Cuánto de la literatura evolutiva sobre el comportamiento humano puede colapsar si incorporamos el principio de exaptación en el corazón de nuestro pensamiento evolutivo? Este colapso podría ser constructivo, porque se ampliaría mucho nuestro rango de hipótesis, y focalizaría la atención en la función actual y del desarrollo (todas propuestas

una adaptación creada por la Selección Natural. Sin embargo, es mucho más sensato pensarlo como una exaptación de una característica que había perdurado sin utilidad alguna: la masculinización de los órganos sexuales femeninos. Este sería, según los autores, un ejemplo de una no-aptación. Sin embargo, este carácter no seleccionado es, a su vez, entendido como el efecto secundario de otra adaptación: el elevado nivel de andrógeno en las hienas hembras. Así, la masculinización de los órganos sexuales femeninos de las hienas constituye un ejemplo no trivial de exaptación (como sí lo es la flotabilidad de las margaritas), pero cabe señalar que es mucho menos extremo que la afirmación de que hay características que perduren porque sí; más bien parece un tipo de relato que todo investigador sensato debería estar dispuesto a aceptar.

testeables), en lugar de dejarnos llevar por fantasías imposibles de probar (Gould y Vrba, 1982, p. 13).

En el siguiente apartado reconstruiré las explicaciones desarrolladas por Dennett para dar cuenta del origen de la conciencia mostrando que estas constituyen un buen ejemplo del valor que puede tener continuar adhiriendo, pese a las críticas de Gould y colegas, a las hipótesis evolucionistas en el campo de la filosofía de la mente y del comportamiento humano, a fin de obtener hipótesis heurísticamente prometedoras. Asimismo, intentaré mostrar que, al adoptar hipótesis evolucionistas, Dennett no cae en los errores que acabamos de señalar. Por una parte, en el desarrollo de su teoría de la evolución de la conciencia, Dennett no deja de lado la posibilidad de encontrar exaptaciones; antes bien, las incluye como parte del proceso. Dicha incorporación es consecuente con un reconocimiento teórico general de dicho fenómeno por parte del filósofo, que se deja ver en discusiones como su crítica a la *falacia genética*: “el error de inferir la función o significado actual de la función o significado ancestral” (Dennett, 1995, p. 456); lo que nos lleva a pensar que su decisión de no acuñar el término estaría vinculada a que no lo considera una verdadera innovación: “como el propio Gould nos ha *recordado*, uno de los mensajes fundamentales de Darwin es que los artefactos pueden reciclarse para cumplir nuevas funciones” (Dennett, 1995, p. 450, la cursiva es mía).³²

Por otra parte, con respecto a la crítica al pan-adaptacionismo, Dennett realiza una defensa explícita del adaptacionismo como modelo para buscar explicaciones científicas. En un interesante capítulo titulado “Los sistemas intencionales en la etología cognitiva: Defensa del ‘Paradigma panglossiano’” (Dennett, 1987), compara los argumentos del Gould y Lewontin contra el adaptacionismo, con los esgrimidos por Skinner (1964) en contra del “mentalismo”. En la defensa de las perspectivas criticadas, Dennett traza una analogía entre el adaptacionismo y la actitud intencional, en función de que ambos dependen de la adopción deliberada de

³² En realidad, sí he encontrado el empleo del término en un texto sobre el lenguaje y las capacidades cognitivas. Después de muchas referencias a la idea de Gould, en distintos textos, en las que Dennett señala que se trata simplemente de algo que ya había dicho Darwin, y parece haber una decisión deliberada de no emplear el término *exaptación*, el filósofo acude a él en un texto mucho más reciente, que puede ser tomado quizá como una señal de que ha dado por agotada la discusión: “se han desarrollado a partir de algo previo, de algunas capacidades que pueden ser empleadas o *exaptadas* (para usar el término de Stephen Jay Gould) para las nuevas tareas de componer, salvar, recuperar datos, revisar, comparar estos nuevos objetos internos” (Dennett, 1993, p. 542, mi traducción).

modelos óptimos. No voy a realizar aquí una reconstrucción exhaustiva de los argumentos que se han puesto en juego en esta controversia. No obstante, sí señalaré que Dennett reconoce que algunos de los puntos indicados por la ofensiva son válidos (como, por ejemplo, el hecho de que ambos modelos dependen de supuestos difíciles de verificar). Sin embargo, dichos puntos válidos no serían motivo suficiente para abandonar las perspectivas en juego, dado que, aunque el riesgo que se corra sea alto, es mucho mayor el rédito que podemos llegar a obtener por medio de su utilización. El hecho de que tales objeciones no sean relevantes dadas sus ventajas, está vinculado a que dichas perspectivas no deben ser entendidas como teorías:

El adaptacionismo y el mentalismo (teoría del sistema intencional) no son teorías en un sentido tradicional. Son actitudes o estrategias que sirven para ordenar datos, explicar interrelaciones y generar preguntas para formularle a la Naturaleza. (Dennett, 1987, p. 235)

En otros términos, en lugar de tesis profundas acerca de cómo es el mundo natural, son buenas guías provisionales para futuras investigaciones empíricas. Dejarlas de lado nos haría tener que conformarnos exclusivamente con explicaciones causales, por ejemplo, intentando reconstruir una crónica histórico-arquitectónica de la evolución de alguna especie, sin ningún tipo de referencia a las ventajas adaptativas. Si bien este tipo de explicaciones puede ser muy valiosa, no se contrapone al tipo de explicación que apela a razones. Estas últimas le parecen muy importantes a Dennett, por los que las defiende en el ámbito de la biología (defendiendo al adaptacionismo) y en el de la psicología (defendiendo la teoría de los sistemas intencionales).

Aunque en su explicación de la evolución de la conciencia encontramos referencias a ventajas adaptativas, también encontramos señales de que Dennett es consciente de que sería erróneo llevar este supuesto de adaptación al extremo. De hecho, considera que la mayoría de los teóricos de su área también lo hacen:

Por lo que sé, la tesis de que toda propiedad, de que toda característica de lo que se encuentra en el mundo viviente es una adaptación, no es una tesis que ninguno haya tomado en serio, o haya sido sugerida por alguien que fuera tomado en serio. (Dennett, 1995, p. 453)

Por ende, una lectura caritativa nos permitiría decir que la apelación a ventajas adaptativas es para Dennett la mejor estrategia para explicar aquellos rasgos con cierto grado de complejidad elevada, cuyo diseño como resultado azaroso es altamente improbable en términos estadísticos. Del mismo modo que elaborar hipótesis científicas y proceder experimentalmente para su contrastación no implica un desconocimiento del problema de la inducción, la heurística evolucionista dennettiana no tiene porqué implicar un compromiso con el “planglossianismo”: “nunca es un error utilizar los adaptacionistas «porqués» aunque la respuesta verdadera sea que no hay razón para ello” (Dennett, 1995, p. 452). En definitiva, tomar la pregunta con sustento evolutivo: ¿qué puede tener de bueno? como vía para intentar responder a la pregunta ¿cómo funciona? es, para Dennett, una estrategia heurística y no un análisis basado en la convicción de que el proceso de selección natural siempre ha dado por resultado organismos dotados de un funcionamiento óptimo.

La evolución de la conciencia y el “yo”

Fiel a su naturalismo, Dennett abordará el interrogante acerca de qué tipo de cosa es el “yo” desde una dimensión histórica, en el sentido más amplio del término. Esta explicación supone que su *explanandum* -el “yo”- existe (como intentamos exponer en el capítulo 3), pero, a diferencia de la mayoría de las explicaciones filosóficas, la explicación que Dennett elabora en este caso toma en cuenta que, si bien ahora los “yoes” existen, hubo un tiempo, hace miles de millones de años, en el que no existían (por lo menos en este planeta). Esto significa que debe haber una historia sobre cómo llegó a haber organismos con un “yo”. “A diferencia de la mayoría de explicaciones en ciencia, las explicaciones evolucionistas son, esencialmente, narraciones que nos llevan desde un tiempo en que algo no existía a un tiempo en que ese algo ya existía siguiendo una serie de pasos que la narración explica” (Dennett, 1991a, p. 186). Estas historias tienen un alto componente especulativo, muchos límites para encontrar evidencias empíricas que las avalen y, en los casos en los que reconocen la posibilidad de que un carácter pueda haber perdurado por diferentes causas (es decir, que no son pan-adaptacionistas), resultan particularmente complejas.

Por otra parte, las explicaciones evolucionistas muchas veces desafían preconceptos y hábitos que tenemos a la hora de dar explicaciones. Dennett recupera un argumento de la Enciclopedia de Stanford (1975) para ilustrar un tipo de error que cometemos a la hora de reflexionar sobre ciertos temas. El razonamiento reza así: “Todo mamífero tiene un mamífero por madre, pero sólo hay un número finito de mamíferos, así que debe haber existido un primer mamífero, lo cual contradice nuestra primera premisa, así que, y al contrario de lo que parece, ¡los mamíferos no existen!”. Las explicaciones evolucionistas, que intentan dar cuenta de fenómenos como el que permitió que los mamíferos empiecen a existir, prescinden de enunciados universales y *de todo o nada*, según los cuales siempre sufrirían problemas de circularidad como ilustra el ejemplo de la Enciclopedia. En palabras de Dennett, uno de los obstáculos que esta perspectiva tiene que afrontar es el gusto de los filósofos “absolutistas” por “las líneas bien definidas, los umbrales, las esencias y los criterios” (Dennett, 1991a, p. 431), que llevan a pensar que tiene que haber habido, por ejemplo, un primer mamífero, como también un primer “yo” humano.

Se puede realizar una comparación entre la explicación de la emergencia del “yo” y la que responde al origen de la sexualidad. La historia de esta última se podría trazar desde el momento en el que sólo había reproducción asexual, hasta la realidad actual de la compleja e intrincada sexualidad humana; incluyendo la aparición de sexos biológicos y el comienzo de la reproducción sexual en organismos basales. No hay nada erótico [*‘sexy’*] en la vida sexual de los insectos, pero podemos reconocer en sus rutinas mecánicas y aparentemente carentes de placer los orígenes de nuestro mucho más interesante mundo sexual. Del mismo modo, no hay nada egoico [*‘selfie’*] en los precursores primitivos de los “yoes” humanos, pero sobre ellos descansan las innovaciones y complicaciones particulares que ocurrieron en nuestra especie.

En el caso de la narración sobre la evolución de la conciencia y el “yo”, Dennett (1991a) propone un esbozo, que relata de manera muy esquemática una de las posibles maneras en las que pudieron haber aparecido de estos fenómenos. Su exposición comienza en los orígenes de la vida misma. Antes de que existiesen los seres vivos, podría decirse que en el mundo no había razones, dado que no había nada que tuviera intereses. Con la aparición de los primeros replicadores simples surge, por primera vez, en un mundo de causas, una dimensión teleológica. El surgimiento de entidades que pueden mostrar una conducta, al menos muy

primitiva, de evitar su disolución y descomposición, viene acompañado de la emergencia de un punto de vista desde el cual los acontecimientos del mundo pueden clasificarse como favorables, desfavorables y neutrales. Estas categorías quedan determinadas por las tendencias de dichos organismos a buscar las primeras, alejarse de las segundas e ignorar las terceras. Cuando una criatura tiene intereses (no en todas las posibles acepciones de esta palabra, pero sí al menos desde nuestra perspectiva y en cierto sentido), el mundo y sus acontecimientos empiezan a constituir *razones* para ella. Esta afirmación puede hacerse contemplando que las criaturas podrían no reconocer dichas razones como tales desde el comienzo; antes bien, es de suponer que el primer desafío consistiría en aprender a reconocerlas y actuar de acuerdo con las mismas. La tarea de auto-conservación lleva, por una cuestión básica de economía energética, a que se trace una primer frontera: un límite entre el “yo” y el mundo, distinción que está en la base de cualquier proceso biológico. Obviamente, en este momento del proceso no se está hablando del ‘Yo’ que origina nuestra pregunta, el que nos hace ser quienes somos; pero esta es la palabra que Dennett utiliza, de manera metafórica si se quiere, para referirse a esa forma primordial de egoísmo. Tal distinción entre un interior y un exterior se puede encontrar en procesos biológicos como la digestión y la excreción, pero también en procesos como el llevado a cabo por el sistema inmunológico, donde hay millones de anticuerpos destinados a resolver el problema del auto-reconocimiento. Como es sabido, el funcionamiento de los anticuerpos no precisa de un control de mando centralizado, desde donde se dirijan todos los procesos. Sin embargo, esto no impide que estos puedan identificar agentes patógenos de tipo completamente diferente y llevar a cabo los mecanismos de defensa pertinentes para cada caso. La metáfora elegida por Dennett para ilustrar esto es la de que los anticuerpos que conforman el sistema inmunológico representan a sus enemigos del modo en que un montón de candados representan a las llaves que las abren. Así, el sistema inmunológico constituye un buen caso de cómo los actos de reconocimiento pueden ser puestos en práctica por una miríada de rutinas mecánicas y ciegas, sin necesidad de que haya un alto mando o cuartel general.

Esta distinción básica de todo proceso biológico está lejos de definir fronteras inmóviles y absolutas; antes bien, se trata de fronteras permeables y flexibles. Basta con pensar que inclusive dentro de los cuerpos humanos habitan miles de organismos diferentes. Algunos de ellos (como las bacterias del sistema digestivo)

son tan imprescindibles para nuestra supervivencia como los mismos órganos de nuestro cuerpo. Hablamos, entonces, de límites porosos e indefinidos, otro ejemplo del criterio de eficacia utilizado por el proceso evolutivo.

En un artículo temprano, Dennett (1989) llama 'Yo mínimo' (p. 176) a la distinción entre un adentro y un afuera en un sentido biológico. Dicha distinción no es algo dentro del organismo, ni todo el organismo; más bien es una abstracción: un principio de organización que tiende a distinguir, controlar y preservar determinadas porciones del mundo, es decir, que crea y mantiene fronteras. En nosotros mismos encontramos distintos tipos de pruebas de este principio de organización. Algunos experimentos psicológicos dan cuenta de una especie de mandato innato que nos llevaría a regular nuestras conductas de acuerdo con un principio de este tipo. Si se nos pide que traguemos en este momento la saliva que tenemos en la boca, podemos hacerlo sin problemas, pero si se nos pide que escupamos en un vaso limpio y luego nos traguemos lo que escupimos, la idea nos produce asco. Así, pareciera que tendemos a tratar de un modo diferente a las sustancias cuando están dentro de lo que consideramos nuestro cuerpo que cuando han salido de él.

En *La Conciencia Explicada* (1991a), Dennett añade otro factor que no es considerado en los artículos tempranos sobre el origen del "yo", como lo es "The Origin of Selves" (Dennett, 1989). Se trata de un método que complementa este principio de organización que permite responder a la pregunta acerca de qué cosa del mundo soy yo. Para ilustrarlo, el filósofo recurre a la anécdota en la que las autoridades del puerto de Nueva York habían decidido poner un radar común dirigido a los propietarios de pequeñas embarcaciones. Si un bote estaba perdido en la niebla, podía mirar su televisor y saber que era uno de los puntos en movimiento de la pantalla. El problema de saber cuál de todos esos puntos representaba el propio barco se podía resolver con la sencilla estrategia de hacerlo girar en círculo rápidamente, y ver cuál de los puntos realizaba en la pantalla un movimiento en forma de 'O'. Como podemos imaginar, el método no es infalible, pero sí funciona la gran mayoría de las veces. Así, cuando un organismo se encuentra en algún dilema o momento de escepticismo sobre si cierta información de entrada es sobre sí mismo, emplea la técnica de "hacer algo y ver qué es lo que se mueve" (Dennett, 1991a, p. 438). En términos generales, podríamos hablar de tales movimientos como actos de autorreconocimiento básicos, necesarios y automáticos.

Millones de años después de la aparición en la tierra de la vida biológica, nos encontramos en nuestro planeta con una de las más grandiosas creaciones de la Madre Naturaleza, diseñada para guiar el comportamiento interesado controlando las propias actividades en el tiempo y el espacio. Se trata de los sistemas nerviosos, con sus distintos grados de complejidad. Todos los cerebros son, de algún modo, máquinas de anticipación. Dentro de esta caracterización general, se podrían identificar conductas más o menos primitivas. Por ejemplo, hay organismos que *tocan* elementos externos y en función de eso retroceden o engullen. Los sistemas nerviosos más primitivos no tienen la posibilidad de manejarse con mensajes "objetivos" o meramente informativos. Al mismo tiempo, tampoco pueden sacar mucho partido del mundo. Sólo son capaces de una conducta apropiada para lo que se va a producir en un futuro inmediato. Los cerebros más complejos extraen información más rápidamente, y la usan para evitar todo contacto nocivo desde un principio, para buscar alimentos o para aparearse. Frente a la tarea de extraer futuro útil, constituye un ahorro de trabajo para un organismo el descubrir las leyes del mundo (o algo que se aproxime a ellas). Existe una gran variedad de sistemas con capacidad de previsión. Los más primitivos de todos serían los implementados por los organismos para permitirle al mundo advertirles cuando están haciendo algo mal; un paso siguiente a este sería el de la anticipación de corto alcance, como la capacidad de esquivar obstáculos.

Un ejemplo de un "anticipador" un poco más complejo lo constituye la sensibilidad a patrones con un eje de simetría vertical presente en los sistemas visuales de muchos animales. En los entornos *naturales* lo único que responde a este patrón es otro animal cuando está mirando de frente –otra vez, no se trataría de sistemas que se caracterizan por su veracidad y precisión, sino más bien por su economía-. Este es uno de los anticipadores que requieren una ulterior respuesta por parte del organismo en el que se activa, cuyo acierto puede determinar la supervivencia. Como se comentó en el capítulo 1, hay ciertos peces en los que la detección de simetrías verticales los lleva a provocar una *respuesta de orientación*: una suspensión del resto de las actividades en curso, para concentrarse en un hecho puntual. Los animales desarrollamos muchas de nuestras actividades en piloto automático; este tipo de alarmas haría que salgamos de ese estado y nos enfrentemos a la emergencia:

El animal (...) lleva a cabo una rápida exploración y se concede a cada órgano de los sentidos la oportunidad de contribuir al conjunto de información disponible y relevante. Se establece un centro de control temporal a través de un aumento de la actividad neuronal (...) Si el resultado de esta puesta al día es que se dispara una «segunda alarma», entonces se moviliza todo el cuerpo del animal con una descarga de adrenalina (Dennett, 1991a, p. 193).

Si no es así, el centro de control se desarma y cada cual vuelve a su especialidad.

Hay quien dice que estos estados se corresponden con nuestros estados de atención consciente. Dennett piensa que esto no es así³³, pero que sí es verdad que estos episodios de interrupción e intensa vigilancia son los necesarios precursores de nuestros estados conscientes: respuestas que se mostraron tan útiles al provocar un panorama informativo global actualizado, que los animales empezaron a entrar en el modo de *respuesta de orientación* cada vez con más frecuencia. Se podría decir que ese mecanismo desencadenado, en principio, por factores externos, empezó a ser provocado internamente, como un hábito. La adopción del mecanismo de emergencia en momentos que excedían dicha necesidad, llevó desde un hábito de vigilancia, a uno de exploración. Esta estrategia fue adoptada por la mayoría de los mamíferos, en especial los primates. Así surgen lo que Miller (1983) ha dado en llamar los *informívoros*: organismos que tienen interés en información que no es inmediatamente útil, tanto sobre el mundo que habitan, como sobre sí mismos. Es preciso señalar que la hipótesis expuesta tiene claramente la forma de lo que Gould y Vbra (1982) reconocen como una *exaptación*: los mecanismos iniciales fueron creados por Selección Natural para determinados objetivos, y luego fueron cooptados para objetivos diferentes. Esto se aplica tanto al mecanismo de *respuesta de orientación*, como a las conexiones nerviosas primitivas que obtenían información inmediata del entorno para producir respuestas de anticipación proximal.

La hipótesis que hemos presentado muy resumidamente, tiene implicaciones interesantes: un vestigio de ese origen evolutivo sería el hecho de que la recolección “desinteresada” de información realizada por los organismos superiores está siempre

³³ Como veremos más adelante, una de las diferencias entre estos estados y nuestros estados conscientes es que en estos últimos el modelo del *pandemónium* (según el cual muchos “demonios” compiten en paralelo por la hegemonía), presente para explicar el modo en el que se procede en la *respuesta de orientación*, no sería suficiente para explicar el tipo de cadenas de razonamiento que guían el flujo de nuestra consciencia.

matizada por elementos emocionales afectivos, dado que los mecanismos encargados de recoger la información del medio fueron diseñados para hacerlo agregando siempre un elemento valorativo. Cuando Dennett habla de las “huellas” de la historia de nuestra conciencia y señala que nuestro mecanismo exaptado de búsqueda desinteresada de información tiene como “reporteros” mecanismos que fueron creados para detectar si algo era “bueno” o “malo”, brinda una perspectiva que recoge la tensión presente en nuestra forma de vincularnos cognitivamente con el entorno: una forma que por momentos excede la utilidad inmediata, pero nunca escapa del todo de ciertos matices emocionales y afectivos. Es importante aclarar que para Dennett todo este proceso se estaría dando dentro de lo que desde un marco teórico evolucionista debe entenderse como el plano de la selección de genotipos. Sin embargo, después introduce otros dos medios evolutivos que habrían funcionado en simultáneo con el proceso descrito: la plasticidad fenotípica y la evolución cultural con la transmisión de sus productos.

Para comprender otra de las líneas trazadas por Dennett (que deben ser entendidas como sólo un hilo del entramado multidimensional que es la historia evolutiva), al relatar el pasaje del cerebro de los simios al de los humanos, es necesario familiarizarnos con el proceso de evolución por plasticidad fenotípica. Los diseños de la Madre Naturaleza están adaptados en función de la idea de que muchos elementos del futuro se comportarán igual que en el pasado: la fuerza de la gravedad, los cambios de estado de agregación del agua, etc. Sin embargo, muchas cosas del medio natural no guardan este tipo de previsibilidad. En ese caso, más que los diseños estereotipados son efectivos aquellos que pueden rediseñarse a sí mismos, en mayor o menor grado, a fin de adaptarse a las condiciones que encuentren. Este rediseño se denomina a veces aprendizaje y a veces desarrollo. Es difícil fijar esta frontera, por lo que Dennett le llama al conjunto que abarca a ambos *fijación postnatal del diseño*, haciendo referencia a la fijación de caracteres que quedan abiertos a una posibilidad de variación luego del nacimiento.

