

# *El origen de la volición*

ÁLVARO MORENO  
*Universidad del País Vasco*

## I. INTRODUCCIÓN

La volición se suele entender como un tipo de proceso por el cual un agente opera de manera racional, libre y consciente. En este sentido la volición es una de las características más fundamentales de la agencialidad específicamente humana, y ha sido objeto de infinidad de debates filosóficos (y, en tiempos más recientes, también en el ámbito de la psicología y la neurociencia experimental). Pero hay también un sentido más amplio del término volición: es el de aquellas acciones llevadas a cabo por un agente bajo control consciente, como cuando decimos que un bebé o un perro “quieren” o “no quieren” algo, por contraposición a cuando se les fuerza o a cuando reaccionan de manera inmediata y automática, en un acto reflejo.

Tomando el término volición en este sentido amplio, la cuestión que quiero abordar aquí es la volición pre-humana; en particular, si el tipo de agencialidad que manifiestan ciertos vertebrados superiores capaces de lo que Edelman (1992) ha llamado “consciencia primaria”, puede calificarse como una forma elemental de volición. No por supuesto en el sentido de acto racional y susceptible de consideración moral; pero tal vez sí como acción bajo control consciente, como realización de contenidos intencionales cargados de valor, es decir, deseos percibidos como tales y no meramente impulsos inconscientes o fuera de control consciente.

Alguien puede objetar que, independientemente del interés de la cuestión, este análisis guarda poca relación con lo que en el ámbito filosófico normalmente se entiende por volición. Pero creo que, bien al contrario, el estudio de los mecanismos agenciales en los animales superiores, desde una perspectiva evolutiva, permite desentrañar claves importantes para la comprensión de la volición humana. No sólo porque ésta ha evolucionado a partir de aquellas formas de agencialidad, sino porque, como espero demostrar, la relación entre las formas inferiores de agencialidad y la propiamente consciente en los vertebrados superiores puede servirnos de modelo para el estudio de la agencialidad humana.

Desde una perspectiva evolutiva la volición humana se nos presenta como la forma más compleja de control de la conducta. Podemos ver entonces la evolución de la agencialidad culminando en la volición. Por ejemplo, la acción de una medusa, cuyo sistema nervioso (SN) carece de centro integrador, manifiesta un nivel de control muy simple (en este caso, el control es el que ejerce su SN sobre su organización metabólica); mientras que un jaguar acechando a su presa implica un nivel de control agencial mucho más complejo: su córtex centraliza, integra y coordina funcionalmente una enorme cantidad de procesos de niveles inferiores. En este nivel, pues, aparece un tipo de acción caracterizada por un alto grado de control que según todos los indicios de la neurociencia ocurre a nivel consciente –esto es, implica una gestión de representaciones intencionales conscientes. La acción humana consciente constituye por supuesto un grado mucho más elevado de control consciente, una consciencia de la consciencia. Pero podemos así ver la aparición de la volición en el marco de un proceso de progresiva complejización y jerarquización de las formas de control agencial, es decir, como una forma de agencialidad caracterizada por un alto grado de control sobre los procesos agenciales de bajo nivel. En todo este proceso evolutivo hay no sólo saltos sino continuidades, cuya comprensión es importante. La volición específicamente humana tiene sus raíces en estas formas de agencialidad consciente que manifiestan ciertos animales, y por consiguiente, avanzar en la comprensión de los mecanismos que la constituyen y de los factores evolutivos que explican su origen creo que puede ayudar a mejorar nuestra comprensión de aquélla.

Dividiré este artículo en tres partes. En la primera abordaré el concepto de agencialidad, tratando de mostrar cómo la autonomía es el marco fundamental para entenderla. En la segunda, estudiaré los factores que operan en la evolución de la agencialidad hacia formas cada vez más complejas. Y en la tercera analizaré el origen, función y características de la agencialidad consciente en el mundo animal, tratando de mostrar cómo constituye una forma elemental de volición.