La fijación postnatal del diseño adopta, en muchos casos, la forma de un proceso de selección dentro del individuo, análogo a la selección natural. Para Dennett el cerebro es uno de los órganos en dónde se sitúan los más relevantes procesos evolutivos de nuestra especie (Dennett, 1974b). Este órgano plástico es capaz de reorganizarse a sí mismo como respuesta a las novedades que el organismo

va encontrando en su entorno. La forma en la que presumiblemente lo hace es contando con algún tipo de selector mecánico y con las cosas dispuestas de forma tal que se presenten candidatos a dicha selección, para poder escoger la más adaptativa. (Cfr: Dennett 1991a, pp. 196, 197).

Esto da una ventaja a cada organismo en particular, pero también realimenta el proceso de evolución genética y lo acelera; fenómeno conocido bajo el nombre de “Efecto Baldwin” (uno de los ejemplos paradigmáticos de las *grúas* de las que hablábamos al comienzo: en las pp. 14-16 de este trabajo). Las especies, a través de la exploración fenotípica individual, pueden evaluar distintos diseños más allá de aquellos que traen pre-configurados al momento de su nacimiento. Si en esa exploración descubren una configuración ventajosa, crean una nueva presión selectiva: los organismos que estén más cerca de la nueva configuración estarán en ventaja con respecto a los que están más lejos. Así las especies con plasticidad evolucionarían más rápidamente y de forma más eficiente.³⁴

La evolución de los primates, en la que se desarrollaron importantes modificaciones de sus cerebros, junto con capacidades de movimiento mucho mayores a las de otras especies y una inclinación a la exploración, ha dado como resultado el ingreso de una gran cantidad de información multimedia a estos cerebros. El ingreso de estas grandes cantidades de información multimedia ha llevado a otro desafío evolutivo: la necesidad de un control de alto nivel. Mientras que organismos más primitivos podían dirigir su conducta buscando equilibrios

³⁴ El *Efecto Baldwin* puede ser mejor ilustrado mediante la siguiente imagen. Supongamos que hay un buen truco, que cambia mucho las ventajas adaptativas. Si ese buen truco sólo lo posee una cantidad muy reducida dentro de una población, las posibilidades de que sea transmitido son casi nulas, por eso se dice que sería “invisible” para la Selección Natural. La plasticidad otorga cierto margen de exploración a cada organismo, de forma tal que ese truco podría ser alcanzado por varios, con mucha más frecuencia en aquellos que nacen por herencia genética con una predisposición cuyas posibilidades de exploración hagan muy factible que el organismo se encuentre rápidamente con el buen truco. Estas nuevas dimensiones (muchos más organismos habrían podido encontrarse con el buen truco) harían que sí sea visible para la selección natural, siendo seleccionadas aquellas configuraciones cercanas al truco, por lo que este se podría transmitir al resto de la población. Para Dennett, debemos agradecer al efecto Baldwin el hecho de que actualmente nacemos con una alta capacidad para aprender un lenguaje. Es menester aclarar que no se identifica con la idea lamarkiana de transmisión genética de los caracteres adquiridos. Se trata de que los organismos que genéticamente están determinados a nacer más cerca en el espacio de exploración de diseño de ese buen truco aprendible tendrán mayor descendencia, que a su vez tenderá a estar más cerca del buen truco. Después de mucho tiempo la mayoría de la población nacerá con el buen truco o muy cerca de él (Cfr. : Dennett, 1991a, pp. 196–200 y Dennett, 1995, pp. 117-122).

sencillos entre un conjunto limitado de acciones, estos organismos que cuentan con una gran cantidad de información disponible para los distintos sistemas encargados de procesarla, ya no tienen sólo el problema de “qué hacer después”, sino también el de “en qué pensar después”. Una vez que se cuenta con subrutinas como la de la respuesta de orientación, que permite centralizar información proveniente de distintas vías, surge el problema de cómo organizar la confluencia de las mismas, en base a qué direccionar la toma de decisiones y cómo encadenar el pensamiento.

En este punto Dennett introduce el modelo del pandemónium -planteado por primera vez por Selfridge (1959) como una arquitectura para explicar la inteligencia artificial-, que se apoya en la metáfora según la cual muchos “demonios” compiten en paralelo por la hegemonía. Como una herramienta teórica resulta útil para pensar el resultado de la organización de las distintas opciones o respuestas que puede dar un organismo complejo en aquellas ocasiones en las que su mecanismo de alarma es desatado por alguna emergencia ambiental. Es decir, una vez que el entorno lleva al organismo a entrar en estado de alarma, las distintas respuestas posibles a la situación de emergencia “compiten” entre sí, hasta que una de ellas “triumfa” sobre las demás y es llevada a cabo. Sin embargo, este modelo sigue siendo muy dependiente de las características del entorno, por lo que sólo funciona en sistemas de previsión limitada. En cambio, sería de esperar que los organismos que desarrollaron el hábito de mantener esas respuestas de orientación de manera continua, y no sólo con dependencia del entorno, hayan desarrollado una manera de resolver de forma endógena el *metaproblema de en qué pensar después*. Dennett recurre una y otra vez a esta críptica frase al explicar el origen evolutivo de la conciencia. Según la explicación dennettiana de dicho origen evolutivo, los antecedentes primitivos de la conciencia eran despertados por un estado de emergencia en el cual se integraba información proveniente de los distintos sentidos con el fin de emplearla para lidiar con el peligro potencial de la situación. La gran innovación evolutiva consistió en adquirir la capacidad de que dicho estado de integración de información no se active únicamente en función de un peligro externo, sino que pueda activarse por estímulos internos. En función de esto, podemos decir que el *metaproblema de en qué pensar después* consiste en encontrar una forma de concatenar dicha información que sea producida endógenamente. Este mecanismo sería el responsable de que tengamos algo parecido a un flujo de la conciencia, que sea capaz de sostener el tipo de cadenas de pensamiento que mantenemos los seres

humanos y que resulte tan distinto del que suponemos que puede tener un pez, por ejemplo.

Para explicar cómo se puede haber solucionado este problema, pueden ser útiles algunas consideraciones en torno las características que tiene el cerebro humano actual como resultado del proceso evolutivo. El cerebro del primer *Homo Sapiens* (desde hace 150.000 años hasta hace sólo 10.000 años) es prácticamente igual al nuestro en su estructura. La gran encefalización se habría producido mucho antes del desarrollo del lenguaje, y de los desarrollos que se considera que atestiguan la expansión de los poderes mentales humanos: la cocina, la agricultura, el arte. Estas adquisiciones son extremadamente recientes: podemos incluirlas en un periodo que comienza en los últimos 10.000 años. Dado que la aparición del lenguaje no se identifica con el crecimiento del cerebro, los lingüistas acuerdan en entender al mismo como un añadido muy reciente, fruto de la explotación de circuitos secuenciales anteriores, acelerada por el efecto Baldwin. En términos generales, podemos pensar que "el impresionante avance del *Homo Sapiens* en los últimos 10.000 años tiene que ser debido casi por entero al aprovechamiento de la plasticidad de ese cerebro de formas radicalmente originales: creando algo parecido a un software para potenciar sus facultades subyacentes" (Dennett, 1991a, p. 204). Nuevamente, estamos ante una explicación que contempla la posibilidad de que características que existían desde hace tiempo, hayan sido cooptadas para nuevas funciones.

Todo el desarrollo y la complejidad que pueden llegar a adquirir las capacidades representacionales de los cerebros humanos sólo puede deberse a una combinación de tendencias innatas, junto con elementos aprendidos³⁵. La plasticidad

³⁵ Si bien Dennett no se dedica en estos textos a desarrollar el problema de la representación, podemos decir que sostiene una concepción gradual de la misma, de acuerdo con la cual habría distintos niveles de representación. El más básico podría ser el que tienen lugar, por ejemplo, cuando una cebra, que se percata de la presencia de un león, no se olvida de donde está este depredador aún cuando no puede seguir viéndolo porque su trayectoria quedó oculta por unos matorrales. Por contraposición, el girasol - que carece de representaciones - no sigue la trayectoria del sol cuando la nube tapa al sol. Así, las representaciones estarían asociadas al hecho de que los patrones de actividad cerebral que siguen una pista sean capaces de continuar dicho seguimiento, incluso cuando se interrumpe la relación causal con el referente en cuestión. Ahora bien, en los humanos la capacidad de representar se habría disparado en complejidad de una manera muy pronunciada. En el caso de los humanos, algunas categorías relevantes para la supervivencia deben estar dadas innatamente, y otras deben ser aprendidas. Con respecto a tesis como las de Fodor (1975) según la cual todos los conceptos son innatos, Dennett sostiene que dicha formulación es claramente falsa en sus

hace posible el aprendizaje, pero eso por sí sólo no sería suficiente si no fuera por aquello que tenemos en nuestro entorno para aprender que es, a su vez, el resultado de un proceso de diseño previo que evita que tengamos que reinventar la rueda cada vez: los objetos culturales. Así, el tercer medio evolutivo, que guarda relación con la plasticidad fenotípica, del mismo modo en que ésta última guarda relación con la variación genética, es el de la evolución cultural y la transmisión de sus productos. Elaboraré aquí una primera exposición sobre esta última, que volveré a retomar en el próximo capítulo.

Las consideraciones de Dennett sobre la evolución cultural como uno de los medios para explicar la evolución de la conciencia (desarrollados principalmente en Dennett, 1991a) sirven para responder a la pregunta formulada en *Why Everyone is a Novelist* (1988), mencionada al comienzo de este capítulo: ¿de qué nos podría servir hablar con nosotros mismos? Esta pregunta no podría ser abordada si no se hubieran señalado antes algunas consideraciones sobre el modelo de la conciencia de Dennett. Es bajo el supuesto de que los módulos del cerebro tienen diferentes capacidades y formas de hacer las cosas, y de que no son perfectamente inter-accesibles, que la explicación que desarrollaremos a continuación acerca de la invención de determinados hábitos de auto-estimulación, cobra sentido.

En la reconstrucción racional de esta parte de la historia evolutiva -que aparece en *La conciencia explicada* (1991a), pero también en textos centrales sobre su idea del “yo” narrativo, como *Why everyone is a novelist* (1988)- Dennett se remite a aquel momento en que nuestros antepasados habrían desarrollado algún tipo de acto de vocalización. En un momento inicial, tales vocalizaciones no deberían considerarse como actos de habla (en los que la intención del emisor de producir un efecto en la audiencia depende de la capacidad de esa audiencia de apreciar dicha intención), pero sí como actos comunicativos en los que los primates serían capaces de distinguir entre emisores y audiencias diferentes, tomando en cuenta creencias y deseos.

A medida que se fueron desarrollando, estos actos comunicativos habrían empezado a incluir el pedido de ayuda y, en particular, el pedido de información.

versiones más radicales (por ejemplo, una que sostenga la idea de que nacemos con el concepto de *iphone* 7), y difícil de distinguir de la tesis opuesta, en sus versiones más moderadas.(Cfr: Dennett, 1991a, pp. 204–206).

Para que la práctica quedase arraigada en la comunidad, cada participante debió estar dispuesto a responder de un modo apropiado cuando le tocase. Así, en algún punto de la evolución del lenguaje, las vocalizaciones habrían adquirido la función de solicitar y compartir información que resultaba útil.³⁶ Imaginemos que, en algún momento, uno de estos homínidos cometió el error de pedir ayuda cuando no había nadie cerca para poder responder a su solicitud. En el relato dennettiano, el oír su propia petición provocó en él el tipo de estimulación que provocaba el mismo pedido cuando provenía de otro homínido. A su vez, bien podría haber terminado por encontrar la respuesta que precisaba, gracias a ponerse en la situación de quien responde la pregunta. La sugerencia de fondo de esta propuesta es que hacerse preguntas a sí mismo es un efecto secundario de la práctica de hacer preguntas a los otros. Esta práctica surgiría de forma natural y tendría una utilidad parecida a la de su predecesora: permitir al individuo realizar un control más eficaz de sus acciones.

La suposición de fondo es que cierta información ya está en el cerebro, sólo que en una parte distinta de aquella parte que necesita esa información. La evolución no habría llegado a establecer una conexión más estrecha entre los distintos subsistemas del cerebro, pero esta comunicación se podría establecer a partir de determinadas acciones. El acto de proferir una pregunta y respondérsela a uno mismo, podría generar una conexión “virtual” entre esos subsistemas que no se encontraban conectados. Decir algo y escucharse decirlo pudo llevar a encontrar la respuesta que buscábamos. La práctica de la auto-estimulación abre, así, una vía entre esos componentes internos del individuo. Este hábito de auto-estimulación empezaría a establecerse en las conductas de las poblaciones humanas como un buen truco. Por un lado, en los hábitos comportamentales aprendidos en la población, por otro, en las predisposiciones genéticas y mejoras posteriores de eficiencia y efectividad, gracias al efecto Baldwin.

Se puede pensar que el reconocimiento de las ventajas de hablar a solas con uno mismo, terminará conduciendo a un discurso silencioso. Con esto, se mantendría el bucle de auto-estimulación, desechándose las partes de vocalización y audición que no cumplían un rol tan importante. Este cambio podría acarrear el

³⁶ Dennett no se compromete con la idea de que un instinto innato de cooperación haya sido seleccionado evolutivamente, como afirman otros teóricos acerca de la evolución del lenguaje y de la cognición humana; antes bien, puede pensarse que los costos y beneficios de participar en dicha práctica podrían ser considerados de forma individual, y haber sido seleccionados por considerarse provechosos.

beneficio adicional de contar con cierta privacidad durante esta práctica. Dennett sostiene que este relato puede dar cuenta de cómo pudo haberse producido la modificación de la arquitectura funcional del cerebro, a partir de una conducta de hablar privadamente con uno mismo. Al surgir de la práctica de hablar con otros, la conversación privada heredaría ciertas características de aquella: sería lineal (un tema por vez) y dependería de las categorías informativas encarnadas por las acciones que ésta explotaba (no habría cosas que los homínidos podrían decirse a sí mismos pero no a los otros).³⁷ Incluso antes del planteo de Dennett, muchos pensadores (Vigotsky, 1934) han defendido la idea de que el pensamiento consciente es una actividad que puede ser entendida como una manera eficiente de hablar con uno mismo en privado. A partir de esto, es fácil conjeturar que la transición evolutiva al pensamiento se dio primero hablando con otros, luego en privado, y luego descubriendo que la vocalización y la audición eran prescindibles, y que a veces era ventajoso evitar que otras personas escuchen.³⁸ Hablar con uno mismo es sólo un ejemplo de las posibles formas en las que los seres humanos somos capaces de inventar vías de comunicación interna. Otro acto de auto-estimulación en el que se puede pensar es el de hacerse dibujos. Hacer dibujos para pensar algo puede haber traído determinadas ventajas, que habrían llevado a la forma refinada de “hacer dibujos *privados* «en el ojo de la mente»” (Dennett 1991a, p. 210). No necesitamos imaginar ejemplos de cómo esto puede haber ayudado; lo sabemos por la experiencia propia de realizar gráficos para resolver un problema, o mapas para trazar un recorrido.

Esta explicación brindada en *La conciencia explicada*, puede leerse como una respuesta elaborada a una inquietud planteada por Dennett en *Why Everyone is a Novelist* (1988) y *The Self as the Center of Narrative Gravity* (1992) en torno a las investigaciones realizadas por Gazzaniga (1978) sobre los sujetos con el cerebro dividido. En estos casos, la extracción del cuerpo calloso genera una división entre los dos hemisferios, que pueden tener informes muy diferentes de lo que está pasando. Lo curioso de este fenómeno es que, lejos de verse muy afectados por la

³⁷ Para una exposición más completa de cómo podría ser para Dennett la arquitectura de la mente, véase el esbozo planteado en Dennett, 1991a, cap: “La arquitectura de la mente humana”.

³⁸ Sin caer en el rebatido supuesto de que la ontogenia necesariamente recapitula la filogenia (Gomila, 2010b), es menester señalar que un proceso de aprendizaje similar se puede observar en los niños pequeños, que luego de aprender a comunicarse verbalmente con otros, pasan muchas horas hablando solos en voz alta (Karmiloff-Smith, 1979).

separación, los pacientes se mueven de forma bastante normal, salvo en circunstancias muy particulares. Dennett y Gazzaniga comparten la lectura de que este comportamiento casi normal no debe ser interpretado como una muestra de que los pacientes hayan conservado su robusta unidad pre-quirúrgica, sino como la prueba de que la unidad de la vida normal es una ilusión (Dennett, 1988, 1992). Una anécdota bastante ilustrativa es la que surge de un experimento en el que se intenta que pacientes con el cerebro dividido identifiquen objetos dentro de una bolsa opaca en donde sólo pueden meter una mano, y los nombren en voz alta. Cuando utilizan la mano derecha, no suelen tener problemas para hacerlo, dado que esta se comunica directamente con el hemisferio izquierdo, responsable del lenguaje. Sin embargo, al utilizar la mano izquierda se encuentran con la dificultad de que pueden sujetar el objeto, pero no pueden decir su nombre. Aquí es cuando intervienen estrategias inteligentes por parte de los pacientes: en ocasiones, cuando el objeto en cuestión era un lápiz, los pacientes clavaron la punta del mismo en la palma de su mano y, dado que el dolor viaja ipsilateralmente, al hemisferio izquierdo le llega la pista de que se trata de algo punzante. Las distintas hipótesis que el hemisferio izquierdo elabora, pueden ser “escuchadas” por el derecho, el que puede enviar estímulos para rechazar o aceptar las hipótesis enviadas por el izquierdo (por ejemplo, enviar la orden de “sonreír” para aceptar la hipótesis) (Cfr: Dennett, 1991a, p. 211). Si bien éste complicado estratagema habría sido elaborado por alguien con un problema muy particular, se podría tomar como una prueba de un virtuoso talento de exploración presente en el resto de los humanos.

La idea de fondo es que estas estrategias encuentran su razón de ser en el hecho de que no hay una unificación armoniosa y perfecta de las distintas partes de la mente. Así, el Modelo de las Versiones Múltiples cobra mucha relevancia en el planteo dennettiano del origen de la conciencia. Dado que los módulos o sistemas de la mente no son accesibles entre sí al mismo tiempo, han sido necesarias rutas ingeniosas e indirectas para resolver los problemas de comunicación entre ellos. Cuando esta concepción del cerebro era apenas un esbozo, Dennett advierte que “Si esto es verdad (y yo creo que lo es), podría llegar a proveernos la respuesta a una de las más enigmáticas preguntas sobre el pensamiento consciente: ¿qué tiene de bueno?” (Dennett, 1988, p. 22, mi traducción).

Como dijimos al comienzo, incluso antes de elaborar su teoría de la conciencia, Dennett (1988, 1992) anticipa que la pregunta va a requerir de una respuesta evolutiva. Nuestra interpretación de que la apelación a la historia evolutiva debe ser considerada como una metodología heurística, y no como un análisis basado en la convicción de que el proceso de selección natural ha dado por resultado organismos con una adaptación óptima, se ve apoyada por declaraciones del filósofo en el marco de esta discusión particular:

Podríamos seguir divirtiéndonos imaginándonos otras situaciones posibles para explicar la invención de modos útiles de autoestimulación, pero ello podría ocultar el hecho de que no todas estas invenciones tendrían por qué ser útiles para sobrevivir. Una vez que el hábito general de la autoestimulación se hubo inculcado de una u otra manera, éste podría haber engendrado una serie de variaciones no funcionales, (aunque no necesariamente disfuncionales). (Dennett, 1991a, p. 211)

Y, más adelante

Existen, después de todo, muchas variedades de autoestimulación y automanipulación que seguramente no tienen ningún efecto útil sobre la cognición o sobre el control, pero que, por razones darwinianas estándar, no pueden ser eliminadas e incluso pueden derivar hacia la fijación definitiva (cultural o genética) en ciertas subpoblaciones. (Dennett 1991a, pp. 211, 212, la cursiva es mía)

Los casos, como los trabajados por Gazzaniga, en los que una lesión cerebral es superada sin reparar nada en el cerebro parecen brindar buenos argumentos a favor de la existencia de un talento humano para inventar vías de comunicación interna, a partir de estrategias de auto-estimulación, si la ocasión lo exige. Ese talento exploratorio para comunicar distintas partes de nuestra mente sería el responsable no sólo de la unidad de la conciencia, sino, en un plano posterior y característicamente narrativo, de la aparición de algo así como un Yo.

La auto-estimulación que ahora potencia la organización cognitiva es en parte innata y en parte aprendida. Habría una suerte de exploración semiconsciente de distintas técnicas de auto-estimulación cognitiva, que llevaría a desarrollar un estilo personal, con pros y contras. Si bien se acepta la idea de que puede haber personas mejores que otras en esto, la gran mayoría de estas estrategias de auto-estimulación

son transmitidas de unas personas a otras mediante la enseñanza y el aprendizaje. Así, la transmisión cultural,

puede allanar la cumbre de la colina de la idoneidad, creando una meseta que amaina la presión selectiva por transferir el truco al genoma. Si casi todo el mundo se hace lo bastante bueno como para arreglárselas en el mundo civilizado, la presión selectiva por transferir buenos trucos al genoma desaparece o se ve sustancialmente reducida (Dennett, 1991a, p. 212).

Dicho esto, queda abierto el camino para sostener que la principal forma de transmisión de este “sistema” es a través de la cultura. Veamos de qué se trata.

Según este enfoque, la cultura se ha convertido en un depósito y en un medio de transmisión de innovaciones que, entre otros procesos, forma parte del origen de la conciencia, dado que las técnicas de auto-estimulación y estimulación mutua estarían profundamente enraizadas en la educación y la cultura de nuestras sociedades actuales. Después del nacimiento, el ajuste posnatal del diseño se da siempre en relación a la cultura específica en donde el individuo se desarrolla. El proceso se produce tan rápido que debemos esperar que estén las bases genéticas para facilitar el aprendizaje del lenguaje (lo que sería por otro lado esperable, dado el efecto Baldwin).