## II. ¿QUÉ ES UN AGENTE?

Incluso en su forma elemental, el concepto de volición implica una forma altamente compleja de agencialidad. Por ello mismo, y dada la dificultad de la cuestión, creo que lo más razonable es comenzar analizando el concepto de agencialidad en su forma más básica. Pero tampoco la agencialidad, en su sentido más genérico, es fácil de definir. Dentro del campo de la Robótica Autónoma y la Vida Artificial se ha intentado formular un concepto de agencialidad susceptible de verificación experimental a través de sistemas físicamente construidos y/o

simulados (Moreno & Etxeberria 2005). Según T. Smithers (1997) los agentes son sistemas que pueden iniciar, sostener y mantener un continuo proceso de interacciones con su entorno como una parte esencial de su funcionamiento normal, y para P. Maes (1994) un agente es un sistema que trata de cumplir un conjunto de objetivos en un entorno complejo y dinámico.

Se pueden encontrar muchas definiciones similares en la literatura especializada. Los elementos que en casi todas ellas se destacan son: la actividad del sistema tiene fines u objetivos, el sistema de alguna manera está relacionado o concernido (para su continuidad) con las consecuencias de sus acciones, el sistema tiene una perspectiva propia. En suma, las acciones de un agente no son meras interacciones físicas, sin sentido, pues su continuidad se ve afectada por ellas. Estas interacciones con el entorno se entienden de alguna manera como una prolongación del funcionamiento de la organización del sistema. Por eso, el fundamento de la capacidad agencial reside en la aparición de sistemas cuya organización es capaz de automantenerse modificando activamente su entorno inmediato.

Podríamos pensar que este tipo de enfoques está demasiado sesgado por una visión biológica. ¿Por qué no considerar como agente a un sistema si lo que hace es funcional, aun cuando no lo sea para él mismo? Tal vez no podamos decir que un sistema así sea autónomo, pero sí un verdadero agente. Este planteamiento sin embargo tiene un problema: si el sistema no está concernido por lo que hace, los fines de sus acciones le son ajenos, el verdadero agente es su diseñador o usuario, y el sistema en cuestión no es a la postre sino una prótesis de éstos. Podríamos pensar en otra posibilidad: un sistema capaz de establecer sus propios fines, pero de un tipo que no redundara en acciones favorables (ni directa ni indirectamente) a su mantenimiento. Algo así es en teoría concebible, pero en la práctica tales sistemas sólo podrían aparecer como resultado de la acción de otros agentes, y, a no ser que sus creadores velaran continuamente sobre ellos, estarían condenados a desaparecer. Hay pues una dependencia recíproca entre lo que define al agente y las acciones que se derivan de su ser, el ser del sistema es indisociable de su hacer. Por estas razones, el concepto de agente parece difícilmente separable del de sistema autónomo.

Se han realizado muchos esfuerzos en el campo de la robótica para crear agentes artificiales autónomos, y sin embargo, el resultado hasta la fecha es más bien decepcionante. Nadie ha conseguido emular lo que hace una simple bacteria. ¿Por qué? ¿Qué tiene de especial su organización?

Supongamos una bacteria moviéndose en un medio acuoso. En este medio hay muchas sustancias químicas; algunas de ellas atraviesan la membrana de la bacteria simplemente por factores físicos, por ejemplo, porque su desplazamiento ocurre a favor del gradiente de concentración de las moléculas. Pero en otros muchos casos el transporte ocurre contra gradiente, es decir, los pro-

ductos son literalmente bombeados, y para ello, como es obvio, la bacteria ha de poner en marcha una serie de reacciones acopladas que consumen recursos energéticos. La diferencia entre el primer tipo de transporte y el segundo estriba en que aquél es termodinámicamente espontáneo (es decir, se produce por la acción de las leyes físico-químicas) y éste sólo puede ocurrir por la acción de un mecanismo (que reside en el interior de la bacteria). Ambos procesos pueden ser considerados como “interacciones”, pero sólo el segundo, en la medida en que es explicable por la existencia de un mecanismo, es una interacción causada por el sistema; es decir, es una acción.

¿Qué consecuencias tiene esa acción? Sencillamente, contribuir al auto-mantenimiento de la organización metabólica de la bacteria. Si ésta, a pesar del coste que le supone, no efectuara el bombeo selectivo de sustancias, no podría continuar funcionando. Por ejemplo, sin expulsar protones la bacteria entraría en una crisis osmótica y estallaría. En realidad, prácticamente todo cuanto la bacteria hace (es decir, lo que requiere que la bacteria emplee su organización para alterar sus relaciones con el entorno) es “funcional”, es decir, contribuye a su mantenimiento como individuo y como especie.<sup>1</sup> Si no fuera así, las bacterias (y todos los organismos en general) desaparecerían, pues su organización es lo que los físicos llaman “una estructura disipativa”<sup>2</sup> (Nicolis & Prigogine 1977): su cohesión sólo se mantiene a través de un continuo intercambio de materia y energía con su entorno. Por consiguiente necesitan hacer continuamente determinadas cosas sólo para mantener su organización. Como habíamos dicho antes, sus acciones son efecto de su identidad (su organización específica), pero ésta depende también de sus acciones.

Estas consideraciones nos llevarían a apoyar un concepto de agencialidad mucho más en línea con lo propuesto por autores como S. Kauffman (2000), quien define un agente como un sistema que hace algo para sí mismo. Es decir, la agencialidad mínima estaría ligada a /se derivaría de/ la organización básica de la vida.

Sin embargo, la definición de Kauffman no recoge un aspecto importante de la agencialidad, a saber, la relación entre el agente y las variaciones del medio. Una bacteria como *E.Coli* en un medio acuoso donde existe azúcar es capaz de nadar en dirección a la fuente (es decir, hacia la zona donde existe mayor concentración de dicha sustancia). Evolutivamente la bacteria ha desarrollado un mecanismo de detección y movimiento congruentes con su organización metabólica, de manera que, dentro de cierto rango, es capaz de desencadenar

1 Lo cual, en última instancia, es una forma de asegurar su mantenimiento a largo plazo.

2 Una estructura disipativa es una configuración o patrón de materia que mantiene una cohesión sólo en condiciones alejadas del equilibrio termodinámico.

acciones preventivas cuando detecta condiciones de peligro: la bacteria nada en dirección al azúcar sólo si la necesita. El azúcar tiene pues un *significado* para la bacteria, ya que es una sustancia fundamental para el mantenimiento de su organización. En otras palabras, los seres vivos (SV) operan en un espacio de significados, es decir, en un espacio en el que crean distinciones entre lo que es o no relevante para su continuidad y supervivencia.

Imaginamos en cambio un robot moviéndose en una habitación, captando información visual de su entorno y procesando dicha información en un ordenador que lleva incorporado, enviando señales a sus motores para coordinar su movimiento: el robot evita los obstáculos, busca su objetivo (pongamos, una fuente luminosa para recargar su batería mediante sus paneles solares cuando el nivel de ésta desciende por debajo de cierto umbral), etc. El robot puede ser adaptativo e incluso llegar a aprender a través de la experiencia; puede por ejemplo, encontrar la salida de un laberinto. Sin embargo, como ha señalado Di Paolo (2003), por “inteligente” que pueda ser el robot, no está realmente preocupado por lo que hace: si falla en sus operaciones no mostrará ningún síntoma de preocupación: los fallos no afectan a su estructura, porque los objetivos buscados no lo son por el robot, sino por sus diseñadores.

En esto se diferencian de forma radical los SV de los pretendidos agentes artificiales (como los robots) pues las acciones que generan éstos son sólo funcionales para los que los construyen o utilizan. En cambio, como la estructura de los SV –el metabolismo– es un continuo proceso de auto-fabricación (es decir, son organizaciones autopoieticas), sus relaciones con el entorno son, en su forma básica, la prolongación de este hecho: intercambio selectivo de materia y energía que asegura el mantenimiento de la organización metabólica. La actividad del sistema se convierte en algo intrínsecamente relevante para el mismo, pues su existencia depende de los efectos de su propia actividad.