En este momento de la evolución de la conciencia, en el cual se explica que esta no es el resultado exclusivo de la evolución biológica, sino también de una evolución cultural, es cuando Dennett introduce su teoría de la memética. El concepto de “meme” refiere a la unidad mínima de transmisión cultural, que se debe entender en analogía con la unidad de transmisión biológica: el gen. Un meme es, entonces, una idea que se difunde de una persona a otra dentro de la cultura. “Una vez que los cerebros han abierto las vías de entrada y salida para los vehículos del lenguaje, enseguida se ven atacados por parásitos, por unas entidades que han evolucionado precisamente para medrar ese espacio: los memes” (Dennett 1991a, p. 213). La idea de que podemos hablar de evolución en este campo de una manera no meramente metafórica se justifica desde una perspectiva en la que la evolución por selección natural puede entenderse como un proceso que necesita de tres componentes, con independencia de que intervengan elementos orgánicos en el mismo. Uno de ellos sería la variación, referente a la aparición continua de elementos

diferentes. El segundo refiere a la idea de que tales elementos tengan la capacidad de replicarse o heredarse. El último consiste en que posean una idoneidad diferencial; esto es, que el número de copias que de cada elemento sean creadas variará en virtud de las interacciones entre las características de ese elemento (aquello que lo hace distinto a los otros) y las características del entorno en que persiste. Con esto, se podría decir que recientemente existen en la tierra un nuevo tipo de replicadores – los ejemplos van desde ideas como la rueda, a la séptima sinfonía de Beethoven. Estas nuevas unidades estarían en su infancia, pero se evolución y desarrollo está avanzando a un paso muy rápido. Si bien es verdad que la evolución cultural sólo pudo haber tenido lugar gracias a que la evolución genética creó una especie con un cerebro que puede dar cobijo a los memes y con hábitos comunicativos que pueden transmitirlos, evolución genética y cultural son procesos que deben tomarse en ambos casos en sentido literal y que avanzan por medios distintos y a ritmos distintos.³⁹

Una buena comprensión del proceso de Selección Natural enseña que la replicación de los elementos heredables no se produce por un objetivo en particular (en biología se dice que la variación no está dirigida al *fitness*). Así, la idea de meme excede observaciones simples acerca de cómo los elementos de nuestra cultura nos afectan y se afectan entre sí, e incluye la realidad de que podría no existir ninguna conexión necesaria entre el poder replicativo del meme y nuestra conveniencia. Un meme, como cualquier replicador en la evolución, podría haber evolucionado de la manera en la que lo ha hecho simplemente porque es un buen replicador. Aquí encontramos otro motivo por el cual resulta inadecuado pensar que Dennett es pan-adaptacionista: al incorporar rasgos de la selección cultural, lo hace con absoluta conciencia de que el éxito de los mismos no guarda una relación estrecha con la ventaja adaptativa que pueda conferirnos. Evidentemente este hecho tiene un límite: un meme que lleva al suicidio a quien lo acarrea tiene muchas más probabilidades de extinguirse. Y, si bien hay memes que se replican que son inútiles o perniciosos, hay

³⁹ En discusiones actuales sobre evolución cultural, la memética es una teoría dentro de un paraguas más amplio que incluye a aquellas que entienden que lo que motoriza el avance de la evolución es una selección de rasgos culturales. Este conjunto de teorías se diferencia de la postura según la cual la evolución cultural debe ser entendida como un proceso de transformación en el que los individuos recrean los rasgos culturales cada vez que los transmiten (Sperber, 1996, Sterelny, 2006). Recientemente, Dennett (2006) ha señalado que dichos procesos no serían incompatibles, y que cada uno de ellos involucra capacidades cognitivas específicas, por lo que el reconocimiento del segundo tipo de procesos no invalidaría la teoría memética.

muchos que lo hacen con nuestra estima y, a veces, incluso a causa de nuestra estima. Los memes que son buenos replicadores y los memes que son buenos para nosotros, no se identifican, pero tampoco están completamente desvinculados: "en gran medida, los buenos memes son también los que son buenos replicadores" (Dennett 1991a, p. 218). Si los memes que eligiéramos fueran completamente azarosos con respecto a la ayuda que pueden brindarnos, no sobreviviríamos. Por esto, aunque Dennett hace hincapié en que no debemos identificar los memes ventajosos para nosotros con sus capacidades de replicación, tales aclaraciones dejan ver que los efectos que tienen sobre los organismos que funcionan como sus "soportes biológicos" (los seres humanos) también tienen algún tipo de incidencia en su replicación; en palabras de Brook y Ross (2002), la perdurabilidad de los memes puede entenderse como una relación entre contenido, ambiente y predisposición biológica.

Dado que las mentes humanas son un recurso limitado, con espacio limitado, la competencia por entrar a las mentes humanas es la principal presión selectiva de la memosfera. Dennett brinda algunos ejemplos de memes que serían buenos replicadores, aunque muchos de ellos no sean buenos para nosotros, como la fe, las teorías conspirativas, la libertad de expresión. Lo que tienen en común es que su efecto fenotípico tiende a desactivar las presiones selectivas dirigidas contra ellos (por ejemplo, los argumentos racionales en contra de una creencia religiosa sólo fortalecen la idea de que la fe es el camino que hay que tomar para comprenderla, o los argumentos en contra de una teoría conspirativa pueden leerse como un signo de que la conspiración es tan poderosa como para depositar falsas pistas). Por esto es que Dennett advierte que una de las principales fuentes de resistencia a esta idea proviene de que resulta desagradable pensar que los memes que han "invadido" nuestra mente son, en parte, los responsables de que sostengamos una idea.

La evolución de los memes posee el potencial de producir, con gran rapidez, sustanciales mejoras de diseño en los mecanismos subyacentes del cerebro, en comparación con la lenta evolución genética. En algún punto, esto permite responder a la inquietud que producen las explicaciones que apelan de modo exclusivo a la evolución genética: cómo se produce tan rápido la inclusión de nuevas características. Según Dennett, es la constatación de que la evolución muchas veces avanza más rápido que lo que lo haría si dependiera exclusivamente del proceso de

selección natural de genotipos lo que habría dado lugar a la falsa hipótesis de la herencia genética de los caracteres adquiridos. Su respuesta a la inquietud es que, en primer lugar, el efecto Baldwin acelera la evolución, dado que favorece el movimiento de buenos trucos que son descubiertos individualmente hacia el genoma. El modo en el que lo hace es, como ya explicamos antes, por la vía indirecta de crear nuevas presiones selectivas gracias a la amplia adopción de esos buenos trucos por parte de muchos individuos. Sin embargo, la evolución cultural funciona mucho más rápido y además permite que los individuos adquieran buenos trucos por parte de sus antepasados que ni siquiera son sus antepasados genéticos. Tanto es así, que la evolución cultural habría acabado con la Presión Selectiva fruto del efecto Baldwin: "Es probable que las mejoras del diseño que uno recibe de la propia cultura limen gran parte de las diferencias genéticas en el diseño del cerebro, eliminando así las ventajas de aquellos que eran un poco mejores en el momento de nacer" (Dennett, 1991a, p. 221). Con estas consideraciones acerca del papel que desempeña la evolución cultural podemos afirmar también que Dennett no incurre en otro de los errores que han sido atribuidos a las explicaciones evolucionistas para explicar capacidades humanas: el de entender que en la actualidad las habilidades humanas deben ser pensadas enteramente a partir de una competencia por la supervivencia entre nuestros antecesores no tan lejanos y una selección de los mejores caracteres genéticos. En definitiva, la velocidad del proceso de la evolución humana sería explicada por Dennett apelando a que la misma es el resultado de un proceso de selección de genotipos, acelerado por el efecto Baldwin (posible por la plasticidad fenotípica), y por el proceso de evolución cultural (que sería actualmente responsable de que las potenciales "ventajas" del diseño cerebral a la hora de nacer sean irrelevantes).

La evolución genética, la plasticidad fenotípica y la evolución cultural son tres medios evolutivos que han contribuido sucesivamente al diseño de la conciencia humana a velocidades cada vez mayores. La evolución cultural es un fenómeno extremadamente reciente, convertido en fuerza poderosa sólo en los últimos cien mil años, que ha estallado con el desarrollo de la civilización hace menos de diez mil años, y que sólo se da en nuestra especie.

Para entender cómo se ha dado esta evolución cultural, Dennett propone entender a los memes de la conciencia como una máquina virtual que se ha instalado

en cada uno de nosotros. Uno de los supuestos de su historia especulativa es que nuestros antepasados sentían el mismo placer que sentimos nosotros con los métodos de autoexploración: habría una especie de curiosidad e inquietud innatas que lleva a que investiguemos y experimentemos con el entorno y con nosotros mismos. A partir de ello, se vuelve plausible que hayamos dado con estrategias de auto-manipulación y auto-estimulación que llevaron a la inculcación de hábitos y disposiciones que alteraron la estructura comunicativa interna de nuestros cerebros. Estos descubrimientos se habrían vuelto parte de nuestra cultura (serían memes) y estarían disponibles para todos.

Si bien estas diferencias funcionales deben tener un correlato físico microscópico, al menos hasta ahora son (y quizás sigan siendo) invisibles para los científicos del cerebro. Por ende, para hablar de la arquitectura funcional creada por la invasión de memas, es necesario recurrir a un nivel más alto de descripción: el nivel de descripción del software, que se toma “prestado” de la informática y las ciencias de la computación.

Uno de los conceptos centrales al que Dennett acudirá en el nivel informático es el de la máquina de von Neumann. La máquina de Turing, formalización en la que se asienta la máquina de von Neumann, surge del intento de identificar qué es lo que de hecho hacemos cuando calculamos, es decir que está inspirada en una de las funciones de nuestras mentes. Estos diseños son los precursores de las actuales computadoras digitales, que característicamente poseen un procesamiento lineal. A diferencia de cómo la prensa difundió la noticia cuando las máquinas de von Neumann fueron creadas, más que "cerebros electrónicos gigantes", deben entenderse como "mentes electrónicas gigantes". El cerebro, en cambio, funciona con un procesamiento paralelo, de millones de canales actuando al mismo tiempo. Sin embargo, nuestro acceso introspectivo a nuestro modo de razonar no nos familiariza con procesos ocurriendo al mismo tiempo por canales diferentes; en su lugar, una vez depurado el razonamiento da como resultado algo similar a lo que obtuvo Turing al depurar el suyo: un esquema lineal como el que sustenta las máquinas de von Neumann. Partiendo de estas consideraciones, el problema a resolver es cómo puede ocurrir un fenómeno serial joyceano -nombre que emplea Dennett en honor a las descripciones que James Joyce realiza en sus novelas sobre cómo se da la secuencia de contenidos mentales en la conciencia- o von neumanniano, es decir, cómo puede

ocurrir un fenómeno de procesamiento lineal, a partir de la arquitectura paralela del cerebro. De hecho, en relación con esto se formula una de las principales objeciones a la inteligencia artificial: ¿cómo se podría crear una inteligencia parecida a la humana, que depende del procesamiento paralelo del cerebro, sobre máquinas de von Neumann? Se puede responder que para ello es preciso entender al cerebro como una máquina de procesamiento paralelo masivo, y bajo el principio de que cualquier máquina computadora puede ser imitada por una máquina virtual sobre la base de una máquina de von Neumann, concluir que el cerebro también podría ser imitado por una máquina de von Neumann. Aunque, como debería sustituir tiempo por espacio, necesitaría muchísimo más tiempo para funcionar. Una máquina virtual debe entenderse como “un conjunto temporal de regularidades altamente estructuradas impuesto sobre el hardware subyacente por un programa: una receta estructurada con cientos de miles de instrucciones que dotan al hardware de un enorme conjunto de hábitos interconectados y de disposiciones a reaccionar” (Dennett, 1991a, p. 229). Lo que piensa Dennett es que del mismo modo que se puede simular un cerebro paralelo sobre una máquina de von Neumann, se puede simular una máquina de von Neumann sobre un cerebro paralelo: “las mentes humanas conscientes son máquinas seriales implementadas -de forma ineficiente- sobre el hardware paralelo que la evolución nos ha legado” (Dennett, 1991a, p. 231).

La analogía con el software tal y como lo conocemos en el campo de la informática tiene algunos límites. Mientras que los programas de computadora tienen una instalación instantánea, el programa de la conciencia a instalar en el hardware del cerebro necesita de un largo periodo de formación para irse urdiendo (que incluye los procesos de auto-estimulación de los que venimos hablando). En el mismo sentido, la instalación resultará en un programa mucho más rígido en las computadoras que en el cerebro, dada la forma en la que quedan fijadas las reglas o instrucciones en cada uno de los casos. Las máquinas virtuales del flujo de la conciencia presentes en cada uno de nosotros deberían su grado de uniformidad a ciertos métodos de transmisión que son de carácter social y dependientes del contexto, pero también, en cierto punto, son auto-organizados y auto-correctivos.

Si bien existen muchas diferencias entre la conciencia humana y el software de una computadora, Dennett piensa que trazar un paralelismo entre ambas es útil porque ayuda a entender las siguientes características de la conciencia que, si no,

quedarían sin explicación. Una es que la conciencia humana es una innovación demasiado reciente como para estar preconfigurada en la maquinaria innata. Por otra parte, es en gran medida producto de la evolución cultural, que es transmitida a los individuos como parte de la formación ontogenética durante los primeros años de vida. Además, el éxito de la formación de la conciencia está determinada por un sinfín de micro-disposiciones en la plasticidad del cerebro, lo que hace que los rasgos funcionalmente relevantes sean invisibles a un examen neuroanatómico; o sea, es necesario un cambio de nivel para explicar el contenido fenoménico de la conciencia, del mismo modo que una diferencia de software nunca reducirse a las diferencias de voltaje entre computadoras.

Si los seres humanos podemos “compartir un software” (en el sentido de que tenemos instalado el mismo “programa de la conciencia”) sin que nadie tenga conocimientos muy complicados de configuraciones, debe ser porque los sistemas compartidos poseen un elevado grado de labilidad y tolerancia a los distintos formatos. Entre los métodos por los que se llegar a compartir el mismo software podemos incluir procesos como el aprendizaje por imitación, el aprendizaje como resultado de un "refuerzo" y el aprendizaje como resultado de un proceso de formación explícito. De hecho, la escritura y el dibujo también deben cumplir un rol fundamental en el tipo de programa que instalamos. Necesitamos implementar ciertas configuraciones en nuestra memoria y reconocimiento de patrones que nos permitan almacenar información en soportes externos; ciertas configuraciones serían posibles sólo porque existen estos recursos. Se estima que la mentalidad ágrafa podría comportar arquitecturas virtuales muy distintas a las de sociedades analfabetas.

La máquina virtual von neumanniana instalada en nosotros sería un complejo de memes que se han replicado de modo profuso. Se dijo que la replicación de los memas sólo depende de que esta sea buena para ellos mismos. Ahora bien, ¿es buena para nosotros? Un meme podría ser sólo un buen replicador, pero completamente nocivo para nosotros. Dennett parece optar por una opción intermedia. Se preocupa en señalar que, el hecho de que encontremos a nuestro alrededor datos que evidencian algunos de los beneficios que nos proporciona la conciencia, no nos debe llevar a creer que todas y cada una de las partes tienen una función. "Hay espacio suficiente para ciertos hechos brutales que carezcan por completo de justificación

funcional. Algunos memes de la conciencia podrían ser, simplemente, memes egoístas" (Dennett 1991a, p. 234). Sin embargo, el resguardo teórico sobre cómo efectivamente funciona la evolución no inhabilita para el filósofo la pregunta heurística de la ingeniería inversa: ¿para resolver qué tipo de problemas fue diseñada esta máquina?

Dennett retoma la hipótesis de Julian Jaynes (1976) de que la auto-exortación y auto-remembranza que permite la conciencia son requisitos necesarios para el autocontrol que se precisa para prácticas como la agricultura, el proyectar edificios, etc. Por su parte, Douglas Hofstadter (1985) señala que el flujo de la conciencia consta de un tipo de autosupervisión que hace que no caigamos demasiado a menudo en nuestros propios errores. Y Nicholas Humphrey (1986) la rescata como el medio gracias al cual se pueden explotar las simulaciones sociales (usando la introspección para guiar nuestras intuiciones sobre lo que los otros están pensando o sintiendo).

Más allá de estos hábitos especializados, Dennett sostiene que la conciencia es responsable de la capacidad básica de resolver el problema de "en qué pensar después". Platón también vio de algún modo este problema, al preguntarse si era posible "poseer un conocimiento sin tenerlo" ("Teeteto", 197c), usando la metáfora de una persona que tiene aves agrestes (como el alma puede poseer conocimientos) pero no puede hacer que acuda el ave correcta cuando se la llama. Al problema de hacer que venga a nosotros el conocimiento adecuado para resolver un problema determinado lo empezamos a solucionar gracias al *razonamiento*. El flujo de la conciencia puede ser pensado como una especie de razonamiento interno que permite vincular información con la que el sistema, de algún modo, ya contaba, pero a la que no podía acceder en momentos en los que hubiera sido útil. "Aprender a razonar es, de hecho, aprender estrategias de recuperación del conocimiento" (Dennett, 1991a, p. 236). Del mismo modo que hablar o escribir puede ayudar a encontrar esa respuesta, el flujo de la conciencia, o la conversación que se tiene con uno mismo, lo puede hacer de modo más adecuado.

Incluso un conductista como Ryle, en su libro póstumo, *On thinking* (1979), ha dicho, para responder a qué sucede cuando pensamos, que el pensamiento consciente debe ser entendido como un modelo de auto-enseñanza, o mejor, una auto-educación o entrenamiento. Si bien él no tiene mucho que decir acerca de cómo se da este mecanismo, ya se sugiere aquí que no somos los capitanes de la nave, no

hay un “yo” consciente que está comandando los recursos de la mente sin ningún problema. Más bien somos algo desunificado. Nuestros módulos componentes tienen que actuar de formas oportunistas y sorprendentemente novedosas para producir una mínima unidad comportamental, la que es entonces mejorada por la ilusión de una unidad mayor. Así, desde Ryle ya se puede pensar que esa unidad ilusoria de la conciencia mejora el desempeño de nuestra la unidad comportamental desarrollada.

Aún aceptando la idea de Dennett de que el razonamiento es lo que nos permite hacer “que acuda el pájaro adecuado”, sigue cabiendo preguntar por qué funciona de tal modo, y por qué a veces no funciona. Margolis (1987) dice que dado que un ser humano hoy en día no puede mantener más que por unos segundos su atención en un problema, y que los problemas que resolvemos muchas veces precisan de un periodo temporal más amplio, es obvio que el modo en que resolvemos ciertos problemas involucra periodos de meditación seguidos de periodos de recapitulación, que van conduciendo a resultados intermedios. Al formular los resultados intermedios, los vamos introduciendo en la memoria, lo cual es necesario para que no se pierdan en el flujo de la conciencia. Estos hábitos individuales de auto-estimulación (consistentes en decirnos a nosotros mismos los resultados intermedios) son los que nos llevan a suplir el hecho de que nuestra memoria no sirve, según su configuración innata, para ser super-fiable, de rápido acceso, de acceso aleatorio; por lo que en la configuración de la máquina von Neumanniana se utilizan esos trucos para sustituir estas carencias. De la misma naturaleza serían los ensayos, las repeticiones, las rimas, los movimientos del cuerpo que hacemos para pensar o recordar cosas.

Ahora bien, toda esta comparación entre la máquina de von Neumann y la conciencia no explica una diferencia básica entre ambas: la primera no es consciente. La respuesta de Dennett es que la máquina de von Neumann, al estar configurada desde el principio con conexiones de máxima eficiencia, nunca tuvo que convertirse en objeto de percepción para sí misma. En cambio, la máquina joyceana tiene a sus propios trabajos como objetos de percepción, tanto como cualquier objeto del mundo exterior. O sea que la diferencia entre la máquina de von Neumann y la conciencia radica en la reflexividad de ésta última. Y esa reflexibilidad y experiencia fenoménica son efectos de que seamos imperfectos, y de que no hayamos sido diseñados de manera premeditada, sino que seamos fruto de un accidente.

Finalmente, podemos interpretar la analogía de la conciencia con el Software como un modo de dar respuesta a una de las inquietudes que se pone en cuestión cuando Dennett elabora su explicación de la conciencia: la fuerte intuición que tenemos de que, efectivamente, hay en nuestras mentes algo así como un lugar al que todo acude: el teatro cartesiano. En el proceso de I + D (investigación y desarrollo) realizado por alguien que diseña softwares, se va realizando un esfuerzo para conseguir formas cada vez más intuitivas y útiles de interfaces de usuarios, esto es, del espacio en el que se establece la interacción entre la máquina y la persona. De un modo similar puede verse al proceso de I +D realizado por la madre naturaleza, que dió lugar a interfaces de usuarios cuando empezó a existir la comunicación entre individuos. Así, una de las características de nuestro software sería crear una ilusión de Teatro Cartesiano que las otras especies no tienen. Así pues, el Teatro Cartesiano es presentado como estrictamente falso, bajo la explicación dennettiana, en la medida en que se niega que exista tal cosa en términos físicos (un cuartel central en el cerebro), pero, pese a ello, es reconocido, y explicado, en tanto ilusión.

Los relatos evolucionistas que nos condujeron hasta aquí toman como criterio, más que la posibilidad lógica, la plausibilidad biológica. Como hemos visto, esta explicación no excluye que se consideren los componentes culturales, que son vistos como una continuación de los biológicos, y no exige que los mismos sean reducidos a términos propios de la biología, dado que se mantiene una distinción de niveles de complejidad, que están vinculados entre sí pero que tienen autonomía explicativa.

Las presas de los castores, las grandes telas de las arañas, las hermosas y sofisticadas glorietas de los pájaros jardineros, son todas creaciones que ocupan un lugar central en el modo en el que estas especies viven. Estos hábitats pueden ser entendidos como ampliaciones de los límites territoriales, ya sea para dificultar transgresiones negativas, ya sea para facilitar transgresiones positivas. El resultado de esas construcciones puede pensarse desde Dawkins (1982) como fenotipo ampliado (o extendido), bajo la idea de que la presa del castor y el nido del pájaro no deben ser entendidos como meros *productos* del fenotipo, sino como *partes* del fenotipo. Dicho concepto se aplica de igual modo al caso de un caracol que utiliza parte del calcio de los alimentos que consume para producir una concha, que al del cangrejo ermitaño, que prefiere apropiarse de una concha producida por otro animal -por lo que este principio biológico no debe ser entendido desde una perspectiva de

la producción orgánica, sino que también incluye la apropiación. El fenotipo ampliado puede llegar también a incluir a otros individuos de la especie.⁴⁰

A diferencia de las arañas y los caracoles, los humanos no exudamos el material con el que construimos nuestros hábitats. Nuestro fenotipo ampliado más notable consiste en tejer palabras: son las narraciones nuestra forma de crear un hábitat y de proveernos autoprotección, de brindarnos sustento y de favorecer nuestra perspectiva para las relaciones sociales.

Dijimos que el “yo” biológico o mínimo debe ser entendido como una abstracción, como patrones de comportamiento que permiten, por ejemplo, que los animales no se coman a sí mismos, que no protejan un objeto externo que no les interesa, o que actúen en consecuencia con el hecho de que el dolor que sienten es de ellos. Esos patrones son el reflejo de un principio de auto-cuidado en torno al cual fueron diseñados (Dennett, 1989, p. 5). El “yo” narrativo o psicológico es otra abstracción, mucho más compleja pero igual de robusta que la abstracción biológica.