Esta intrínseca relevancia genera un criterio para determinar lo que el sistema “debe” hacer. La organización como un todo “debe” ocurrir de una determinada manera, en el sentido de que, si no fuera así, el sistema se desintegraría. Por consiguiente, las condiciones de existencia del sistema son las normas de su propia actividad y en ese sentido, los fines no son algo externamente impuesto al sistema, sino las condiciones intrínsecas de su propio mantenimiento. Así pues, la lógica de la circularidad causal de un sistema disipativo genera una teleología intrínseca: aquellos aspectos del entorno del sistema que contribuyen a su mantenimiento son “buenos” para dicho sistema, y los que, por el contrario, lo alejan de su automantenimiento, son “malos”. Por eso, como ha propuesto C. Emmeche (2000), la agencialidad implica el establecimiento de alguna forma de “sentencia”, es decir, la creación de un punto de vista “subjetivo”.

### III. LA EVOLUCIÓN DE LA AGENCIALIDAD

De acuerdo a lo que hemos visto, todos los seres vivos son agentes autónomos. Incluso organismos unicelulares muy simples, como las bacterias *E. Coli*, pueden moverse direccionalmente, buscando alimentos o evitando sustancias nocivas. Son pues organismos capaces de desplegar una forma básica de comportamiento senso-motriz adaptativo. Pero el hecho de que todos los seres vivos sean agentes autónomos, capaces de normatividad intrínseca, no significa ni mucho menos que su acción sea volitiva. El fenómeno de la volición, como hemos dicho en la introducción, implica al menos alguna forma de control consciente en la actividad agencial, y esto a su vez, requiere un largo proceso evolutivo.

Ahora bien, la historia de la vida en nuestro planeta muestra que la evolución ha seguido muchos caminos. Animales y plantas por ejemplo muestran muy diferentes formas de agencialidad adaptativa (Keijzer 2006) y dentro del mundo animal, hay también grandes diferencias en la complejidad de su agencialidad. Pero, ¿qué significa hablar de grados de complejidad en la agencialidad? La evolución, como tal, no tiene una dirección definida: tiende a explorar todas las posibilidades organizativas viables. En la mayor parte de los casos la aparición de nuevas formas de vida no ha llevado aparejadas nuevas o más complejas formas de agencialidad.

Viendo la evolución desde nuestra perspectiva, puede parecer evidente que la aparición de más complejas capacidades agenciales otorga ventajas adaptativas. Sin embargo, y contrariamente a una idea bastante difundida (más, por cierto, entre los estudiosos de la cognición que entre los biólogos) el desarrollo de más complejas capacidades agenciales (“cognitivas”<sup>3</sup>) no es evidente que proporcione ventajas adaptativas directas. Biológicamente hablando, las capacidades cognitivas sólo se pueden considerar ventajosas (en términos adaptativos) en el contexto de circunstancias muy específicas. Por ejemplo, en la evolución de los homínidos el aumento del tamaño del córtex sólo podía suponer una suficiente ventaja evolutiva en el caso de ir acompañada de un conjunto de factores que compensaran sus inconvenientes, tales como un considerable aumento del gasto metabólico.

Por consiguiente, es necesario plantearse dos cuestiones: 1) qué es específicamente lo que evoluciona cuando pensamos en la historia de la agencialidad

3 Es frecuente asociar la evolución de la agencialidad senso-motriz con la aparición del fenómeno cognitivo, seguramente porque, a medida que aumenta la complejidad de aquella, aparecen una serie de rasgos y capacidades más y más cercanos a los procesos cognitivos específicamente humanos. Desde esta perspectiva, la aparición del sistema nervioso parece marcar una clara direccionalidad en la evolución de la agencialidad.

que, pasando por la vida animal, llega hasta nosotros; y 2) cuáles son los factores que explican esa línea evolutiva.