Cuando contábamos la anécdota del radar común y los barcos (en la p. 89 de este trabajo), intentábamos dar cuenta del proceso de identificar signos externos de movimiento corporal. Al menos en los seres humanos la necesidad de autoconocimiento va mucho más allá de esta dimensión: necesitamos conocer sobre nuestros estados internos, nuestra toma de decisiones, nuestras tendencias, nuestros defectos y virtudes. En este ámbito, la forma que los humanos tenemos de “hacer algo y ver qué se mueve” es a través de prácticas que involucran momentos de contar historias y de verificar historias, tanto reales como ficticias (podemos suponer que cada una de ellas tienen un tipo de verificación diferente: las reales se verificarían con respecto a los hechos concretos que constituyen nuestra biografía, las ficticias, con respecto a lo que producen en nosotros, y a qué dice de nosotros que las mismas produzcan tal o cual cosa). Estos procesos, que en los niños adoptan la forma del juego, la simulación, la formulación verbal en voz alta, en la mayoría de los adultos se dan en forma silenciosa, con una pretendida distinción entre las reflexiones serias y las que giran en torno a nuestras fantasías.

⁴⁰ Para ilustrar la idea de fenotipo ampliado, Dennett dice que las enciclopedias de Ciencias Naturales donde se representan, entre otros animales, a una hombre y una mujer desnudos, caen en un error del mismo tipo que si representarían a un oso andando en bicicleta y vestido de payaso: ninguna de las dos representaciones refleja el modo en que dichos animales se encuentran en la naturaleza!

Estos flujos narrativos apuntan a un mismo origen, no por el hecho de haber sido narrados por una misma persona, sino por producir el efecto de postular un agente unificado al cual atribuirle dichos relatos. El “yo” es la importante abstracción que nos permite recopilar la información de las personas sobre sí mismas. En ese sentido, decimos que tal fenómeno es el centro de gravedad narrativa de todas las historias que conforman la biografía del cuerpo viviente de alguien.

Ross y Brook (2002) han señalado algo interesante respecto de esta tarea de unificación que vendría a cumplir la conciencia para Dennett: “Nótese que la explicación de Dennett de la génesis de los “yoes” sólo es posible si la gente no tiene acceso directo a los estados informacionales de sus cerebros, y por ende debe predecir y explicar su propio comportamiento usando los mismos recursos que están disponibles para las terceras personas” (p. 30, mi traducción). Por ende, esta explicación del “yo” es coherente con la idea señalada al comienzo de que la conciencia no es transparente y que el de la primera persona no es ningún lugar privilegiado para la explicación del fenómeno, por lo que resulta pertinente intentar brindar una explicación objetiva del mismo.

Al aunar todos estos órganos especializados, que evolucionaron independientemente, ante una causa común y dotando, así, al conjunto de unos poderes muy mejorados, la máquina virtual, este software del cerebro, hace posible la emergencia de una suerte de organización centralizada en medio de tanto caos. Según el modelo de las versiones múltiples, en nuestro cerebro se dan procesos de fijación de contenido, que generan distintos “borradores”, fragmentos narrativos, etc. Luego, estos borradores atraviesan procesos de selección, en los que los borradores “ganadores” van generando conexiones (a veces más y a veces menos coherentes entre sí), que dan como resultado “capitanes virtuales” temporarios, que podríamos pensar como modos o criterios de razonar que se tornan hegemónicos. En palabras de Dennett: “la máquina virtual, este software del cerebro, lleva a cabo una especie de milagro político interno: crea un capitán virtual para la tripulación, sin ascender a ninguno de ellos al rango de dictador vitalicio” (Dennett, 1991a, p. 242). Estos “capitanes virtuales” deben tener algún vínculo con el “yo”, sin embargo es difícil de entender cómo se da tal vínculo dado que, como se dijo, se trata de “cargos” que son sólo temporales. Ahora bien, la explicación dennettiana sobre cómo se dan las “sucesiones” puede ser iluminadora: “¿Quién está al mando? Primero una coalición y

luego otra, en una alternancia que no es caótica gracias a unos buenos meta-hábitos que tienden a producir secuencias coherentes y resueltas en vez de una interminable y atropellada carrera por el poder” (Dennett, 1991a, p. 242). Es gracias a la formulación lingüística, ya sea expresada en voz alta o solamente “mental” (o sea, el razonamiento), que se imponen los contenidos que son más “populares” en nuestras mentes, y van elaborando una secuencia medianamente coherente (el término ‘populares’ para referirse a los contenidos mentales que adquieren preponderancia frente a los demás, es adoptado en Dennett, 2005). Uno de los recursos para obtener esa coherencia consiste en que el modo en el que organizamos nuestro discurso produce la ilusión de que hay un autor del texto: un comandante único a cargo que ha decidido producirlo. Esa ilusión es el punto de autorepresentación en torno al cual nuestro cerebro organiza toda la información que tiene sobre nosotros. Así, las historias que dan coherencia a la sucesión de “capitanes virtuales”, o de fragmentos narrativos que provienen de distintos módulos o circuitos, como siendo creadas por un agente unificado, nos conforman como un “yo”. En esta línea, Dennett dirá que, aunque parezca que nosotros creamos tales historias, en realidad son ellas las que nos crean a nosotros.

La explicación dennettiana del funcionamiento de la conciencia está muy vinculada con el modelo que el filósofo elabora para explicar la producción del lenguaje -señalando que abundan las teorías sobre la comprensión del lenguaje, pero no sobre su producción-, es decir, un modelo de cómo podemos entender lo que ocurre en nuestras mentes cuando efectuamos emisiones lingüísticas (Dennett, 1991a, cap: “Cómo las palabras hacen cosas con nosotros”). Si bien no abordaré este tema en el presente trabajo, sí es preciso señalar que en dicho modelo se vuelve a rechazar la existencia de un “centro de mando” en el cerebro, y se otorga un rol central a la formulaciones lingüísticas de las ideas más que a las ideas desnudas, en la medida en que, al parecer, muchas veces adoptamos una idea porque nos gusta cómo está formulada. En lugar de pensar que tenemos una idea previa, escrita en alguna suerte de lenguaje del pensamiento, Dennett sostiene que muchas de las cosas que decimos las decimos más porque nos gusta cómo se oyen que por el hecho de que el significado de las mismas sea algo a lo que hayamos arribado como conclusión de nuestros razonamientos. También hay veces en las que descubrimos qué pensamos

sólo después de decirlo y encontrar que no hay correcciones a realizar.⁴¹ Esta forma de explicar la producción del lenguaje en relación con su papel en la conformación de nuestra identidad personal es complementaria de su concepción de la evolución cultural: la memética. A fin de cuentas, las palabras, frases y expresiones que adoptamos porque nos gusta cómo suenan, no son más que los cuerpos que adquieren los memes que invaden nuestra cultura, y, en palabras de Dennett:

no podría haber un medio más agradable en el que los memes puedan replicarse que un sistema de producción del lenguaje en que los burócratas supervisores han dimitido en parte⁴², cediendo una gran cantidad de control a las palabras mismas, quienes luchan entre ellas por una oportunidad de estar en el candelero de la expresión pública (Dennett 1991a, p. 257).

Sin embargo, el modelo de producción del lenguaje de Dennett también señala, en vínculo con lo desarrollado en el párrafo anterior, que las producciones lingüísticas terminan constituyendo un hilo coherente. Como dijimos, no hay un ente unificado que haya decidido producir el discurso, pero el discurso se produce *como si* lo hubiera. Por ejemplo, lo que proferimos en primera persona tiene por sujeto a un “Yo” (aunque en el caso del español muchas veces sea tácito). Ese “Yo” es una abstracción en torno a la cual nuestro cerebro estructura la información que produce, aunque en última instancia la misma provenga de diferentes circuitos o subsistemas. En función de todo esto se torna comprensible la famosa frase de Dennett según la cual: “Nuestras historias se urden, pero en gran parte no somos nosotros quienes las urdimos; ellas nos urden a nosotros. Nuestra conciencia humana, nuestra egoticidad narrativa es su producto, no su origen” (Dennett 1991a, p. 428).

⁴¹ Ahora bien, ¿qué pasa cuando no tenemos una sensación de descubrimiento al escucharnos decir lo que pensamos? Una interpretación posible es decir que es porque justamente nosotros somos la fuente de esos significados. La otra, defendida por Dennett, es que no tenemos esa sensación simplemente porque es muy obvio lo que queremos decir con cosas como “por favor, pásame la sal”; no se necesita un acceso privilegiado para entender esto de manera inmediata.

⁴² Con un modelo “burocrático” Dennett se refiere, en este contexto, a uno en el que habría una estructura jerárquica con un “centro de mando” en el que se toman las decisiones acerca de qué mensaje transmitir y otros subsistemas encargados de ejecutar las órdenes enviadas por el centro de mando, sin modificarlas.

Yoes múltiples

Desde esta explicación naturalista acerca de cómo se origina un “yo”, la puesta en cuestión de su unicidad puede constituir un problema social o psicológico, pero ya no uno metafísico: la precariedad del mismo no es algo que contradiga la naturaleza del fenómeno. Martin y Barressi (2006) señalan que a finales del siglo XX parece haber acuerdo en que el “yo” no es otra cosa que multiplicidad. Sin embargo, esto va en contra de fuertes intuiciones básicas que tenemos respecto de que el “yo” debe tener una naturaleza unificada (cuya mayor cristalización quizá sea la que se deja ver en el idea de alma). Dennett elabora una visión en la que no se niegan ninguna de las dos cosas. Se reconoce la multiplicidad constitutiva que subyace al “yo”, al tiempo que se presenta a la unidad del mismo como una adquisición que es resultado de un proceso.

El modo en el que Dennett se relaciona con su objeto de análisis e investigación se deja ver en el siguiente pasaje:

Si pudiéramos ver cómo sería que dos (o más) yoes compitieran por el control de un único cuerpo, podríamos comprender mejor lo que es realmente un único yo. Como investigadores del yo, nos gustaría llevar a cabo experimentos controlados, en los que, variando las condiciones iniciales, pudiéramos observar lo que tiene que ocurrir, en qué orden y con qué requisitos, para que emerja un yo parlante (Dennett 1991a, pp. 429, 430).

Por cuestiones éticas no es posible llevar a cabo experimentos en los que se intente que emerja un nuevo “yo” para pelear con otro por el control de un cuerpo, pero podemos aprovechar los *experimentos hechos por la naturaleza* y extraer conclusiones. Mencionamos que lo normal, pero no necesario, es un “yo” por cada cuerpo. Esto no sucede en los casos de Trastornos de Personalidad Múltiple. Este trágico fenómeno puede resultar muy ilustrativo para comprender el de la identidad personal. Del mismo modo que la fuerza de gravedad es difícil de asimilar conceptualmente sin la idea de que hay partes del espacio en las que la gravedad actúa con otras magnitudes sobre los cuerpos, un “yo”-simple es difícil de discernir sin recurrir a la idea de un “yo”-múltiple. Dennett recurre a los casos de Trastorno de Personalidad Múltiple no para explicar la anomalía, sino para intentar ver la continuidad con los casos corrientes, mostrando que dan cuenta de un patrón de

comportamiento que ya está presente, aunque en menor grado, siempre que hay un “yo”. Para algunos, que existiesen casos en los que un mismo cuerpo fuera compartido por más de un “yo”, sería un fenómeno paranormal. Si el “yo” fuera un alma, como sustancia separada, el único modo de entender el fenómeno en el que, en un mismo cuerpo parecen convivir múltiples “yoes”, sería pensando que ese cuerpo fue poseído por una o más almas foráneas. Sin embargo, Dennett señala que la existencia de más de un “yo” en un cuerpo no debería resultarnos un fenómeno tan extraño en el plano metafísico, pues, en el mismo sentido ya es sorprendente que exista un “yo” por cada cuerpo. Eso no quita que el hecho de que existan cuerpos capaces de producir personalidades diferentes -cada una con sus propias narraciones, patrones de acción y de percepción del mundo, con algún grado de coherencia interna- nos pueda parecer sorprendente a nivel humano. Las explicaciones más exitosas del trastorno en cuestión son compatibles con la perspectiva narrativista. Lo que sucedería, según determinadas investigaciones, es que ante situaciones muy difíciles de afrontar, algunos pacientes se “bifurcan”, dando lugar a algo que se puede entender como la creación, de forma temporal, de un segundo centro de gravedad narrativa (Humphrey, 1986; Humphrey y Dennett, 1989). La existencia de este último tiene una temporalidad más acotada, por lo que no contaría con el tipo de autobiografía que poseen los “yoes” corrientes. Estos casos patológicos en los que se construye más de un “yo”, muestran que algo casi igual de sorprendente pasa en los casos “normales” en los que se construye un solo “yo”. Todos estamos todo el tiempo narrándonos a nosotros mismos la historia de nuestras vidas (o historias parciales, de aquellos segmentos que nos interesan de ellas). Somos espontánea y continuamente, “inventivos novelistas autobiográficos” (Dennett, 1992, p. 6, mi traducción).

Al comienzo de este capítulo nos referimos a algunas intuiciones que tenemos acerca del “yo”, que caracterizamos como obstáculos para desarrollar una mirada naturalista del problema. Los Trastornos de Personalidad Múltiple constituyen un buen ejemplo de cómo algunas de estas intuiciones pueden ser cuestionadas. La idea de que no puede haber “cuasi-yoes” o especies de “yoes”, con lo cual el “yo” es algo de todo o nada; y la intuición de que debe haber un número entero de “yoes” (igual a 1) asociados a un cuerpo, entran en conflicto con la evidencia que brinda este trastorno. La perspectiva dennettiana nos invita a tratar este problema no del modo en que se

trabaja con una sustancia unificada y definida, sino realizando un cambio en nuestra concepción del “yo”, que deberá producir, a su vez, un cambio en el tipo de preguntas que realizamos sobre el fenómeno en cuestión. Dennett señala que todavía perviven algunos vestigios de la idea de alma en nuestra concepción del “yo”, por lo que lo que a menudo lo entendemos como una ‘perla mental’, es decir, como una entidad con existencia continua e independiente (Dennett, 1991a). Desde un abordaje naturalista y evolutivo del “yo”, estas características deberán ser dejadas de lado.

Con la reconstrucción de la explicación del origen de la conciencia y el “yo”, espero haber contribuido a mostrar ciertas ventajas del naturalismo dennettiano. A lo largo de este capítulo se tomaron en consideración aspectos de explicaciones evolucionistas que no caen en el problema de asumir un pan-adaptacionismo, al tiempo que, por el lugar que otorgan a la evolución cultural, no incurren tampoco en el error de querer explicar aspectos de la cultura humana recurriendo a una idea de selección de genotipos que aumentan el *fitness*. Por otra parte, aunque dichas narraciones evolutivas deben estar sujetas a revisión y están lejos de ser infalibles, permiten mostrar una historia rica en complejidad que exige, a su vez, un cambio en el modo de pensar ciertos fenómenos. Los razonamientos analíticos que nos hacen entender los fenómenos en términos de “todo o nada”, se ven desdibujados ante la evidencia de la historia natural, que muestra el carácter procesual y de límites difusos que han llevado a las cosas a ser lo que son.

Por otra parte, en la explicación de Dennett, la especie humana no es ubicada en un pedestal, de modo que otorgue a sus características y habilidades una historia evolutiva absolutamente diferente a las del resto de las especies. Sin embargo, eso no va en detrimento de que Dennett presente dichas características y habilidades como resultado de un camino evolutivo muy particular. Una mayor especificación de en qué consisten esas continuidades y esas diferencias será desarrollada en el capítulo siguiente.

5.El acto de narrar: Instintivo y deliberado

Como se señaló al comienzo de este trabajo, uno de los principales desafíos de las teorías naturalistas en filosofía de la mente se puede resumir diciendo que las mismas pretenden construir explicaciones que eviten brindar una imagen excesivamente deflacionada de las características y capacidades de la mente humana, al tiempo que tampoco incurran en el error de presentar a las mismas como originándose a partir de una suerte de salto milagroso, o sostenidas por un “gancho celeste”. En este capítulo intentaré presentar algunas piezas que pueden considerarse claves para defender que Dennett no está cometiendo ninguno de estos errores: ni presentando una teoría que precise de un salto “mágico” para dar cuenta de nuestras habilidades, ni proponiendo una visión demasiado empobrecida de estas últimas. La peculiar decisión de Dennett de mantenerse en el punto medio de esta tensión, sin renunciar a ninguna de las dos pretensiones, es lo que hace que Clark (2002) hable de una *perspectiva dual* en su filosofía: “Dennett presenta las mentes humanas tanto profundamente diferentes de, como en profunda continuidad con, las mentes de otros animales y agentes simples” (p. 187, mi traducción).

Considero que el aspecto continuista de esta perspectiva dual también se manifiesta de un modo particular en su teoría del “yo”. No sólo porque, según la teoría de los sistemas intencionales, nosotros, los perros y los termostatos somos capaces de creer y desear fundamentalmente del mismo modo, sino porque, en la explicación dennettiana, el modo en el que construimos nuestros “yoes” es comparado al modo en que una araña teje su tela, y un pájaro construye su glorieta. Este es otro aspecto que resulta sumamente anti-intuitivo de la teoría y entra en tensión con algunas nociones fundamentales que tenemos respecto a la constitución de nuestros “yoes”.

En una primera instancia, pretendo defender que la versión narrativista de Dennett se distingue (o singulariza) por proponer una explicación de la elaboración de las narraciones que nos constituyen como “yoes” con componentes de carácter biológico. La presencia de explicaciones de carácter biológico en una instancia en la que no sólo estamos hablando de la constitución de algunas bases de la conciencia, sino de la construcción de complejos entramados lingüísticos, nos devuelve una

imagen en la que no aparecemos como creadores conscientes, antes bien, somos como arañas exudando una tela de forma relativamente instintiva. Esta imagen se complementa de forma coherente con la teoría de la intencionalidad derivada de Dennett, según la cual no somos dadores últimos de significado, sino que “si ha de haber alguna intencionalidad original –original sólo en el sentido de no derivar de ninguna otra fuente ulterior- la intencionalidad de la selección natural se merece el honor” (Dennett 1987, p. 281). Por otra parte, esta lectura de Dennett se ve apoyada por su teoría de la memética y su modelo de producción del lenguaje según el cual “las palabras hacen cosas con nosotros”, como se explicó en el capítulo anterior. Además de resultar sumamente anti-intuitivo, esto parece negar por completo toda posibilidad de tener una intervención consciente sobre aquello que nos hace ser quienes somos (problema directamente conectado con el del libre-albedrío, que Dennett trabaja en extensos libros y que no será abordado de modo directo en esta oportunidad). Pienso que una reconstrucción de la teoría de Dennett sobre el rol del lenguaje en la inteligencia, y la evolución cultural, permite negar esta presunta implicancia, y mostrar una imagen más aceptable de la construcción de nuestros “yoes”, que en su origen resulta instintiva y automática, pero que deja lugar para intervenciones deliberadas.

Exhudando narraciones

Podría pensarse que la idea de que sólo somos los relatos que construimos sobre nosotros mismos no es tan difícil de aceptar. Sin embargo, creo que hay otra arista de la propuesta del “yo” narrativo de Dennett, que no ha sido lo suficientemente señalada y que resulta un tanto más disruptiva. “Somos todos virtuosos novelistas” es una de las frases que encontramos reiteradas veces en los textos que abordan esta temática (Dennett 1988, 1989, 1992). Sin embargo, esa no es la respuesta a la pregunta: ¿qué somos?, en el sentido de: ¿qué son nuestros “yoes”? Esa es la respuesta a la pregunta de: ¿de dónde provienen todas esas narraciones?, y da cuenta de nuestra contribución a su creación que es, en gran medida, inconsciente.⁴³

⁴³ De hecho, se puede pensar que ese es el motivo por el cual en la segunda versión de su artículo sobre el Yo, Dennett cambió el título de “Why Everyone is a Novelist” (1988) por “The Self as the Center of Narrative Gravity” (1992).

Pretendo defender que, en la teoría del “yo” narrativo de Dennett, la forma en la que nos vinculamos con ese *corpus* de narraciones que nos hace ser quienes somos no es la de ser sus Autores, en un sentido fuerte. En *El Banquete* de Platón, Diótima habla de las obras, a través de Sócrates, como fruto de la sabiduría de los poetas y demás artistas, que son obras del espíritu preferibles a las obras del cuerpo, los hijos naturales, dado que son más bellos y más inmortales:

¿Y qué es lo que toca al espíritu producir? La sabiduría y las demás virtudes que han nacido de los poetas y de todos los artistas dotados del genio de invención... De esta manera el lazo y la afeción que ligan el uno al otro son mucho más íntimos y mucho más fuertes que los de la familia, porque estos hijos de su inteligencia son más bellos y más inmortales, y no hay nadie que no prefiera tales hijos a cualquiera otra posteridad, si considera y admira las producciones que Homero, Hesiodo y los demás poetas han dejado... (Platón, El Banquete, 208e).

Las obras de arte serían, en este fragmento, *hijas de la inteligencia*, objetos que sólo son capaces de producir aquellos que fueron *dotados con el genio de la invención*. En la teoría de Dennett, el entramado de narraciones que nos hace ser quienes somos no lleva, en un sentido análogo a las producciones de los artistas de los que habla Diótima, nuestro sello y firma de autoría. Podríamos distinguir así dos formas distintas de ser autores: el modo en el que comúnmente consideramos que Homero lo es de la Odisea⁴⁴, y el modo en el que comúnmente consideramos que yo lo soy de mis sueños, y el hornero de su nido⁴⁵. Desde la perspectiva narrativista, el “yo” es más bien una obra en el segundo sentido. Dos experimentos mentales que se encuentra en distintos textos de Dennett pueden ser ilustrativos al respecto.

Cuando Dennett (1988) compara al “yo” con un personaje ficticio de una obra literaria, señala que muchos nos vemos tentados a establecer una distinción nítida, de carácter metafísico, entre ambos, afirmando que los personajes de las novelas

⁴⁴ Retomamos el ejemplo de la Odisea de Homero porque es el que aparece en el texto platónico, dejando de lado las discusiones sobre si efectivamente existió un único autor de dicha obra.

⁴⁵ Esta distinción entre dos sentidos distintos de la autoría la formulo deliberadamente en el plano de las concepciones de sentido común, dado que existen discusiones filosóficas que ponen en cuestión nuestras concepciones corrientes sobre la autoría de las obras de arte y la idea misma de creatividad. Se pueden encontrar en Dennett (1995, 2001a, 2001b) reflexiones al respecto.

dependen, para su creación, de “yoes” reales, sus Autores o Creadores. Pareciera que para crear un “yo” ficticio es preciso un “yo” real. Sin embargo, Dennett argumenta en contra de esta arraigada intuición, a partir de un experimento mental.

Inicialmente, el experimento consiste en que imaginemos una máquina, producto de una investigación en inteligencia artificial, que tiene la capacidad de escribir novelas. Podemos imaginarnos que se le ha dado un gran stock de información, pero dado que no ha sido diseñada para escribir ninguna novela en particular, dicha información es utilizada en forma aleatoria. Si bien Dennett no lo explicita, es de suponer que deberían agregársele otros condicionantes, como simuladores que garanticen la continuidad de la historia sin contradicciones. Al producir una novela, los personajes que en ella aparecen serían personajes ficticios que no fueron creados por un “yo” real, es decir, no hay un ideólogo detrás del plan. Podríamos responder que, sin embargo, la máquina sí fue creada por diseñadores humanos; pero esto no es una objeción, dado que ellos no diseñaron a estos personajes ficticios. Ahora bien, Dennett señala que podríamos realizar el mismo tipo de exégesis literaria con esta historia que con cualquier otra. De hecho podríamos incluso leerla sin saber que fue realizada por una máquina. Nuestra reconstrucción interpretativa de lo que aprendemos sobre los personajes al leer la novela, gira en torno a ellos como personajes ficticios de cualquier otra novela. Luego Dennett complejiza el experimento, agregando que dicha novela, que tiene un narrador en primera persona que es el protagonista de la historia (Gilbert), en lugar de ser escrita en un papel es reproducida sonoramente por la máquina, a la que además hay que agregarle ruedas que le permiten desplazarse y una especie de ojo de televisión. Por la forma en la que se habría hecho la programación en esta versión complejizada del experimento, las aventuras de Gilbert ahora guardarán una fuerte relación, presumiblemente no casual, con las aventuras sobre ruedas de este nuevo robot. Así, si alguien golpea al robot con un bate, ahora la novela incluye el relato por parte de Gilbert en el que ha sido golpeado con un bate por alguien que lucía como esa persona. Si el robot se cae a un pozo, nos pedirá ayuda. En ese momento alguien podría preguntarse quién está pidiendo ayuda, si el robot o Gilbert, el personaje. Al ser ayudado, el robot nos envía un mensaje que dice: “Gracias por ayudarme. Saludos, Gilbert”. Alguien podría todavía resistirse a identificar al robot con Gilbert. ¿Por qué no podemos decir que Gilbert el robot es, de algún modo, un personaje ficticio, como nuestros “yoes”? Pareciera que la computadora del robot (su cerebro,

metafóricamente hablando), no sabe nada acerca del mundo; es sólo una tonta computadora que no sabe lo que está haciendo: “Ni si quiera sabe que está creando un personaje ficticio”, y aquí es dónde Dennett realiza un pequeño comentario que supone un cambio de perspectiva que tira por la borda el razonamiento anterior: “Lo mismo es verdad para tu cerebro, el tampoco sabe qué está haciendo” (Dennett, 1988, p. 19, mi traducción). De cualquier modo, los patrones de comportamiento que están siendo controlados por la computadora-cerebro del robot son interpretados por nosotros, al vincularnos con él, como una narración biográfica creciente de un “yo”. Además, el personaje ficticio de Gilbert no tiene sólo intérpretes externos, él también realiza una auto-interpretación, al proveerse a sí mismo de reportes sobre su actividad en el mundo.

Lo más perturbador de este experimento, presentado con el objetivo de mostrar la plausibilidad de la existencia de Gilbert, el robot, son sus consecuencias sobre el lugar que deben ocupar nuestros propios “yoes”. Las novelas de las que éste es protagonista en primera persona, son como las creaciones de nuestro cerebro (que guardan un correlato con las aventuras efectivas de nuestro cuerpo por el mundo), el cual ni siquiera sabe lo que está haciendo. El siguiente extracto de *The Origins of Selves* (Dennett, 1989) parece indicar que, efectivamente, este es el modo en que debemos leer dicho experimento mental:

Ahora estamos listos para la idea más extraña de mi artículo, pero también, creo, la más importante: hay una similitud más fuerte entre las arañas, los castores y nosotros. Las arañas no tienen que pensar conscientemente y deliberadamente, sobre cómo tejer sus redes; es simplemente algo que los cerebros de las arañas están diseñados para hacer que las arañas hagan. Igualmente, los castores, a diferencia de los ingenieros humanos profesionales, no planean consciente y deliberadamente las estructuras que construyen. Y finalmente, nosotros (a diferencia de los contadores de historias humanos profesionales) no resolvemos, conscientemente y deliberadamente, qué narraciones contar y cómo contarlas; como las redes de la araña, nuestros cuentos son tejidos por nosotros; nuestras conciencias humanas y nuestros yoes narrativos son sus productos, no sus causas. (Dennett 1989, pp. 168, 169, mi traducción)

Esta imagen también es consecuente con la que podemos encontrar en esta otra cita:

«Llámenme Dan» oyen salir de mis labios, y ustedes obedecen, pero no llamando Dan a mis labios, o a mi cuerpo, sino llamándome Dan a mí, la ficción de un teórico creada por... bueno, no por mí, sino por mi cerebro, actuando en concierto durante años con mis padres, hermanos y amigos. (Dennett, 1991a, p. 439)

La idea de que nuestros cerebros no saben lo que hacen cuando construyen un “yo” tiene el problema de ir contra la intuición de que somos autores conscientes de aquello que somos y de aquello que decimos que somos. La imagen de nuestros órganos cerebrales excretando narraciones, sin entender lo que hacen, suena terriblemente reduccionista, o al menos parece dejar de lado algo que consideramos esencial en la conformación de las personas: la posibilidad de intervenir conscientemente sobre quiénes somos.

Algo similar ocurre con un experimento mental formulado en *La Actitud Intencional* en el marco de una polémica sobre la idea de la intencionalidad intrínseca (Dennett, 1987, cap: “La evolución, el error y la intencionalidad”), que es una discusión que puede verse en relación con la pregunta sobre qué grado de autoría nos corresponde. En el marco de la controversia acerca de si existen hechos profundos que nos permitan distinguir nuestros estados mentales como siendo de un tipo diferente de aquellos que simplemente atribuimos a otros animales o artefactos tecnológicos, Dennett nos invita a imaginarnos que tenemos el deseo de ver cómo es el mundo en el año 2401. Las tecnologías disponibles permitirían cumplir tal deseo sólo si estamos dispuestos a ser congelados y permanecer en estado comatoso durante todos esos años. Dado que no podemos pretender disponer de la ayuda de descendientes cercanos que cuiden nuestro cuerpo durante ese período, deberíamos construir un dispositivo capaz de hacerlo. Deberá ser lo suficientemente resistente, y deberá poder ser autosustentable energéticamente, por lo que contará con dispositivos de transformación y aprovechamiento de la energía. Con respecto a la localización, hay dos estrategias posibles: la de encontrar el lugar indicado y confiar en que vamos a poder hacer uso del mismo todo el tiempo necesario, o bien, la de proveer al robot de un medio de traslado que le permita encontrar nuevos lugares con provisiones energéticas y escapar de los peligros. Cabe notar, por otra parte, que dichas estrategias se corresponden con las que se llevan a cabo con plantas y animales,

respectivamente, como resultado del proceso evolutivo. Dado que nosotros estaríamos completamente incapacitados para realizar cualquier toma de decisión, deberíamos proveer al dispositivo de cierta capacidad de autocontrol, que le permita derivar metas subsidiarias vinculadas al objetivo final de preservarnos, en función de los distintos contextos. El resultado del diseño será un robot con cierto grado de autocontrol, que mientras persigue el objetivo final de auto-preservación del organismo en su interior, se verá involucrado en distintos proyectos (muchos de los cuales pueden terminar resultando contraproducentes para nosotros). Sin embargo, el robot sólo tendría intencionalidad derivada, no tomaría *verdaderamente* decisiones. Dennett se explaya aún más en las características que podrían tener estos robots que pondrían en problemas a nuestro deseo de establecer una línea de corte entre la caracterización que hacemos de la situación hipotética de ellos y la nuestra. Sin embargo, el *quid* de la cuestión se encuentra en que la conclusión de que los robots sólo poseen intencionalidad derivada también debería extenderse a nosotros, dado que el experimento mental recién relatado es sólo una versión con variaciones de la propuesta de Dawkins (1976) de entendernos a nosotros mismos como máquinas de supervivencia que fueron diseñadas para perpetuar la existencia de sus genes *egoístas*. Según esta visión, es coherente sostener que su preservación es nuestra razón de ser original, “¡De manera que nuestra intencionalidad deriva de la intencionalidad de nuestros genes “egoístas”!” (Dennett 1987, p. 264).

Los dos experimentos mentales contribuyen a la idea que estamos sosteniendo, según la cual no somos autores en un sentido fuerte de los relatos que nos constituyen; es decir, nuestra autoría sobre nuestras narraciones no es del mismo tipo que aquella que le atribuimos a Homero por la *Odisea*.

Ahora bien, esta idea podría resultarnos, al menos hasta ahora, poco convincente, dado que los relatos elegidos para transmitirla parecen estar proponiendo una teoría de un grado de reduccionismo extremo o de un antropomorfismo extraño. La idea de que nuestras narraciones las escribe nuestro cerebro parece reducirnos, a nosotros como creadores, a las operaciones ciegas de un órgano que *no sabe lo que hace*. Por otra parte, parece que la imagen de los genes egoístas de los que se deriva nuestra intencionalidad, debe ser tomada con muchas pinzas. Siendo más precisos, el de los genes sería un caso en el que no podemos hablar de intencionalidad, ni intrínseca ni derivada, sino más bien de intencionalidad

como si: nosotros encontramos teorías o interpretaciones que les dan sentido a esos procesos mediante una atribución mentalista que no es consecuente con las otras cosas que sabemos del fenómeno en cuestión⁴⁶. En otras palabras: no parece que nuestra intencionalidad y sus complejidades sea derivada de la de los genes; decir que fuimos diseñados por nuestros genes que nos dotaron de intencionalidad, parece un antropomorfismo completamente injustificado.

Tomando en consideraciones estas dificultades, en lo que sigue querría proponer una interpretación de ambos experimentos que considero adecuada. Sin embargo, antes me voy a permitir examinar otro uso de conceptos, por parte de Dennett, que podríamos considerar, al menos en primera instancia, como “inapropiado”. Este se centra en en una figura antropomórfica, introducida por Dennett, que ya fue mencionada anteriormente: la de la Madre Naturaleza (que como se expuso en el capítulo 1, se puede abordar desde la actitud intencional, atribuyéndole intenciones y buenas ideas).

¿Son estos argumentos dennettianos subsidiarios del uso metafórico de ciertos términos? ¿Qué tan dependientes son de la adopción de la actitud intencional? ¿Pueden ser enunciados sin utilizar las metáforas anteriores? Afirmar que son nuestros cerebros, sin saber lo que hacen, quienes tejen las historias sobre nosotros mismos y los otros; y la de que somos máquinas al servicio de la perpetuación de nuestros genes egoístas, puede vincularse ahora con la tesis de que nuestra intencionalidad es derivada de la de la Madre Naturaleza. La idea de fondo es, nada más ni nada menos, la de que somos el resultado de un proceso evolutivo. Esto explica el hecho de que nuestro cerebro empieza a tejer historias antes de que podamos tener una idea acabada de qué es una historia y de que nosotros tomemos alguna decisión al respecto, así como también el que tengamos muchos instintos que son útiles para nuestra supervivencia, con independencia de que conozcamos su utilidad. Antes de poder tomar cualquier decisión de forma consciente, contamos con una batería de herramientas cognitivas y comportamentales para relacionarnos en el mundo sin las cuales, sin duda, no sobreviviríamos. Así, el gusto por las cosas dulces puede ser pensado como un rasgo seleccionado naturalmente por su capacidad para

⁴⁶ Dennett resalta la importancia de poder utilizar esta estrategia en biología para hablar de genes, enzimas, membranas, etc. Si bien son procesos pueden ser explicados en términos físicos o mecánicos, las atribuciones mentalistas nos sirven para entender qué es lo que intentamos explicar, hasta el momento en que logremos hacerlo.

llevarnos a consumir alimentos altos en glucosa (que después esto nos lleve a preferir consumir tortas de chocolate en lugar de manzanas es una reconversión de ese estímulo en un estímulo “supra-normal”). Algunas de estas herramientas tienen su origen en la evolución biológica de las especies por selección de genotipo, como el ejemplo del gusto por lo dulce recién mencionado, mientras que otras tienen un origen vinculado a aspectos que fueron mencionados al hacer referencia a la evolución de la conciencia: el lenguaje y la evolución cultural. Sin estos componentes adicionales, sería imposible obtener una explicación que incluya todo aquello que en nuestra especie forma parte de nuestro fenotipo extendido. Vivimos inmersos en una cultura, un lenguaje, un conjunto de valores. Todo lo que somos tiene una historia, biológica y cultural, que nos ha sido legada como puro resultado. En la medida en la que somos beneficiarios de esa herencia, inclusive antes de tomar ninguna decisión al respecto, podemos pensar que incluso los componentes lingüísticos y culturales contribuyen a la idea de que el proceso de creación de un “yo” no debe ser pensado como estrictamente análogo al que lleva a cabo un novelista cuando decide escribir una novela; más bien, se parece al de un castor construyendo una represa, o a dos niños imitando a sus padres el modo de relacionarse entre sí. Sin embargo, como veremos en el siguiente apartado, dichos componentes son los que nos permiten completar y enriquecer una explicación que, sin ellos, nos devolvería una imagen de nosotros mismos bastante anti-intuitiva (en la medida en que el carácter instintivo ha impregnado no sólo nuestro vínculo primario con el alimento y la protección, sino que parece estar en el seno de la construcción narrativa de nuestros “yoes”).

Por otra parte, esta lectura según la cual, para Dennett, en la elaboración de narraciones predomina un comportamiento instintivo antes que uno deliberado es consecuente con su teoría de la memética y su modelo de producción del lenguaje presentados en el capítulo anterior. La concepción según la cual *las palabras hacen cosas con nosotros*, es la que predomina en su propuesta de la memética y el motivo por el cual, según él, dicha teoría encuentra tanta resistencia a ser aceptada.⁴⁷ La idea de los memes resulta desagradable, en la medida en que pareciera que nos quita el estatus de autores y críticos de todas las ideas que poseemos; más que usina de

⁴⁷ La teoría de la memética ha sido blanco de fuertes críticas. Independientemente de si Dennett está en lo cierto al decir que el motivo último de las mismas es la imagen que nos devuelve de nosotros mismos, debemos decir que las principales apuntan a que no hay un claro criterio para individuar los memes (Sterelny, 2006) y a que la evolución cultural progresa de modo no análogo a la evolución biológica (Dupré, 2003).

creaciones, nuestras mentes parecen un hotel de paso de diferentes habitantes sobre los que no tenemos mucho control. Estas dificultades son las que se hallan en el centro de las confusiones que rodean a la idea de “yo”: ¿quién manda, nosotros o nuestros memes? Tanto desde el enfoque de la memética, como en términos generales, la perspectiva dennettiana parece ir en contra de nuestras intuiciones sobre la responsabilidad que nos cabe sobre eso que somos.

En función de lo expuesto hasta aquí sobre la explicación de Dennett de cómo elaboramos las narraciones que constituyen nuestros “yoes”, podemos considerar que la misma es coherente con una mirada de los seres humanos que se preocupa por entender a los mismos como una especie animal entre otras, pero que es defectuosa a la hora de conservar algunas de nuestras intuiciones de sentido común sobre el “yo”, en la medida en que presenta a dicha elaboración de narraciones como una actividad exclusivamente instintiva. Sin embargo, creo que si ponemos a dialogar lo dicho hasta el momento con otro aspecto de la teoría de Dennett -específicamente, la relación que establece entre lenguaje e inteligencia- podemos negar esta presunta implicancia negativa. Ese es el tema de los siguientes apartados.

El “yo” como artefacto cultural

“¿Por qué la importancia o excelencia de alguna cosa le debe llegar desde las alturas, desde algo más importante, como un regalo de Dios?” (Dennett, 1995, p. 98)

Conceptos

Esta visión teórica, que coloca a la especie humana en continuidad respecto de otras especies, no implica que no podamos establecer grandes diferencias entre una y las otras. El verdadero desafío naturalista es, antes que diluir esas diferencias, esbozar buenas explicaciones de las mismas. Con ese objetivo Dennett introduce un esquema, simplificado e idealizado, para entender las capacidades mentales de distintos organismos pertenecientes al mundo natural. El llama a este esquema “la torre de generación y testeo” (Dennett, 1994, 1995). Dicha “torre” está estructurada en diferentes niveles que clasifican a los organismos según sus capacidades cognitivas. Cada uno de estos niveles se diferencia progresivamente de los anteriores, en función de presentar una mayor complejidad en cuanto a las estrategias que pueden emplear los organismos pertenecientes al mismo para adaptarse al entorno.

A su vez, cada uno de estos niveles incluye las capacidades presentes en los estratos inferiores⁴⁸. En el primer piso de la torre se encuentra lo que Dennett decide llamar *criaturas darwinianas*: se trata de las criaturas que fueron creadas por el proceso de evolución de las especies, según el principio básico que produce organismos con variaciones por recombinación y mutación genética, de los que sobreviven los mejores diseños para los contextos dados. Por ende, el primer piso de la torre incluye a todos los organismos biológicos. Sin embargo, las criaturas que sólo pertenecen a este nivel son muy pocas, dado que estamos hablando de organismos que, a lo sumo, pueden detentar repertorios conductuales útiles, pero que no tienen la capacidad de adaptarlos o modificarlos de ningún modo en el caso de que el entorno presente una situación inesperada. Para ello, sería preciso contar con una capacidad que distingue a los organismos darwinianos de los del nivel que sigue: la capacidad de aprender a partir de la experiencia, que permite tener mayor flexibilidad conductual.

Una de las características a las que la selección natural dio lugar es a la plasticidad fenotípica, de la que hablamos en el capítulo anterior. Esto es, la característica de determinados organismos de no nacer con un diseño acabado e inmodificable, sino de poder cambiar ciertos rasgos, lo que les permite hacer una nueva selección entre diferentes variables en el desarrollo ontogenético. El sub-grupo de criaturas darwinianas que cuentan con esta plasticidad sujeta a procesos de condicionamiento, recibe por parte de Dennett el nombre de *criaturas skinnerianas*. Dicho nombre se debe a Skinner, padre del conductismo psicológico, quien sostuvo que el condicionamiento operante no es sólo análogo a la selección natural, sino que además es continuo a ella. Estos organismos, que no tendrían una configuración acabada de sus repertorios conductuales, despliegan diferentes conductas e incorporan de modo permanente aquellas que les brindan buenos resultados -o, en otros términos, aquellas de las que han obtenido “estímulos positivos”. Como cabe imaginar, aunque esta posibilidad puede implicar importantes beneficios comparados con la adaptación que pueden desarrollar las meras criaturas darwinianas, especialmente en ambientes cambiantes, tiene también algunas desventajas. Una de ellas es que el aprendizaje que se lleva a cabo requiere de mucho tiempo para poner a prueba las distintas opciones. Por otra parte, dado que obtienen

⁴⁸ Dennett aclara que la estructura ascendente de la torre no debe ser leída con connotaciones referentes a que una especie es mejor que la otra, o que va a sobrevivir más tiempo, sino que es una escala referente a la inteligencia, entendiendo a la misma en el sentido en el que se desarrolla en los distintos niveles.

este aprendizaje luego de poner ciertas opciones en práctica, pueden sufrir graves daños en dichos intentos, o incluso morir. Otro límite de las capacidades cognitivas las criaturas skinnerianas se puede explicar señalando que poseen lo que Dennett (1993) ha llamado “aprendizaje ABC”⁴⁹. A partir de un entrenamiento gradual, los animales obtienen complejas y poderosas capacidades discriminatorias. Sin embargo, las mismas están “ancladas” [*embedded*] en los tejidos específicos que son modificados por el entrenamiento y no pueden ser transportados a otro dominio, con el fin de ser utilizado en una situación completamente diferente.

Una mejora considerable con respecto al mecanismo skinneriano es posible gracias a la adquisición de la capacidad de representar internamente el medio. Esto es lo que hacen aquellos organismos que pertenecen al siguiente piso en la torre. En lugar de poner a prueba las distintas opciones comportamentales y correr el riesgo de morir en el intento, estos organismos cuentan con la posibilidad de realizar una preselección “interna” que les permite descartar aquellas opciones que evalúan como riesgosas o absurdas, antes de ensayarlas. El subgrupo de las criaturas skinnerianas que posee estas características recibe el nombre de *criaturas popperianas*, en honor a la idea de Popper de que una estrategia del tipo de la que Dennett atribuye a estas criaturas permite que nuestras hipótesis mueran en lugar de nosotros. Esta estrategia precisa, entonces, de algún dispositivo interno al organismo (como nuestros cerebros) con la capacidad de almacenar información acerca del mundo externo, de manera tal que la preselección tenga sentido. Dennett señala que esto no implica que se precise de una representación idéntica, que contemple cada aspecto del mundo externo, antes bien, lo importante es que se cuente con información que tenga relevancia de algún tipo para poder predecir con un grado suficiente de ajuste, el éxito en el empleo de una conducta determinada. Si bien se trata de capacidades muy beneficiosas, no encontramos aquí una explicación de la especificidad humana, dado que compartimos este peldaño de la escalera no sólo con otros simios, sino inclusive con pájaros, reptiles y peces: todos ellos poseen la aptitud para extraer información del medio para seleccionar su conducta antes de probarlas efectivamente. Sin embargo, entre la variedad de diseños que encontramos aquí, se puede establecer

⁴⁹ El nombre se debe a que es un tipo de aprendizaje que se pudo comprender a partir del esfuerzo acumulativo de teóricos de distintos ámbitos: asociacionistas (A), conductistas (B - en inglés-), y conexionistas (C).

una diferencia que divide aguas y que permitirá que dicha capacidad pueda ser explotada en grados de magnitud muy diferentes.