En un reciente trabajo, W. Christensen (2007) ha propuesto una interesante respuesta a la primera de las cuestiones. Este autor argumenta que cualquier teoría sobre la cognición debería intentar responder a una cuestión previa fundamental, a saber: cómo se explica la distribución empírica de las capacidades cognitivas en el ámbito biológico. En sus propios términos, “una teoría de la cognición debe explicar la evolución de la organización senso-motriz y de la conducta en los metazoos. Antes de pretender abordar y resolver cuestiones más complejas acerca de la arquitectura de la cognición humana (una buena teoría de la cognición) debe ser capaz de establecer (...) qué es lo que está bajo selección cuando evoluciona la cognición.” (p 258). La respuesta que este autor da a dicha pregunta es la siguiente: *lo que está bajo selección en la evolución específicamente cognitiva es el conjunto de factores que generan formas de control de alto nivel*. Y apoya esta afirmación en dos argumentos: 1) los datos empíricos que ofrece la historia evolutiva de los sistemas sensomotrices en los vertebrados; y 2) los descubrimientos de las neurociencias, que han puesto en evidencia cómo el córtex prefrontal realiza funciones de alto control jerárquico. Christensen considera que estos hechos invalidan la teoría de que la cognición está basada en formas de organización altamente distribuidas y abren la puerta a un programa de investigación de la evolución de la cognición en términos de control jerárquico.

La segunda cuestión es más difícil y compleja, seguramente porque hay diferentes factores en juego tanto externos como internos (Diéguez 2005). El gran cambio en la historia evolutiva de la agencialidad está ligado a la aparición de un tipo nuevo de células –las neuronas– que pueden formar redes (el sistema nervioso) capaces de permitir una agencialidad senso-motriz rápida, flexible y versátil en los multicelulares. Sin el sistema nervioso (SN) los multicelulares no habrían podido explorar los nichos basados en el movimiento, por lo que tiene una ventaja selectiva. Cabe pensar que el origen del SN es un caso de reclutamiento de una determinada variedad celular, cuyo origen es anterior.<sup>4</sup> Al mismo tiempo, la invención del SN ha permitido una mucho mayor versatilidad y flexibilidad de la agencialidad senso-motriz que la que existía en unicelulares. Pero esto no significa que, una vez aparecido, el SN fuera a evolucionar en la dirección de una mayor complejidad. De hecho, el desarrollo evolutivo del SN ha seguido una multitud de direcciones diferentes. En la mayoría de los fila animales la tendencia a incrementar la complejidad del sistema nervioso se

4 Las neuronas derivan (filogenéticamente) con mucha probabilidad de las células epiteliales, las cuales ya presentan funciones de comunicación a través de cambios de polaridad en sus membranas.

ha encontrado con cuellos de botella debido a las constricciones de sus planes corporales (Moreno & Lasa 2003). Sólo donde no ha sido así, como es el caso de los vertebrados (Moss 2006), ha podido darse este proceso de integración y jerarquización del que habla Christensen, que en términos morfológicos toma la forma de encefalización.

El proceso de encefalización permite por un lado cada vez mayores niveles de integración y control de los procesos sensomotrices; y por otro, un control de las tareas de regulación interna (del sistema cardio-respiratorio, digestivo, etc.). Como es sabido, una de las características de la evolución del SN de los vertebrados es el desarrollo de un “sistema nervioso del interior” (SNI) –el llamado sistema nervioso autónomo<sup>5</sup>– dedicado al control del funcionamiento de la organización interna. El desarrollo de este SNI ha resultado fundamental a su vez en el desarrollo de organizaciones corporales más complejas, y, sobre todo, ha posibilitado, como vamos a ver, una nueva forma de evolución del cerebro.

Para entender las funciones de este sistema voy a poner un ejemplo. Supongamos una liebre en un bosque que, de repente, oye un ruido. En literalmente el tiempo de uno o dos latidos del corazón, todo el funcionamiento de su cuerpo va a alterarse. Su ritmo cardiaco se acelera, su presión sanguínea aumenta, los vasos sanguíneos de sus músculos se dilatan, incrementando el flujo de oxígeno y nutrientes. Al mismo tiempo, los vasos sanguíneos de su tracto intestinal y de su piel se contraen, reduciendo el flujo sanguíneo en esas zonas y disponiendo más sangre para los músculos. Las pupilas se dilatan, mejorando la visión. La digestión se ralentiza, y el suministro de glucosa desde el hígado se facilita. Todo ello prepara al animal para enfrentarse a un potencial peligro o huir de él. Y todos estos cambios son desencadenados por el SNI.