Para explicar esta suerte de punto de inflexión que divide a las criaturas popperianas en nuevos subgrupos, se puede apelar a otra imagen empleada por Dennett en los mismos textos (1994, 1995) sobre dos modos de proceder que podría tomar un cuerpo de agentes que actúan en conjunto con una misión en común. El primero de estos modos es el del espionaje, en el que cada agente debe saber sólo lo mínimo y necesario para cumplir su parte en la misión. El otro, es el del principio de comando, según el cual todos los miembros del equipo deben saber todo lo que sea posible acerca del plan, de forma tal de poder responder del mejor modo posible a imprevistos. Dicha distinción permite explicar una diferencia que parece ir en la dirección opuesta a la de las afirmaciones que comparaban la actividad humana de construir narraciones con la de otros animales en la producción de su fenotipo ampliado. En esta ocasión, Dennett señala que un niño podría copiar, paso por paso, el método con el que un castor construye una presa y obtener un resultado exactamente igual. Sin embargo, esto no invalida el hecho de que, incluso en estos casos, “hay dos métodos muy diferentes para construir presas: el método que utilizan los castores y el método que utilizamos los humanos” (Dennett, 1995, p. 614). La diferencia no radica en el resultado, sino en las estructuras de control que tienen los cerebros que desarrollaron la actividad. Mientras que el cerebro de los castores está diseñado según el principio del espionaje, nuestros cerebros humanos están estructurados según el principio de comando en equipo. La primera afirmación se ve apoyada empíricamente por experimentos como el de Wilson (1974) que, poniendo a prueba la reacción negativa que los castores parecían tener al ruido del agua que corre, colocó altoparlantes que reproducían grabaciones del fenómeno, a lo que los castores respondieron tapando los parlantes con barro. En cambio, los niveles de control más elevados permiten un grado de versatilidad que sólo presentan las criaturas que ascienden un piso más en la torre. Como se puede predecir en función de lo expuesto en el capítulo anterior, la adquisición de una estructura de control de mayor complejidad está vinculada con el lenguaje. En los párrafos siguientes explicaré cuáles son las características distintivas de las criaturas que habitan el último piso de la torre de generación y testeo: los seres humanos.

Nosotros tenemos algunas conductas rutinarias e instintivas. Muchas otras que son imitativas, es decir que aprovechan exploraciones realizadas no sólo por nuestros ancestros, sino también por nuestros grupos sociales (una forma de transmisión no genética de información que también compartimos con otras especies). Sin embargo, nuestra capacidad de planear actos deliberadamente es algo que cuenta no sólo con información reunida y transmitida por seres de nuestra misma especie de diferentes culturas y diferentes momentos históricos, sino además por objetos que contienen una cantidad de información que ningún ser humano individualmente considerado ha incorporado o entendido por completo. Esto nos hace ascender como especie un piso más en la torre de generación y testeo, y formar parte del sub-sub-subgrupo de las criaturas darwinianas que reciben el rótulo de *criaturas gregorianas*. El aporte de Gregory (1981), según Dennett, es haber entendido la importancia de la información, que él llama inteligencia potencial, en la inteligencia que llama inteligencia cinética. Desde la perspectiva de Gregory, el uso de herramientas no sólo requiere inteligencia, sino que además la confiere, dado que las mismas se entretajan con destrezas y habilidades previas de modo tal que las extienden, las amplifican y las enriquecen. Mientras más trabajo de diseño posea una herramienta, más información potencial confiere su utilización.

Cuando hablamos de herramientas debemos comprender que entre ellas, y en un lugar de gran relevancia, debemos ubicar a las herramientas mentales: las palabras. Las palabras confieren a las criaturas gregorianas una capacidad para ejercitar la preselección interna presente en todas las criaturas popperianas de un modo mucho más sofisticado, pudiendo realizar una cantidad exponencial de reflexiones sin encontrar un límite. En consonancia con lo explicado en el capítulo anterior, respecto de las capacidades del efecto Baldwin, cuando las palabras arriban por primera vez a los cerebros humanos actuales no encuentran una *tabula rasa*, sino un medio que es el resultado de presiones evolutivas más tempranas y que, según estudios actuales, favorece la apropiación de ciertos conceptos. El lenguaje es lo que permite las largas cadenas de pensamiento que hacen posible este modo tan fructífero de exploración interna. En definitiva, no estamos diciendo algo muy distinto a lo dicho en el apartado anterior, sólo que ahora se nos permite ver al lenguaje como una parte de la herencia que tenemos en la historia de la sofisticación de los mecanismos que los organismos vivos hemos utilizado para adaptarnos al ambiente. Aunque no tengamos incorporado un mecanismo interno que nos permita

un proyecto a largo plazo, como el del castor para construir una presa, podemos construir la presa y muchas cosas más, porque el lenguaje nos permite un tipo de encadenamiento del pensamiento que posibilita esta continuidad orientada a un objetivo.

En la punta más alta de la torre encontramos un mecanismo que ha superado la capacidad de los distintos organismos en la actividad de generar y poner a prueba: la ciencia. Esta se sirve de las herramientas mentales con las que contamos los seres humanos (las palabras). La práctica científica (que podemos entender como “el altillo” al que podemos subir las criaturas que habitamos el piso más alto de la torre, esto es, las criaturas gregorianas) se caracteriza por poseer muchos dispositivos que nos permiten hacer mejor precisamente aquello que marca la diferencia en el aprovechamiento de las capacidades popperianas: contemplar las opciones para someterlas a prueba. Del mismo modo que una fotografía nos permite contemplar mejor algo para incorporar esa información, el lenguaje puede entenderse como una tecnología que crea una clase de objetos completamente nueva, que pueden revisarse en cualquier orden y a cualquier velocidad. Además de esta nueva posibilidad de poner a prueba distintas opciones y de observar los errores, que acarrea la adquisición del lenguaje, la ciencia significa otro cambio cualitativo en la medida en que constituye una práctica en la que se trata ya no simplemente de observar nuestros propios errores, sino de hacerlo en público: “Cometer errores para que los vean todos, con la esperanza de conseguir ayuda de otros con las correcciones” (Dennett, 1995, p. 627). La adquisición del lenguaje ha posibilitado que cada uno de nosotros, gracias a nuestros lazos comunicativos, seamos beneficiarios de las labores cognitivas de los otros, lo que nos otorga un poder sin precedentes. Por ejemplo, consideremos el caso de la habilidad de emplear la actitud intencional. Dennett (1996) defiende que la capacidad para emplear esta beneficiosa actitud sería propia de las especies que se encuentran en los estratos más altos de la torre. Hay muchos primates no humanos que han demostrado poseer esta habilidad; sin embargo, nunca se han visto envueltos en discusiones acerca de qué atribución sería más adecuada en un caso en particular. Así, mientras estos psicólogos naturales no tengan posibilidad de compartir sus “tomas de notas”, de intercambiar opiniones sobre posibles razones, no hay motivo para que necesiten representarse a sí mismos tales razones, por lo que no van a necesitar acceder a un diseño de comando en equipo.

Para Dennett, la comparación fisiológica entre un cerebro humano y el cerebro de un chimpancé no sería muy útil para entender las diferencias entre las capacidades cognitivas de dichas especies, dado que, como intentamos exponer, las ventajas del primero tienen que ver con que está unido con miles de otros cerebros humanos de forma tal que puede ser beneficiario de logros cognitivos foráneos. Este sistema cognitivo colectivo es posible gracias al lenguaje. El lenguaje es lo que hace posible que tengamos el tipo de mente que tenemos: “un tipo de mente que, al ser capaz de cooperar con otras mentes y grabar los resultados de la cooperación para que otros puedan construir por encima, puede comprender la física del universo, encontrar curas a las peores enfermedades, construir el telescopio de Hubble y los canales de televisión, crear dispositivos artificiales que interpretan discursos y resuelven problemas, y mucho más” (Brook y Roos, 2002, p. 9, mi traducción).

Si bien es la práctica científica y la acumulación cognitiva lo que permite que hayamos llegado a elaborar conocimientos tan sofisticados como los citados en el párrafo anterior, la importancia de la adquisición del lenguaje para las criaturas gregorianas no radica en la elaboración de estas complejas teorías, sino que se manifiesta de un modo mucho más básico, en la transformación que se da en cada una de nuestras mentes. El beneficio más importante que extraemos de las experiencias cognitivas de otros tiene que ver, como señalamos anteriormente, con aprovechar la sabiduría que está corporizada en las herramientas mentales que otros han inventado, mejorado y transmitido. Según Clark (2002), debemos entender como una herramienta mental a cualquier constructo diseñado (o heredado culturalmente) que ayuda a transformar y simplificar la resolución de problemas en el entorno (mental) interno. Dennett (1996, 1998) brinda algunas pistas acerca de cómo podemos pensar el funcionamiento de tales herramientas. Esta nueva incorporación por parte del sistema cognitivo de un ser humano puede pensarse como el resultado de la apropiación de este instrumento. En una instancia previa, el sistema cognitivo disponía de la capacidad de agrupar cierta información relevante, pero es recién con la incorporación de las palabras que esa información se vuelve más “saliente” y susceptible de ser manipulada en futuras interacciones con el entorno. La adquisición de un lenguaje se asemeja de algún modo a la adquisición de etiquetas (Dennett, 1993) que nos permiten formular conceptos al aplicarlas a ciertos nodos de información que ya habíamos adquirido, en función de anteriores interacciones con el mundo.

En función de lo que veníamos exponiendo, podemos decir, siguiendo a Dennett, que parece que la gran diferencia entre el modo de actuar del cuco⁵⁰ cuando tira del nido los otros huevos, y el modo de pensar nuestro, tiene que ver con que el del cuco es conocimiento *en* el sistema, pero no conocimiento *para* el sistema. Podemos decir por eso que se trata del “aprendizaje ABC”, al que hicimos referencia antes. La “inteligencia” de su comportamiento está en algún sentido integrada *en* el cableado innato que logra este efecto tan robustamente, pero el cuclillo no tiene una clave acerca de esta “razón de ser”. A diferencia del mero entrenamiento laborioso al que se circunscribe este tipo de aprendizaje, nuestros cerebros están equipados de modo tal que pueden enriquecerse desde adentro al volver a representar el conocimiento que ya se han representado antes. El factor fundamental que permite este cambio es la adquisición de símbolos. Los símbolos son medios de representación explícitos, que permiten apreciar, entender y explotar conocimientos tácitos entretejidos en las redes neuronales. Así, podemos decir que la red conexionista que es la base de nuestro sistema cognitivo, está conectada a un sistema de símbolos que es exclusivo de nuestra especie.⁵¹

Resulta tentador objetar a esta propuesta que contiene cierta circularidad. La idea de que algo que nosotros mismos hemos creado nos vuelve más inteligentes, podría ser circular en la medida en que parece que para crearlo deberíamos contar con esa inteligencia previamente. No obstante, este no es un problema dado el concepto de diseño de Dennett, derivado de entender a la Selección Natural como un algoritmo ciego, que demuestra que los buenos diseños no precisan necesariamente de una planificación previa (1995, 2001b).

Si bien Dennett piensa que todavía no disponemos de las herramientas para elaborar una teoría sobre cómo se originó este sistema simbólico y cuál es su relación con la red conexionista subyacente, sí elabora una propuesta a modo de bosquejo

⁵⁰ El cuco (*Cuculus canorus*) es una especie de ave insectívora que practica el parasitismo en el momento de la puesta de huevos. Las hembras de cuco ponen sus huevos en los nidos de otras especies de aves, aprovechando la ausencia de sus dueños. Al nacer, los polluelos de cuco se comportan siempre de la misma manera: se deshacen o bien de los otros huevos que hay en el nido, si no han nacido, o bien de los pequeños polluelos de la especie parasitada.

⁵¹ Un ejemplo de filósofos que han defendido algo así son Clark y Karmiloff-Smith (1993), quienes propusieron que las capacidades para usar el lenguaje dependen de una capacidad subyacente en la arquitectura re-descriptiva de nuestro aparato cognitivo. Hay una diferencia de énfasis con respecto a ellos en Dennett (1993), quien señala que esa capacidad de redescrición que necesitamos para usar el lenguaje depende, en algún punto, de lo que el niño adquiere al recibir las estructuras prediseñadas del lenguaje natural.

tentativo de tales cuestiones (Dennett, 1993, 1996). La analogía con la cámara de fotos que mencionamos anteriormente puede volver a resultar útil. Esta tecnología significó un avance científico en la medida en que permitió observar con más detenimiento fenómenos que en tiempo real ocurren muy rápidamente. La analogía que quiere explotar Dennett es que el artefacto que permite este incremento de poder cognitivo mediante la captación de determinado fenómeno, es “estúpido”. Según el filósofo, algo similar ocurre en el proceso de la instalación del sistema simbólico que conforma los rasgos propios de las mentes humanas: para adquirir nuestra capacidad cognitiva que incluye tanto la de percibir como la de entender las representaciones simbólicas, necesitamos pasar por una instancia anterior en la que somos capaces de percibir las pero no de comprenderlas. La adquisición de un lenguaje implica un largo proceso en el que los niños se acostumbran a oír palabras e incluso a repetir las antes de comprender su significado. Muchas veces esas vociferaciones se realizan en circunstancias determinadas (vociferar “no hay que tocar porque está caliente” cuando se está cerca de la estufa). Estas prácticas son las que dan lugar a la asociación entre procesos auditivos y articulatorios, por una parte, con patrones de actividad o circunstancias específicas, por el otro, que nos hacen familiarizarnos con determinados rótulos. Lo que hace un rótulo, en términos de Dennett, es volver más “salientes” ciertos nodos de información, produciendo que de algún modo queden anclados en la memoria (Cfr. Dennett, 1993, p. 544).⁵² En el avance y sofisticación de este proceso, se van desarrollando nodos cada vez más salientes, con la capacidad de asociarse entre sí; se trata, nada más y nada menos, que de los conceptos. Una vez adquirida esta capacidad, la formación de un concepto puede no precisar de la asociación con una palabra, pero ellas sí son las necesarias precursoras de los primeros conceptos que podemos manipular. El resto de las especies podrá tener nodos de información respecto de ciertas cosas en el mundo que podemos llamar conceptos, pero no podrán considerar tales conceptos; eso se debe a que al carecer de lenguaje natural no tienen cómo “arrancar” los conceptos de las redes conexionistas

⁵² Dennett también profundiza en algunas características que deberían tener estos rótulos. Una de las condiciones para que las marcas que hacen de rótulos cumplan la función de poder ser soportes cognitivos, es que sean distintivas. Esto significa que tienen que poder distinguirse de algo que se haya originado accidentalmente. Por otra parte, debería ser algo que se pueda distinguir aunque se lo observe aisladamente, sin necesidad de estar observando todo el conjunto de símbolos del que forma parte. Así, por ejemplo, si yo estoy marcando cajas que ya revisé para distinguirlas de las que no revisé con el fin de no hacer trabajo innecesario, será mejor que haga una cruz en la tapa en lugar de abollar un poco una esquina.

interconectadas. En cambio, para los seres humanos los conceptos se vuelven objetos de manipulación y de escrutinio. Así, los símbolos, entendiéndolos simplemente como la representación perceptible de una idea, tienen de especial que pueden manipularse, trasladarse, hacer un aporte significativo diferente en un pensamiento diferente. Con esto, el lenguaje (y los códigos públicos) nos han provisto de un nuevo ámbito de objetos reconocibles y manipulables sobre los cuales se vuelven más sofisticadas las capacidades básicas de reconocimiento, imaginación y aprendizaje. Las herramientas mentales servirían para explicar por qué accedemos a ciertos niveles de abstracción, comprensión y aprendizaje.⁵³

En definitiva, retomando el dilema del *Teeteto* de Platón, ahora podemos explicitar mejor en qué consiste la habilidad de hacer que acuda el pájaro adecuado:

¿Cómo hacemos eso? Por medio de la tecnología. Construimos complejos sistemas de asociación mnemotécnica (señaladores, etiquetas, rampas y escaleras, ganchos y cadenas). Refinamos nuestros recursos mediante ensayos y componendas incesantes, transformando nuestro cerebro (y todos los accesorios periféricos que adquirimos) en una red de habilidades inmensamente estructurada. Todavía no se han encontrado evidencias que demuestren que haya ningún otro animal que haga lo mismo (Dennett, 1996, p. 179).

Autocontrol

Hasta aquí he explicado cómo, para Dennett, las herramientas mentales dan lugar a las capacidades cognitivas de abstracción, comprensión y aprendizaje, que

⁵³ Clark recupera un experimento que puede ser bastante ilustrativo al respecto. Thompson, Oden y Boyson (1999) han realizado un estudio de resolución de problemas en chimpancés, en el que los animales lograban resolver una nueva clase de problemas abstractos a partir de un entrenamiento. Los chimpancés eran entrenados para usar un marcador arbitrario de plástico (un triángulo amarillo) para señalar pares de objetos idénticos entre sí (por ejemplo, dos tazas) y otro marcador arbitrario (un círculo rojo) para señalar objetos diferentes entre sí (una taza y un zapato). Luego del entrenamiento los chimpancés eran capaces de realizar una tarea similar pero involucrando relaciones de un orden superior. Así, por ejemplo, frente a dos pares distintos de objetos idénticos (dos tazas por un lado y dos zapatos por el otro) los chimpancés debían señalar con el triángulo amarillo que se trataba de la misma relación (identidad), o frente a un par de objetos idénticos y un par de objetos distintos (dos tazas por un lado y una taza y un zapato por el otro) señalar con el círculo rojo que se trata de relaciones distintas (identidad en el primer par y diferencia en el segundo). El éxito de los chimpancés en esta difícil tarea se puede explicar apelando al entrenamiento anterior, suponiendo que han interiorizado mentalmente los antiguos marcadores.

serían imposibles sin aquellas. Sin embargo, esto no parece una respuesta satisfactoria al problema planteado anteriormente, a saber, que la práctica instintiva de narrar, entendida dentro de un esquema según el cual *las palabras hacen cosas con nosotros*, parece no dejarnos lugar para intervenir conscientemente en la constitución de quienes somos. Por ende, en esta última parte del capítulo intentaré señalar que la adquisición de los nuevos objetos de manipulación a los que dan lugar las herramientas mentales también están vinculadas a nuestra capacidad para una agencia consciente. Si bien no abordaremos aquí la teoría filosófica de Dennett respecto del libre albedrío (a la que dedica importantes libros, como *Freedom Evolves* (Dennett, 2003a), sí intentaremos identificar algunos aspectos relevantes sobre el vínculo entre las capacidades específicas de las criaturas gregorianas y la posibilidad de tener cierto tipo de autocontrol moralmente significativo (Dennett, 1984, 2003b).

Según lo expuesto anteriormente, las palabras son herramientas mentales capaces de objetivar conceptos y relaciones, de modo tal que podemos hacer con nuestro pensamiento cosas que nunca antes hubieran sido posibles. En definitiva, es el hecho de contar con el lenguaje lo que nos vuelve capaces de tener activamente pensamientos sobre nuestros pensamientos. Ahora bien, es razonable pensar que participar de intercambios comunicativos y cooperativos nos hace obtener recursos de una naturaleza completamente nueva, dado que en ocasiones nos obliga a preguntarnos cosas como “¿qué razones tengo para tal o cual cosa?”, “¿son buenas razones?”, “¿qué tan fiable es la evidencia que tengo para creer tal cosa y orientar mis acciones en virtud de esa creencia?” Ese tipo de auto-cuestionamiento se vuelve posible cuando las razones del agente se vuelven objetos para el mismo agente; esto es, cuando el agente tiene a su disposición cierta representación de las razones que le permite diseñarlas, editarlas, revisarlas, manipularlas, defenderlas. En función de lo expuesto, parece que las objetivaciones de soporte lingüístico de nuestros estados mentales guardan algún vínculo con la aptitud para la moral (y para la reflexión y evaluación orientada a otros dominios también), en la medida en que son necesarias para el proceso de poner en cuestión las propias creencias y razones para la acción, que da lugar al autocontrol. En definitiva, se está suponiendo que cierto tipo de autocontrol moralmente significativo depende de la capacidad para tratar

reflexivamente a los propios pensamientos, creencias y razones como objetos.⁵⁴ Dicha adquisición puede pensarse como el resultado de un proceso en el cual las criaturas van desarrollando de manera creciente la capacidad de autocontrol. De hecho, Clark (2002) señala que se puede pensar que hay, en Dennett, una cascada de autocontrol análoga a la de las estrategias adaptativas.

Uno de los desafíos tempranos con el que se enfrentan los seres humanos en su desarrollo ontogénico es el de aprender a controlarse a sí mismos. Se puede pensar que el grado más básico del autocontrol es el de controlar la motricidad del propio cuerpo. Unos peldaños más arriba nos encontraríamos con la capacidad de tomar decisiones de forma deliberada. Sin embargo, la toma de decisiones de forma deliberada todavía se diferencia de un piso ulterior, en el que encontramos un componente específicamente moral en la realización de elecciones. Para explicar en qué consistiría tal distinción, Dennett (1984) señala que un jugador de ajedrez se debe “hacer cargo” de su jugada, en la medida en que afirma que no estaba siendo controlado por nadie más al momento de realizarla, lo que lo convierte en artífice de sus elecciones. Este es uno de los aspectos en el que somos autores de aquello que somos: somos responsables de las acciones que realizamos y que conforman nuestra historia. Sin embargo, esto no es suficiente para la responsabilidad moral, en la que se contempla a un tipo de toma de decisiones que es el resultado de un proceso diferente, que Dennett (1984) llama la “autoevaluación radical”. Dicho proceso consiste en una indagación mediante la cual se intenta decidir qué es lo realmente importante y cuáles son los valores de fondo en los criterios de la criatura que la realiza. Así, en las autoevaluaciones, se ponen en tela de juicio los criterios que la misma criatura emplea en las evaluaciones corrientes. Por ejemplo, podemos pensar que si una persona intenta tomar algunas de sus decisiones en virtud de cuál de las opciones disponibles conduce a consecuencias más justas, en la autoevaluación radical pondrá en cuestión qué es lo “justo” para ella. Dennett señala, siguiendo a Taylor (1976), que no contamos con un lenguaje de meta-nivel que haga posible tal tipo de indagación. En ese sentido, es un proceso análogo al pensamiento filosófico, en la medida en que durante el desarrollo de la exploración se va transformando la pregunta misma. Según Dennett y Taylor, en el intento de formular los propios

⁵⁴ Nótese que, como señalamos en el apartado anterior, lo que permite distinguir una máquina joyceana de un software informático es la capacidad reflexiva de la primera, que es el resultado de estar provista de un diseño mucho más impreciso. Aquí la reflexividad es presentada como una condición de posibilidad de la toma de decisiones.

valores, quien realiza la indagación busca formular una idea de *qué es lo que realmente importa* en torno a la cual podrá ordenar los propios criterios, sin embargo, se trata de una idea que de algún modo ya está allí, pero no de modo definido y articulado, sino como una intuición borrosa. El proceso de conseguir formular esa intuición es lo que termina siendo constituyente de los “yoes”; por eso Dennett dice que: “creamos a nuestros valores al tiempo que nos creamos a nosotros mismos” (Dennett, 1984, p. 108). En definitiva, habría valores previos, implícitos, no articulados, que los sujetos buscan explicitar, articulándolos, para que ello les permita ordenar los criterios para tomar decisiones o evaluar acciones particulares. En función de esto, podemos decir que sólo las criaturas “gregorianas” de Dennett son capaces de realizar una radical autoevaluación. Esa autoevaluación requiere la articulación explícita de los valores e ideas inherentes en nuestras acciones y proyectos y, junto con ello, el desarrollo de habilidades deliberativas que después mejorarán nuestra capacidad para las acciones controladas (Dennett, 2003b).

Considero que sería consecuente con el planteo de Dennett, sostener que el formato narrativo es importante para la práctica de incorporar lo obtenido en la autoevaluación radical a nuestros comportamientos futuros. Entre la cantidad de información infinita que tenemos respecto de los comportamientos que llevamos a cabo, las narraciones permiten discriminar “una forma y un fondo” a partir de patrones de actividad. Así, podemos detectar estrategias de operación generales en dichos patrones de actividad y nos podemos hacer preguntas de meta-nivel sobre tales estrategias y evaluar cuán bien sirven a sus deseos, objetivos y necesidades. Nótese que este grado de asimilación de la autoevaluación es la cúspide de un proceso que sólo es posible porque se cuenta con un lenguaje auto-descriptivo, que le brinda al agente la posibilidad de representarse las propias creencias, intenciones y principios en un medio “objetivo”, de forma tal que pueden ser objeto de escrutinio y evaluación.

Por otra parte, es importante señalar que para Dennett la responsabilidad moral no se restringe a aquellos casos en los que hay un reporte consciente de la toma de decisión. Experiencias como aquellas en las que nos damos cuenta a posteriori de que ya hemos tomado una decisión, muestran que ésta no se identifica con su reporte consciente, por lo que podemos hablar de la misma sin presencia de esto último. La perfecta introspección, que muchos han convertido en una de las

características de lo mental, es una generalización injustificada de una capacidad que cabe considerar como una adquisición evolutiva tardía, que sólo empleamos en ocasiones limitadas, y que no siempre desempeña el mismo papel epistemológico o psicológico. Sin embargo, sí está en nuestras manos equiparnos de forma tal que desarrollemos habilidades deliberativas y la capacidad de acudir a ellas cuando sea importante.

En función de lo expuesto, podemos sostener que es la constitución de un “yo” lo que da lugar a que el autocontrol adquiera otra categoría: aquella a la que se accede únicamente cuando los patrones de actividad se vuelven transparentes o autoconscientes, y pueden ser contrastados y revisados en cada proceso de autoevaluación radical. En un artículo más reciente, Dennett (2003b) habla del “yo” como un *artefacto*. Tomando este término en su acepción más general, esto es, como un objeto fabricado con cierta técnica para desempeñar una función específica, podemos decir que una de las funciones de este nuevo artefacto es la de funcionar como un foco de autocontrol y de toma de decisiones. En ello se cifra la potencialidad de la emergencia de un “yo” responsable. Y es por eso que los “yoes” tienen proyectos, intereses y valores, que crean en el curso del desarrollo de sus habilidades autocontroladas y autoevaluadas.

Con esto podemos decir que examinar la teoría de Dennett sobre el rol del lenguaje en la inteligencia como parte de la evolución cultural permite explicar de modo naturalista cómo tenemos capacidades específicamente humanas de: i) abstracción, comprensión y aprendizaje, y ii) autocontrol moralmente significativo.

En este capítulo hemos intentado explicar algunas imágenes metafóricas a las que recurre Dennett, como la de que nuestra intencionalidad deriva de la de nuestros genes egoístas, exponiendo su concepción sobre los seres humanos como un resultado de un largo proceso evolutivo. Desde esta visión, muchos de los aspectos de nuestras capacidades y comportamientos son explicados en función de su asentamiento genético, por haber resultado útiles anteriormente (para la supervivencia de nuestros ancestros). Sin negar esto, la perspectiva de Dennett sobre el rol del lenguaje en la inteligencia nos permite sostener, al mismo tiempo, que esa herencia no opera en nosotros como una fatalidad. Por eso no es contradictorio que Dennett mismo sostenga que:

Nuestros intereses tal y como los concebimos y los de nuestros genes pueden muy bien diferir, aun cuando si no fuera por los intereses de nuestros genes no existiríamos: su preservación es nuestra raison d'être original aun si podemos aprender a ignorar ese objetivo e idear nuestro propio summum bonum, gracias a la inteligencia que nuestros genes han instalado en nosotros (Dennett 1987, p. 264).

La realidad de que los distintos elementos que nos constituyen sean el resultado de una historia, no impide que seamos artífices de la misma desde el momento en el que empezamos a incidir en el mundo. Nuevamente, en un escenario donde la idea de alma o “perla cerebral” (Dennett, 1991a) no tenían nada que aportar en términos explicativos, una comprensión naturalista del problema puede buscar pensar de qué manera los individuos desarrollan sus auto-representaciones, consiguiendo equipar los cuerpos que controlan con unos “yoes” responsables.

Sin embargo, esto todavía deja sin responder otra cuestión que planteamos al introducir la teoría de la memética. Si bien podemos ir contra el instinto de supervivencia y reproducción a partir de nuestra capacidad de razonar, parece que en gran medida no somos los responsables de nuestros razonamientos, ya que los mismos tienen lugar a partir de una selección de memes. En función de esto habíamos planteado la pregunta con respecto a quién manda: ¿nosotros o nuestros memes? Si bien la teoría memética en su formulación más esquemática es quizá una de las partes menos interesantes del planteo de Dennett sobre la conciencia, y debería ser, como mínimo, objeto de ulteriores refinamientos (como parece reconocer él mismo en Dennett, 2006), la respuesta que brinda el filósofo a esta pregunta tiene implicaciones interesantes para el tema que nos convoca. Lo que él dice es que es preciso recordar que los memes de algún modo reestructuran los cerebros humanos (Dennett, 1991a). Es por eso que son tan diferentes las mentes francesas y las chinas, las alfabetizadas y las analfabetas, porque, en definitiva, las diferencias más notables entre cerebros humanos dependen de modificaciones microestructurales que introdujeron sus memes. Ahora bien, si es verdad que nuestras mentes fueron modificadas (en algún punto creadas) por memes, ya no se puede pensar en términos de memes contra nosotros, dado que muchos de ellos han determinado qué seamos o quiénes somos. A lo sumo, este esquema permite distinguir entre el imperativo de los genes y el de los memes, pero sería estúpido

pensar que uno de ellos se ajusta mejor a lo que *realmente somos*. No hay una esencia de fondo que sea más fundamentalmente “eso que somos”: somos al resultado de un miríada de adquisiciones de distinto tipo, lo que por otro lado, no nos condena de ninguna manera a un destino en particular.

En un razonamiento análogo al del párrafo anterior, que tampoco abordaré aquí en profundidad, en el contexto de una discusión acerca de la responsabilidad sobre acciones motoras en casos en los que no se puede rastrear un reporte de la decisión consciente de realizarlas⁵⁵, Dennett señala que la cuestión de la responsabilidad está vinculada a qué cosas pensamos que se pueden atribuir al “yo”: “Si te haces a vos mismo muy pequeño, puedes externalizar prácticamente todo” (Dennett, 2003b, p. 41, mi traducción). La propuesta de Dennett es reconocer como propios tanto los comportamientos a los que tenemos acceso consciente, como los que no, lo que heredamos genéticamente y lo que heredamos culturalmente. Todos ellos serían rasgos que pertenecen a cada individuo, y de los que, en alguna medida, gracias a las capacidades adquiridas por el desarrollo filo y ontogenético, el individuo puede “hacerse responsable”, en la medida en que puede, en principio, intervenir activamente sobre ellos. A partir de esto, podemos decir que resulta particularmente difícil en el caso del ser humano hablar de una naturaleza última distinta de sus agregados accidentales.

Para concluir este capítulo me gustaría señalar tres aspectos que se desprenden de esta concepción naturalista compleja de las mentes y los “yoes” humanos. En primer lugar, la ya mencionada idea de que hay que entender a la cultura y a las herramientas, que solemos ver como un producto de la naturaleza humana más bien como algo que la constituyen. Es por eso que no tendría sentido intentar dejar de lado la cultura para ver cómo es realmente la mente humana; más bien debería pensarse que sin cultura no hay mente humana. Así, cobraría sentido hablar en Dennett de una constitución híbrida, biológica y no biológica, de lo humano. Con esto no nos referimos a que somos poseedores de una dimensión que no es, en última instancia, biológica; sino a que es necesario un abordaje desde distintos niveles explicativos, entendiendo que algunos aspectos de lo humanos serán

⁵⁵ También se incluyen experimentos sobre acciones motoras en donde hay un reporte de decisión consciente, pero experimentalmente se ha determinado que el momento en el que, según el reporte, se toma la decisión, es posterior a la señal cerebral para ejecutarlo (Ver Libet *et all*, 1983).

mejor explicados desde los términos y principios de la biología misma, mientras que hay otros que han adquirido una dinámica autónoma específicamente cultural y que la distinción entre ambos es inestables, variable en cada individuo y difícil de establecer.

Ya mencionamos en capítulos anteriores que, en algún punto, el “yo” es lo que es por el uso que le damos. Como dijimos, en uno de los artículos recientes, Dennett (2003b) habla del “yo” como un artefacto. Una de las funciones de este nuevo artefacto sería la de lograr un mayor autocontrol, que es posible gracias a las objetivaciones lingüísticas de nuestros propios estados mentales, que contribuyen al proceso de adquirir la capacidad de confrontar las propias creencias y razones para la acción. El artefacto del que estamos hablando, esto es, la ilusión del usuario a la que da lugar el conjunto de narraciones sobre nosotros mismos, nos permite vincularnos con el ambiente y con los demás de modos que serían imposibles en su ausencia. En palabras de Clark: “No somos más que una bolsa de herramientas cognitivas de usuarios menores (algunas naturales, algunas artificiales) reunidas por un tipo de ilusión de mismidad: una ilusión enraizada en la puesta en marcha de nuestra capacidad de hilar narrativamente, adquirida por cortesía de nuestra destreza y nuestro lenguaje” (Clark, 2002, p. 194, mi traducción). Por otra parte, entender al “yo” como un artefacto nos permite concebirlo como algo que no tiene garantizada su estabilidad *a priori*. Más bien está, como los demás artefactos, “sujeto a repentinos cambios de estatuto” (Dennett 1991a, p. 434). Por eso el “yo” es en algún punto un concepto que implica una inestabilidad mucho mayor a la que viene asociada con la idea de alma o de sustancia mental, porque lejos de ser inmutable está en constante transformación. Además, cuenta con la fragilidad de toda construcción humana. Si bien es una construcción sólida para sus propósitos funcionales, en circunstancias más o menos estables, no es indestructible.

Por último, podemos pensar en otro de los cambios que implica haber convertido al alma en un “yo” naturalizado: eso que los filósofos tradicionales creyeron que estaba ahí porque lo había insuflado Dios, y que apareció como misterioso e inexplicable hasta que contamos tanto con la teoría darwiniana de la evolución, como con la posibilidad de interpretar sus alcances de un modo no reduccionista. Creo que lo que intentamos dilucidar en este capítulo son las tensiones propias de entender a los seres humanos como resultado de un largo proceso

evolutivo. La idea de que componentes tan importantes de nuestra constitución ya nos vienen dados, parece atentar contra nuestra capacidad de decidir quiénes somos. Una visión de lo humano comprometida con los resultados de investigaciones empíricas señala que gran parte de nuestro repertorio conductual está basado en preferencias instintivas, comportamientos imitativos y enseñanzas culturales. Lejos de negar la relevancia de estos rasgos heredados, la perspectiva de Dennett nos permite sostener que dicha herencia no opera en nosotros como una fatalidad; antes bien, hace posible que empecemos a operar en el mundo como nunca podríamos haberlo hecho si no contáramos con ella.

Así, podemos decir que en esta teoría en la que se explican de un modo naturalista algunas importantes diferencias entre las mentes humanas y las animales, la perspectiva que tenemos sobre las primeras se ve enriquecida no sólo por los aspectos que se señalan como excepcionales, sino también por aquellos puntos en los que se presentan continuidades con las mentes de otros animales. Si no contáramos con las capacidades adaptativas de criaturas darwinianas, skinnerianas y popperianas, sería imposible explicar nuestras capacidades de pensamiento abstracto y autocontrol moralmente significativo, propios de las criaturas gregorianas.

Creo que en algún punto esto invita a una reconversión de los criterios de valoración que todavía subyacen a muchas de nuestras intuiciones sobre estos temas. Los aspectos que heredamos biológica y culturalmente no se oponen a aquellas instancias en las que pensamos reflexivamente cómo comportarnos, sino que constituyen la base necesaria para que tales procesos sean posibles y nos brindan los materiales con los cuales podemos llevarlos a cabo. Así, el hecho de que nuestra capacidad para responder ante ciertas circunstancias tenga que ver con ciertas condiciones previas que nos hicieron llegar hasta ese momento, puede ser objeto de importantes consideraciones sociológicas y políticas, pero no nos condena a un determinismo filosófico. Lo mismo sucede con nuestras creaciones. La autoría no resulta más “legítima” por partir de la nada. Esta idea de creación *ex nihilo* tiene un merecido lugar en los relatos teológicos, pero nada que hacer cuando se trata de dar cuenta del modo de comportarse de las criaturas naturales. Nuestra capacidad para intervenir en forma activa y (al menos algunas veces) reflexivamente en la historia

que nos conforma, no es algo que surge de un momento a otro, sino que tiene una pesada historia que no sólo la antecede, sino que la hace posible.

Retomando la frase de Dennett citada en el epígrafe de este apartado, en la que se pregunta por qué la importancia o excelencia de algo debe venir desde las alturas, como un regalo de Dios, podemos decir que la explicación brindada por el filósofo constituye un buen ejemplo de una teoría que no apela a regalos divinos ni ganchos celestes, pero que no por eso convierte a sus objetos de análisis en fenómenos menos excelentes; antes bien, al menos desde mi punto de vista, se ven mucho más enriquecidos. Así, a pesar de que no hay una diferencia esencial entre las creencias de los humanos y las de animales, y que no hay hechos profundos que puedan distinguir nuestra intencionalidad de la de los meros artefactos, los poderes transformadores de las herramientas mentales incubadas culturalmente hacen posible la creación de un “yo” “experimentante” y responsable. Lo cual, a su vez, implica que, en un sentido no metafísico, pero quizá más profundo, nuestras vidas mentales sean más interesantes.

Conclusiones

En este trabajo he pretendido evaluar la potencialidad de la concepción narrativista del “yo” de Dennett a la hora de brindar una explicación de las personas y de la mente humana que no sólo sea compatible con un enfoque naturalista darwiniano, sino que también tome en cuenta los rasgos que la psicología popular reconoce como distintivamente humanos. Teniendo en cuenta que, analizada de forma aislada, la propuesta dennettiana del “yo” resulta poco clara (motivo por el cual ha recibido críticas de distintos frentes) he defendido la idea de que al poner en relación dicha propuesta con otros aspectos de la teoría de Dennett la misma se puede comprender mejor, constituyendo una teoría coherente y contundente. Considero que, en el desarrollo de estos vínculos, la teoría del “yo” narrativo ha mostrado ser capaz, tanto de responder a las exigencias de un enfoque darwiniano, como de no presentar un concepto de “yo” tan deflacionado que no preserve ciertas intuiciones *folk* que tenemos sobre dicho fenómeno. Es más, creo que dicha teoría revela algunas virtudes que derivan de adoptar un abordaje evolutivo y narrativista del “yo”, constituyendo un ejemplo a favor de la idea de que una buena explicación naturalista de nuestro “yo” ha de apoyarse en una concepción híbrida (biológica y no-biológica) de los seres humanos. En lo que sigue recapitularé los corolarios a los que he arribado a lo largo de este trabajo y que apoyan estas conclusiones.

Con respecto a la teoría de la identidad personal de Dennett, he presentado una reconstrucción según la cual nuestros “yoes” se encuentran constituidos tanto por aquellas narraciones que elaboramos nosotros mismos, como por las que elaboran otros. Sin embargo, se ha aclarado que, de las narraciones que elaboran otros, sólo se deberían considerar a las que nos afectan, ya sea porque las aceptamos o las negamos, y de ese modo alteran nuestra auto-percepción o nuestro comportamiento. Además, he defendido que el “yo” como centro de gravedad narrativo debe ser entendido como un *Abstracta* en el sentido reichenbachiano, lo cual, permite entender al “yo” como algo que, por una parte, no se puede identificar con un objeto físico con masa o composición química, pero, por otra parte, tampoco es una abstracción como el símbolo π , sino una abstracción capaz de nacer, morir y sufrir modificaciones (dado que depende de los *Concreta*). Con respecto a los tópicos

y características de las narraciones que conforman el “yo”, he defendido que las mismas no se limitan a los relatos que se pueden formular desde la actitud intencional y desde la actitud personal, sino que cuentan con una riqueza adicional (brindada por añadidos históricos, afectivos, estéticos, culturales, geográficos, etc.), responsable de que, con respecto a estos conjuntos de narraciones, no sólo evaluemos la racionalidad, o la aptitud moral, sino que experimentamos sentimientos estéticos, mayor o menor interés, distintos grados de empatía, etc. En consonancia con esto, he defendido que si recuperamos la distinción dennettiana entre las tres clases de psicología intencional, debemos pensar que el fenómeno del “yo” que Dennett está explicando pertenece, propiamente, al ámbito de la psicología *folk*. Esta tesis también cobra fuerza en virtud de que el “yo” es, precisamente, un dispositivo “creado para” hacer posible esa interacción cotidiana. Así, en alguna medida, la respuesta de Dennett a qué tipo de cosa es el “yo” propone un cambio en el eje de la investigación, en la medida en que responder *qué es el “yo”*, implicará responder *para qué se lo usa*.

Si bien el Modelo de las Versiones Múltiples de Dennett sostiene que nuestra ilusión de que la conciencia funciona según un procesamiento lineal controlado por una entidad única (el “yo”) es falaz, su concepción de la conciencia permite, por otra parte, explicar por qué tenemos esa ilusión y para qué sirve. Dicha ilusión es un punto de autorepresentación en torno al cual nuestro cerebro organiza toda la información que tiene sobre nosotros. Así, aunque los fragmentos narrativos que constituyen nuestra conciencia en realidad provengan de distintos módulos o circuitos, estos aparecen como siendo producidos por un único agente. Esto se logra gracias a las formulaciones lingüísticas, de otros y propias, en las que atribuimos eventos y acciones a un sujeto unificado. Es por esto que aunque parezca que nosotros creamos los relatos que proferimos, para Dennett sería más preciso decir que ellos nos crean a nosotros.

La propuesta presentada hasta aquí es desarrollada por Dennett desde una perspectiva en la que se intenta dar una explicación de cómo pueden haber evolucionado la conciencia y el “yo”. En el capítulo dedicado a exponer esta explicación evolutiva, espero haber demostrado que la misma no cae en los problemas por los que se suele objetar la adopción de este tipo de explicaciones. Por un lado, no constituye una teoría enmarcada en un paradigma ‘pan-adaptacionista’.

Por otro lado, por el modo en el que considera e incorpora la evolución cultural, no cae tampoco en el error de querer explicar aspectos de la cultura humana recurriendo a una idea de selección de genotipos que aumentan el 'fitness'. Así, aunque las consideraciones evolutivas dennettianas sobre la conciencia y el "yo" deben estar sujetas a revisión y están lejos de ser infalibles, permiten dar cuenta de fenómenos de complejidad elevada sin proponer la necesidad de una reducción explicativa a los términos de la biología.

A su vez, otra ventaja de dicha perspectiva evolucionista es que la misma conduce a un cambio en el modo de pensar ciertos fenómenos. Los razonamientos analíticos que nos hacen entender los fenómenos en términos de *todo o nada* se ven desdibujados ante la evidencia de la historia natural, la cual muestra el carácter procesual y de límites difusos que han llevado a las cosas a ser lo que son. Así, un paradigma darwiniano permitirá entender mejor fenómenos complejos como el caso de los Trastornos de Identidad Múltiple.

Por otra parte, la especie humana no está ubicada, en esta historia, en un pedestal con respecto a las demás, en la medida en que no se asume que por definición sus rasgos y características van a ser de una naturaleza completamente diferente a los rasgos y características de otros animales avanzados. Así, se presenta al "yo" narrativo como parte del fenotipo extendido de la especie humana. El modo en el que construimos nuestros "yoes" es comparado con el modo en que una araña teje su tela, y un pájaro construye su glorieta; comparaciones en las que la actividad de elaborar narraciones parece ser de carácter exclusivamente instintivo. Este aspecto tornaría la propuesta del "yo" de Dennett sumamente anti-intuitiva, sino fuera porque puede ser complementado con el vínculo que el filósofo establece entre lenguaje e inteligencia. He defendido que una reconstrucción de dicho vínculo, en diálogo con otros aspectos de la teoría de Dennett que ya se habían desarrollado, permite mostrar que el filósofo norteamericano consigue explicar cómo la adquisición de herramientas lingüísticas que se produce con el ingreso a un lenguaje natural y de intercambios lingüísticos dentro de una cultura, nos permiten ser portadores de capacidades específicamente humanas de i) abstracción, comprensión y aprendizaje, y ii) autocontrol moralmente significativo. Con esto, he rescatado la potencialidad de la idea de Dennett de que la capacidad de tratar los propios pensamientos como objetos habilita posibilidades de intervención deliberada frente a

los mismos. Así, el “yo” puede ser entendido como un artefacto, que tiene como una de sus funciones la de operar como un foco de autocontrol y de toma de decisiones. En ello se cifra la potencialidad de la emergencia de un “yo” responsable. Y es por eso que los “yoes” tienen proyectos, intereses y valores, que crean en el curso del desarrollo de sus habilidades autocontroladas y autoevaluadas.

Finalmente, he defendido que de esta concepción del “yo” humano como producto de la evolución biológica y cultural se siguen otros tres aspectos que pueden ser presentados como fortalezas de la teoría de Dennett, que son consecuencia de intentar pensar la especificidad humana desde un marco de continuidad con el resto de las especies. Uno de ellos es relativo a que el “yo”, como producto de una cierta actividad humana, pasa a ser entendido como un artefacto, lo que nos permite concebirlo como algo que no tiene garantizada su estabilidad a priori. Antes bien, el mismo está, como los demás artefactos “sujeto a repentinos cambios de estatuto” (Dennett 1991a, p. 434). Por eso el “yo” es en algún punto un concepto que implica una inestabilidad mucho mayor a la que viene asociada con la idea de alma o de sustancia mental, porque lejos de ser inmutable está en constante transformación. Además, cuenta con la fragilidad de toda construcción humana. Si bien es una construcción sólida para sus propósitos funcionales, en circunstancias más o menos estables, no es un fenómeno indestructible.