Este ejemplo pone bien de manifiesto cómo el SNI controla el funcionamiento de las vísceras. Pero el SNI también está en permanente interrelación con el SN senso-motriz, y esta relación está muchas veces mediada por los estados corporales. Autores como Damasio (1994) o Lewis (2005) han argumentado que la compleja coordinación entre ambos sistemas –el del “interior” y el del “exterior”– está en el origen de las emociones. Las emociones aparecen como un proceso de feed-back entre el entorno, el SN senso-motriz y el SNI, mediado por cambios en los estados metabólicos y viscerales, de manera que los patrones de coordinación viables puedan ser recordados. La formación de estados emocionales desde este punto de vista no ocurriría únicamente en situaciones especiales, sino de forma continua. Además, diversas investigaciones

5 El sistema autónomo es el núcleo de estas tareas de regulación interna, pero en ulteriores fases evolutivas se han desarrollado otras zonas dedicadas también a dichas tareas, como el sistema límbico.



han mostrado que las emociones modulan de manera significativa las propias actividades cognitivas (Damasio op cit.). Las emociones juegan un papel esencial en la reorganización y evaluación selectiva de opciones complejas de comportamiento, pues permiten cierta forma de percepción de las consecuencias hipotéticas para el cuerpo que tendrían unas u otras opciones. El animal puede así anticipar de alguna manera las consecuencias de su comportamiento en el estado de sus vísceras.

En este sentido, el SNI juega un papel importante en la formación de un nuevo nivel de control jerárquico en el desarrollo evolutivo del cerebro de los vertebrados.<sup>6</sup> En algún momento de la evolución, probablemente cuando los vertebrados fueron capaces de movimientos terrestres rápidos y versátiles y su cerebro alcanzó cierto umbral de complejidad,<sup>7</sup> la actividad neuronal alcanzó unos niveles de integración que generaron alguna forma básica de consciencia o vida propiamente mental. Esta forma básica de consciencia no puede desvincularse de los mencionados procesos de coordinación entre el SNI y el SN del exterior (Edelman 1992) implicando incluso a las vísceras (Damasio 1994, 1999). Así, en lugar de una forma de agencialidad rápida y reactiva, la mente aparece ligada a otra forma de control sobre la conducta senso-motriz, más lenta, anticipativa, e implicando también continuos procesos de regulación con los estados corporales (Moreno & Lasa op.cit). Estos procesos, vinculados a experiencias emocionales, permiten empezar a hablar de una nueva forma, más compleja, de autonomía agencial: la agencialidad consciente.

6 La expresión externa de las emociones da lugar a la necesidad de que otros organismos las perciban, interpreten y respondan con sus propias emociones y conductas, contribuyendo así a generar una nueva forma de comportamiento comunicativo y social (Shepherd 1994). Este hecho está en la base de nuevos comportamientos, como el cuidado de las crías, nuevas relaciones presa-predador, y también nuevas formas de competencia y colaboración social. Todos estos comportamientos nuevos, a su vez, generan presiones selectivas para una mayor complejización cognitiva.

7 Por razones todavía no bien conocidas, en la evolución de los mamíferos se produjo un aumento significativo del neocórtex. Se ha sugerido que este hecho podría ser debido a que durante el largo periodo en que los mamíferos fueron predominantemente pequeños animales nocturnos ocurrió un aumento en las áreas olfativas y un cambio en las formas de procesamiento de la información visual en el tálamo, que habrían inducido una complejización del neocórtex (Aboitiz 1992). Este proceso filogenético habría sido posible por la capacidad ontogenética de las neuronas para migrar y formar nuevos patrones de conexión en el SN central (Gerhart & Kirschner 1997). Quizás, pues, lo único que se necesitó para generar este aumento del neocórtex fuese la eliminación de alguna restricción de los mecanismos del desarrollo ontogenético.