En segundo lugar, desde esta forma de entender la evolución cultural y el rol del lenguaje en la inteligencia, no tiene sentido intentar dejar de lado la cultura para ver cómo es “realmente” la mente humana; dado que, desde el enfoque de Dennett, sin cultura no hay mente humana. Así, cobraría sentido hablar, para este autor, de una constitución híbrida, biológica y no-biológica, de lo humano. Con esto no nos referimos a que somos poseedores de una dimensión que no es, en última instancia, biológica; sino a que es necesario un abordaje desde distintos niveles explicativos, entendiendo que algunos aspectos humanos serán mejor explicados desde los términos y principios de la biología misma, mientras que hay otros que han adquirido una dinámica autónoma específicamente cultural.

Finalmente, defendí también que la tensión que intenté dilucidar en el último capítulo, entre una perspectiva evolutiva que nos presentaba como una especie con muchas conductas de carácter instintivo, y nuestra arraigada creencia según la cual somos capaces de tener comportamientos deliberados, es producto de confrontar

cierta idea *folk* sobre lo que es ser un agente autónomo, con una visión que nos muestra como resultado de un proceso evolutivo, con una cantidad de características heredadas. Pienso que la visión de Dennett permite reformular esa concepción idealizada de agencia, que vincula la autonomía con una idea de “partir desde cero”. La perspectiva dennettiana no sólo considera que es posible sostener que dicha herencia (biológica y cultural) no opera en nosotros como fatalidad, sino que explica cómo es gracias a ella que podemos adquirir la capacidad de operar de forma deliberada (incluso en contra de lo que los instintos y enseñanzas heredados mandan). Así, podemos decir que en esta teoría en la que se explican de un modo naturalista algunas importantes diferencias entre las mentes humanas y las animales, la perspectiva que tenemos sobre las primeras se ve enriquecida no sólo por los aspectos que se señalan como excepcionales, sino también por aquellos puntos en los que se presentan continuidades con las mentes de otros animales. Desde esta visión, no habría diferencias esenciales entre los humanos y el resto de las especies y, sin embargo, se puede defender que la humana -gracias a que los poderes transformadores de las herramientas mentales incubadas culturalmente hacen posible la creación de un “yo” “experimentante” y responsable- está dotada de una vida mental mucho más interesante.

Si bien la capacidad de poseer un tipo de autocontrol reflexivo que resulta moralmente significativo está vinculada con la cuestión del libre albedrío, este último tema no fue abordado de manera directa en este trabajo. El concepto de libre albedrío podría considerarse un problema para el materialismo de Dennett apoyado en un universo determinista de las ciencias físicas según el cual, como explicamos en el primer capítulo, todos los fenómenos que se explican desde la actitud intencional deberían poder explicarse desde la actitud física. Si bien realizar este cambio de nivel explicativo (de la actitud intencional a la actitud física) para explicar las conductas de los agentes intencionales implicaría una pérdida (del “patrón” o “modelo” de comportamiento que intentamos identificar), y no sería posible para las capacidades cognitivas humanas, sería incoherente suponer que no hay un correlato físico de aquellos *Concreta* que “abreviamos” con los *Abstracta* de las creencias y deseos. Para defender esta postura, Dennett elabora una bomba de intuición en la que unos marcianos super inteligentes, con acceso los eventos que ocurren en todo nuestro sistema nervioso, podrían predecir todos nuestros comportamientos. Ahora bien, si

un marciano super inteligente puede, por conocer absolutamente todos los estados y movimientos del sistema nervioso de alguien predecir con éxito y precisión todas las acciones que realiza esa persona, ¿no se está negando la posibilidad del libre albedrío? Podemos adelantar que Dennett señala que una ontología que sea estocástica en lugar de determinista no resolvería el problema. Sin embargo, pienso que esta pregunta queda abierta y pienso que es una de las líneas a seguir en futuras investigaciones, para entender mejor la concepción dennettiana de la mente y el “yo”.

Por otra parte, pienso que la intuición dennettiana de otorgarle a las narraciones un lugar fundamental en la creación social de los “yoes” se ha visto constatada por el formato de las redes sociales más populares. Dichas redes sociales están estructuradas en plataformas que, en general, tienen un correlato con un “yo” real (a diferencia de pertenecer por ejemplo, a grupos de personas, o anónimos) y en las que una gran cantidad de recursos (lingüísticos, visuales, audiovisuales) están destinados a caracterizar a este “yo” singular. Por otra parte, otro aspecto llamativo del fenómeno es que dichas redes sociales juegan un rol importante en la vida cotidiana de muchas personas; en otras palabras, lo que “sucede” ahí, les afecta. En función de esto, pienso que también se podrían realizar futuras investigaciones orientadas a especificar y actualizar una teoría narrativista del “yo” teniendo en cuenta las nuevas tecnologías de las redes sociales que acarrearán nuevas formas de narrarse.

En un libro que acaba de ser publicado, Dennett (2017) contrapone dos actitudes posibles a la hora de examinar filosóficamente fenómenos mentales: la actitud científica y la actitud cartesiana. En dicha contraposición, presenta la idea de que existen fuerzas de distorsión de la imaginación que muchas veces impiden los avances de las investigaciones científicas y filosóficas en estos campos. Como es de esperar, según Dennett, la actitud cartesiana es la que incurre en una visión errónea. Para ilustrar el tipo de movimiento que quiere promover con su nuevo libro, introduce una metáfora en la que llama a la más original de estas fuerzas de distorsión “gravedad cartesiana” y que buscará contrarrestar provisto de poderosas herramientas para incitar al lector a dirigirse en la dirección contraria. Se trata del mismo concepto, con uso metafórico, que empleó Dennett en los escritos que analizamos en este trabajo, sobre la identidad personal y el yo, sólo que su contenido va en la dirección "gravitacional" contraria. Con respecto a dicha fuerza dice que:

La gravedad cartesiana, a diferencia de la gravedad de la física, no actúa sobre las cosas en proporción a su masa y proximidad a otros elementos masivos; actúa sobre las ideas o representaciones de las cosas en proporción a su proximidad en cuanto al contenido de otras ideas que desempeñan roles privilegiados en el mantenimiento de los seres vivos (...) La idea de la gravedad cartesiana, tal como se ha presentado hasta ahora, es simplemente una metáfora, pero el fenómeno que yo llamo por este nombre metafórico es perfectamente real, una fuerza disruptiva que molesta (y a veces ayuda) nuestra imaginación, y a diferencia de la gravedad de la física, es en sí mismo un fenómeno evolucionado. Para entenderlo, necesitamos preguntar cómo y por qué surgió en el planeta Tierra. (Dennett, 2017, pp. 21, 22).

Con estas afirmaciones, Dennett vuelve a sostener que una perspectiva filosófica interesante no sólo debe intentar eliminar aquellas perspectivas y explicaciones o concepciones que contienen elementos que podemos considerar falsos, sino comprender cuál es el rol que desempeñan para nosotros, qué aspectos de las mismas resultan relevantes o irrenunciables y por qué motivos. Este modo de vincular explicaciones científicas y concepciones *folk* no tiene como horizonte un reduccionismo explicativo, sino al contrario, defiende la importancia de que convivan distintos niveles de explicación. Espero que este trabajo haya podido dar cuenta de la importancia de abordar temas tan relevantes como el de la identidad personal desde una perspectiva filosófica que comprenda que éste y otros fenómenos emparentados, pueden y deben ser abordados explicativamente en distintos niveles.

Bibliografía

- American Psychiatric Association, & American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed., text revision.). Washington: American Psychiatric Association.
- Brandom, R. (1994) *Making it explicit. Reasoning, representing and discursive commitment*, Cambridge: Harvard University Press.
- Brentano, F. (1874), *Psychology from an Empirical Standpoint*, London: Routledge, 1995.
- Brook, A., Ross, D. (2002) "Dennett's position in the Intellectual World" en A. Brook and D. Ross (eds) *Daniel Dennett*, Cambridge: Cambridge University Press, pp. 3-37.
- Burke, P. (1997) "Representations of the Self from Petrarch to Descartes" en Roy Porter (ed) *Rewriting the Self: Histories from the Renaissance to the Present*, London: Routledge, pp. 17-28.
- Carr, D. (1998) "Phenomenology and Fiction in Dennett", *International Journal of Philosophical Studies*, 6, pp. 331-344.
- Churchland (1986) *Neurophilosophy: Toward a Unified theory of Mind/Brain*, Cambridge: The MIT Press.
- Clark, A. (2002) "That Special Something: Dennett on the Making of Minds and Selves" en A. Brook and D. Ross (eds) *Daniel Dennett*, Cambridge: Cambridge University Press, pp. 187- 205.
- Clark, A., Karmiloff-Smith, A. (1993) "The Cognizer's Innards", *Mind and Language*, 8(4), pp. 540-547.
- Corcuera, A. (2013) "Problemas y aciertos de la teoría del Yo narrativo de Dennett: aportaciones al debate sobre la identidad personal", *Logos. Anales del Seminario de metafísica*, 46, pp. 27-45.
- Davidson, D. (1967) "Truth and meaning", en *Inquiries into Truth and Interpretation*, Oxford: Oxford Clarendon Press, 1984, pp. 17-3.
- Davidson, D. (1973) "Radical Interpretation", en *Inquiries into Truth and Interpretation*, Oxford: Oxford Clarendon Press, 1984, pp. 125-139.
- Davidson, D. (1975) "Thought and Talk", en *Inquiries into Truth and Interpretation*, Oxford: Oxford Clarendon Press, 1984, pp. 155-170.
- Davidson (1988) "What is present to the mind?", en J. Brandl y W. Gombos (eds.) *The mind of Donald Davidson*, Amsterdam: Atlantta.

- Dawkins, R. (1976) *El gen egoísta: las bases biológicas de nuestra conducta*, Barcelona: Salvat, (1994).
- Dawkins, R. (1982) *The Extended Phenotype: The Gene as the Unit of Selection*, San Francisco: Freeman.
- Dennett, D. (1969) *Contenido y Conciencia*, Barcelona: Gedisa, 1994.
- Dennett, D. (1971) "Sistemas Intencionales" Cuadernos de Crítica, #40, Distrito Federal: UNAM Ed., 1985.
- Dennett, D. (1973) "Mecanicismo y responsabilidad," *Cuadernos de Crítica*, #42, Distrito Federal: UNAM Ed., 1985.
- Dennett, D. (1974a) "Coment on Wilfred Sellars", *Synthese*, 27(3-4), pp: 439-44.
- Dennett, D. (1974b) "Why the law of effect will not go away", *Journal of the Theory of Social Behaviour*, 5, pp. 169-187.
- Dennett, D. (1976): "Condiciones de la cualidad de ser persona", *Cuadernos de crítica*, #45, Distrito Federal: UNAM Ed., 1989.
- Dennett, D (1984) *Elbow Room: The Varieties of Free Will Worth Wanting*, Cambridge: The MIT Press.
- Dennett, D. (1987) *La actitud intencional*, Barcelona: Gedisa, 1998.
- Dennett, D. (1988): "Why Everyone is a Novelist", *The Times Literary Supplement*, London: News UK, 4, pp. 16-22.
- Dennett, D (1989) "The Origin of Selves", *Cogito*, 3, pp. 163-173.
- Dennett, D. (1991a) *La conciencia explicada*, Barcelona: Paidós, 1995.
- Dennett, D. (1991b) "Real Patterns", *Journal of Philosophy*, 88, pp. 27-5.
- Dennett, D. (1992): "The Self as the Center of Narrative Gravity," in F. Kessel, P. Cole and D. Johnson (eds.) *Self and Consciousness: Multiple Perspectives*, Hillsdale: Erlbaum.
- Dennett, D. (1993) "Learning and labeling" *Mind and Language*, 8, pp. 540-548.
- Dennett, D (1994) "The Role of Language in Intelligence," en Jean Khalifa (ed.) *What is Intelligence?*, *The Darwin College Lectures*, Cambridge: Cambridge Univ. Press., pp. 161-78.
- Dennett, D. (1995) *Darwin's dangerous idea*, Londres: Penguin Books.
- Dennett, D. (1996) *Tipos de Mente*, Barcelona: Ed. Debate pensamiento, 2000.

- Dennett, D. (1998) *Brainchildren: Essays on Designing Minds*, Cambridge: The MIT Press.
- Dennett (2001a) "Collision, Detection, Muselot, and Scribble: Some Reflections on Creativity" en D. Cope (Ed.) *Virtual Music, Computer Synthesis of Musical Style*, Cambridge: The MIT Pres, pp. 283-291.
- Dennett, D. (2001b) "In Darwin's Wake, Where Am I?, *Proceedings and Addresses of The American Philosophical Association*, 75 (2), pp. 13-30.
- Dennett (2003a) *Freedom Evolves*, London: Allen Lane Publishers, e impreso por Penguin Books.
- Dennett (2003b) "The Self as a Responding—and Responsible—Artifact," *Annals New York Academy of Sciences*, 1001, pp. 39-50.
- Dennett, D. (2005) *Dulces sueños: Obstáculos filosóficos para una ciencia de la conciencia*, Buenos Aires, Katz, 2006.
- Dennett, D. (2006) "From Typo to Thinko: When Evolution Graduated to Semantic Norms", en P. Jaisson and S. Levinson (eds.) *Evolution and Culture*, Cambridge: The MIT Press., pp. 133-45.
- Dennett, D. (2017) *From Bacteria to Bach and Back: The Evolution of Minds*, New York: W. W. Norton & Company.
- Dennett D., Humphrey N. (1989) "Speaking for Our Selves: An Assessment of Multiple Personality Disorder", *Raritan: A Quarterly Review*, 9, pp. 68-98.
- Di Francesco, M., Marraffa, M., Paternoster, A. (2014): "Real selves? Subjectivity and the subpersonal mind", *Phenomenology and Mind*, 7, pp. 90-101.
- Dretske, F. (1985) *Knowledge and the Flow of Information*, Cambridge: The MIT Press.
- Dupré, J. (2003) *El legado de Darwin. Qué significa hoy la evolución*, Buenos Aires: Katz, 2006.
- Fodor, J. (1985) "Fodor's Guide to Mental Representation: The intelligent auntie's vade-mecum", *Mind*, 94 (373), pp. 76-100.
- Fodor, J. (1987) *Psychosemantics*, Cambridge: The MIT Press.
- Gabora, L. (2005) "Creative Thought as a Non-Darwinian Evolutionary Process", *Journal of Creative Behavior*, 39(4), pp. 65-87.
- Gazzaniga, M. (1978) "Is seen believing: Notes on clinical recovery" en L. Finger (Comp.) *Recovering from Brain Damage: Research and Theory*, Nueva York: Plenum Press, pp. 409-414.
- Gomila, A. (2010) "Evolutionary psychology and the proper relationship between ontogeny and phylogeny" en L. Pérez-Miranda y A. Izaguirre (Eds.)

Advances in Cognitive Science: Learning, Evolution and Social Action, Bilbao: University of the Basque Country Press, pp. 233–252

- Gould, S. (1980) *The Panda's Thumb*, Nueva York: W.W. Norton.
- Gould, S., Lewontin, R. (1979) “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme”, *Proceedings of the Royal Society of London B: Biological Sciences*, 205(1161), pp. 581-598.
- Gould, S., Lewontin, R. (1983) “La adaptación biológica”, *Mundo científico*, 3 (22): web. 20 de enero (2013).
- Gould, S., Vrba, E. (1982): “Exaptation: A Missing Term in the Science of Form”, *Paleobiology*, 8(1), pp. 4-15.
- Gregory, R. (1981), *Mind in Science: A History of Explanations in Psychology and Physics*, Cambridge: Cambridge University Press.
- Hofstadter, D. (1985). “On the seeming paradox of mechanizing creativity”, en *Metamagical Themas*, Nueva York: Basic Books, pp. 526-546.
- Hofstadter D, Dennett, D. (1981) *El ojo de la mente: Fantasías y reflexiones sobre el Yo y el alma*, Buenos Aires: Sudamericana, 1983.
- Hume, D. (1739) *Tratado de la naturaleza humana*, Madrid, Editora Nacional, 1980-1981 (2 tomos).
- Hume, D. (1979) *Diálogos sobre la religión natural*, Madrid: Tecnos, 1994.
- Humphrey, N. (1986) *The Inner Eye*, London: Faber & Faber.
- Humphrey, N. y Dennett, D. (1989) “Speaking for ourselves: An assessment of Multiple Personality Disorder”, *Raritan*, 9, pp. 68-98.
- Hutto (2007) “The Narrative Practice Hypothesis: Origins and Applications of Folk Psychology”, *Royal Institute Philosophy Supplement*, 60, pp. 43-68.
- Hutto (2008) *Folk psychological narratives: the sociocultural basis of understanding reasons completar*, Cambridge: The MIT Press.
- Jaynes, J. (1976) *The origins of Consciousness in the Breakdown of the Bicameral Mind*, Boston:Houghton Mifflin Harcourt.
- Kafka, (1915) *La Metamorfosis*, Buenos Aires: Losada, 1962.
- Karmiloff-Smith, A. (1979) *A Functional Approach to Child Language*, Cambridge: Cambridge University Press.
- Lamarque, P. (2004) “On not expecting too much from narrative”, *Mind and language*, 19(4), pp. 393-408.
- Leslie A. (1992) “Pretense, Autism, and the Theory-of-Mind module”, *Current Directions in Psychological Science*, 1, pp. 18-21.

- Libet, B. (1983) "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential) - The Unconscious Initiation of a Freely Voluntary Act", *Brain*, 106, pp. 623–642.
- Lowe, E. (2000) *Filosofía de la Mente*, Barcelona: Idea Books, 2000.
- Margolis, H. (1987) *Patterns, Thinking, and Cognition*, Chicago: The University of Chicago Press.
- Martin, M. y Barresi, J. (2006) *The Rise and Fall of Soul and Self: An Intellectual History of Personal Identity*, New York: Columbia University Press.
- McDowell, J. (1994) *Mind and World*, Cambridge: Harvard University Press.
- Menary R. 2008 "Embodied narratives", *Journal of Consciousness Studies*, 15 (6), pp. 63-84.
- Miller G. (1983) "Informavores", *The Study of Information: Interdisciplinary Messages*, New York: Wiley-Interscience, pp.111–113.
- Nagel T. (1974) "What it is like to be a bat?", *Philosophical Review*, 83, pp. 435- 450.
- Pérez, D. I. (2002): "Naturalismo y mente humana. De las intrincadas relaciones entre la filosofía, la ciencia, el sentido común, el alma humana y la madre naturaleza", en Diana Pérez (ed.) *Los caminos del naturalismo*, Buenos Aires: Eudeba.
- Platón, "Fedón", *Diálogos: Fedón; Banquete; Fedro*, traducción de Lledó, E., Gual, C. G., & Hernández, Madrid: Gredos, 1992.
- Platón, "Teeteto", *Diálogos: Parménides; Teeteto; Sofista; Político*, traducción de A. Vallejo Campos, Madrid: Gredos, 1988, 1992 (1ª rempr.).
- Platón, "Banquete", *Diálogos: Fedón; Banquete; Fedro*, traducción de Lledó, E., Gual, C. G., & Hernández, Madrid: Gredos, 1992.
- Popper y Eccles (1977) *El yo y su cerebro*, Barcelona: Labor, 1982
- Quine, W. O. (1950) "Dos dogmas del empirismo", en *Desde un punto de vista lógico*, Barcelona: Ed. Ariel, 1962.
- Quine, W. (1960) *Word and Object*, Cambridge: The MIT Press.
- Quine, W. O., (1969) "Epistemología Naturalizada", en *La relatividad ontológica y otros ensayos*, Madrid: Tecnos, 1984.
- Reichenbach (1938) *Experiencia y predicción: Un análisis de los fundamentos y la estructura del conocimiento*, Chicago: University of Chicago Press, 1970.
- Ryle, G. (1979) *On Thinking*, Edición a cargo de Kolenda, K., Totowa: Rowman and Littlefield

- Sánchez, A. (2014) “La concepción del Yo en Daniel Dennett: Un análisis de la relación entre la perspectiva heterofenomenológica y el enfoque memético” *Logos: Revista de lingüística, filosofía y literatura*, Vol 24(1).
- Schaeffer (2007), *El fin de la excepción humana*, Buenos Aires: Fondo de Cultura Económica.
- Schechtman, M. (1996) *The Constitution of Selves*, Ithaca: Cornell university Press.
- Searle, J. (1980) “Minds, brains, and programs”, *Behavioral and Brain Sciences*, 3(3), pp. 417-424.
- Searle, J. (1983), *Intencionalidad: Un ensayo en filosofía de la mente*, Madrid: Tecnos, 1992.
- Sellars, W. (1962), *Ciencia, percepción y realidad*, Madrid: Tecnos, 1971.
- Selfridge, O. (1959) “Pandemonium: A paradigm for learning”, *Symposium on the Mechanization of Thought Processes*, London: Stationery Office.
- Skinner B. (1964) “Behaviorism at Fifty” en T. W. Wann (comp.) *Behaviorism and Phenomenology: Contrasting bases for Modern Psychology*, Chicago: University of Chicago Press.
- Sperber D. (1996), *Explaining Culture: A Naturalistic Approach*, Oxford: Blackwell.
- Sterenly, K. (2006), “Memes revisited”, *British Journal of the Philosophy Science* 57(1), pp. 145-165.
- Strawson, G. (2005): “Against narrativity”, en G. Strawson (Ed.) *The Self?*, Hoboken: Blackwell Publishing.
- Strawson, G. (2009) *Selves: An essay in revisionary metaphysics*, Oxford: Clarendon Press.
- Taylor, C. (1976) “Responsibility for the Self” en E. Rorty (Ed.) *The identities of persons*, Berkeley: University of California Press.
- Velleman, D. (2005) *Self to self: Selected essays*, Cambridge: Cambridge University Press.
- Vigotsky, L. (1934) *Pensamiento y lenguaje*, Barcelona: Paidós, 1995.
- Voltaire (1759) Morize, André, (Ed.) *Candide: ou, L'optimisme; édition critique avec une introd. et un commentaire par André Morize*. Paris, 1931.
- Williams G. (1966) *Adaptation and Natural Selection. A Critique of Some Current Evolutionary Thought*, Princeton: Princeton University Press.
- Wilson, L. (1974) “Observations and Experiments on the Ethology of the European Beaver”, *Viltrevy, Swedish Wildlife*, 8, pp. 115-266.