#### IV. LA NATURALEZA DE LA VOLICIÓN ELEMENTAL

Al final de la sección anterior explicábamos cómo los estados emocionales del cuerpo interactúan con las áreas corticales, de forma que el feed-back entre la alteración de los estados corporales desencadenado por el sistema senso-motriz produce una forma de percepción interna de tipo emocional. Este feed-back entre estados metabólicos, sistema nervioso autonómico y senso-motriz está en la base de lo que Edelman llama *consciencia primaria* (Edelman 1992, Edelman & Tononi 2000) y Damasio denomina *percepciones basadas en emociones secundarias* (Damasio 1994, 2003) o *consciencia nuclear* (Damasio 1999).

Tononi y Edelman explican la emergencia de este proceso en relación con la formación de vastos conjuntos de asambleas neuronales, funcionando de manera altamente integrada. Estas estructuras neurodinámicas se forman cuando se alcanza una integración coherente de la actividad cerebral de alto nivel de complejidad (es decir, la integración cruzada y correlacionada de un enorme número de patrones de actividad neuronales, tanto distribuidos como especializados y locales). El proceso que lleva a la integración de la actividad neuronal distribuida ha sido propuesto por Edelman y Tononi como la “marca” indicativa de la aparición de la consciencia primaria (no necesariamente auto-reflexiva), pues está claramente correlacionado con estados de vigilia, mientras que desaparece en situaciones de sueño profundo, anestesia o episodios epilépticos (Tononi & Edelman 1998, Edelman & Tononi 2000).

La formación de estructuras neurodinámicas de este tipo en las áreas talamocorticales ocurre en la banda de frecuencia gamma (Llinás *et al.* 1998). Este hecho ha llevado a algunos autores a sugerir que la sincronización en dicha banda de frecuencia en las áreas visuales puede ser el correlato de la experiencia visual consciente (Crick & Koch 1990) o de la consciencia perceptiva (Engel & Singer 2001). Detrás de todas estas propuestas está la idea de que la formación de una compleja estructura neurodinámica dominante (lo que Edelman y Tononi llaman un “núcleo dinámico”) estaría en la base de la unidad de la experiencia consciente y de su relativa lentitud temporal (sólo se percibe y decide por encima de unos pocos cientos de milisegundos, lo cual viene a corresponder a unos pocos ciclos gamma). Además, esto explicaría la singularidad específica y la riqueza experiencial de cada uno de los procesos conscientes, precisamente por el hecho de que en cada proceso integrativo de ese complejo núcleo dinámico se da una participación diferente de grupos especializados de neuronas.

El núcleo dinámico, con su alto nivel de integración y control, por un lado, evitaría los problemas que enfrentan los planteamientos distribuidos, y por otro, evitaría el problema del “homúnculo” a cargo de todo el proceso. El núcleo dinámico aparece pues como un conjunto de constricciones de alto nivel

que operan sobre formas inferiores de organizacional neurodinámicas. Pero es importante señalar que esta integración funcional de alto nivel no ocurre sólo “horizontalmente”, es decir, en el marco del SN del exterior, sino que a su vez, todos estos procesos están modulados por la actividad del SNI (cuya integración opera en una banda de frecuencia diferente, la teta). De hecho, una de las funciones más importantes del SNI y de los procesos emocionales es precisamente la modulación del SN del exterior para conseguir una mayor coherencia global: atención (amplificación de perturbaciones sensoriales) stress (generación de inestabilidades para provocar nuevas reestructuraciones globales) o satisfacción (refuerzo de las estructuras neurodinámicas que generen conductas que satisfacen condiciones de estabilidad global) (Barandiaran & Moreno 2006, Barandiaran 2008).

El resultado global de todos estos procesos es la emergencia de unas formas altamente integradas de organización neurodinámica, las cuales implican muchos niveles de control. Estas formas serían verosimilmente la base de la actividad agencial consciente en los vertebrados superiores. Que un animal controle su conducta significa que el conjunto de procesos de interacción que realiza como agente (buscar alimento, huir de un predador...) tienen diferentes grados de complejidad y están jerárquicamente organizados.

La naturaleza de esta jerarquía ya fue puesta en evidencia con algunos experimentos llevados a cabo a comienzos del pasado siglo en gatos (Brown 1911, Sherrington 1947). Cuando el tronco cerebral y el cordón espinal son aislados de la zona prefrontal, el gato es todavía capaz de respirar, tragar y andar. Sin embargo, sus movimientos se efectúan de manera estereotipada, robótica. El animal no es capaz de desarrollar actividades dirigidas, ni de responder a estímulos del entorno. Esto significa que el tronco cerebral y el cordón espinal son las zonas responsables de producir la coordinación motora básica, pero no las actividades que implican sensibilidad de alto nivel hacia los estímulos del entorno ni capacidad de dirección hacia objetivos. Un segundo tipo de experimentos muestra que los gatos a los que se les dejan los ganglios basales y el hipotálamo intactos son capaces de moverse espontáneamente evitando los obstáculos, beber, comer y manifestar emociones como la rabia. Pero sin el córtex prefrontal el animal no es capaz de desplegar las formas más complejas de actividad motora, como acechar y planificar estrategias de caza.

Estos experimentos no sólo evidencian la jerarquización de las formas de control agencial en los vertebrados superiores; también nos dan indicios sobre la naturaleza de las formas más complejas de agencialidad en dichos animales, en las que la aparición de una forma embrionaria de consciencia desempeña un papel central.

La agencialidad consciente, a diferencia de la meramente reactiva, es un proceso en el que el agente re-elabora internamente muchos *loops* senso-motri-

ces de bajo nivel. La “escena” o “el presente recordado” de los que habla Edelman son representaciones, es decir, estados mentales que no forman un continuo dinámico de procesos senso-motrices, sino que participan en un proceso más complejo entre el agente y su entorno, pues están mediados por una jerarquía de controles dinámicamente desacoplados. Las representaciones conscientes no están directamente conectadas a los órganos motores, sino en primer lugar a las vísceras, produciendo estados emocionales. El animal incluso construye simulaciones internas: en lugar de usarlas directamente para el control de acciones reales, ciertas configuraciones neuronales son empleadas para desarrollar modelos internos de la realidad que operan como interacciones virtuales con el entorno (Coterill 2001). Como señala Diéguez (op cit), “los organismos con el tipo de sistema cognitivo que capacita para tener representaciones desvinculadas pueden realizar mediante ellas simulaciones subjetivas del medio, lo cual les permite a su vez una previsión de las posibles intervenciones sobre él, evitando los riesgos de poner a prueba realmente ciertas conductas.” (p. 33).

Emerge así un tipo de actividad agencial basada en una forma básica de consciencia que consiste en una escena en tiempo presente pero construida con todas las respuestas, cargadas de valores, de su anterior experiencia vital. El control voluntario de la acción consciente (por ejemplo, el gato al lanzarse sobre su presa después de un proceso de acecho) no es posible sin la integración de una multitud de procesos más rápidos de bajo nivel. La decisión consciente es un proceso global, que implica en definitiva la elaboración de un conjunto de constricciones de control de alto nivel ejercidas sobre una multitud de automatismos inconscientes.

Este esquema es consistente con el hecho bien conocido de que los procesos específicos de control agencial a nivel consciente requieren lapsos temporales mucho más lentos que los procesos de control de niveles inferiores. Diferentes experimentos (Libet *at al.* 1983, Wegner 2002) han mostrado que la experiencia subjetiva de control voluntario de los actos es posterior a su inicio real, y de este hecho se ha querido concluir que la volición es una ilusión. En palabras de Wegner: “El inicio de un acto voluntario resulta ser un proceso inconsciente del cerebro. Claramente, la voluntad libre para elegir actuar ahora puede no ser debida al agente (consciente), contrariamente a la opinión mayoritaria.” (Wegner, op.cit, p. 54). Pero esta interpretación sólo se justifica si concebimos la volición como un acto instantáneo, en lugar de verla como un proceso que incluye diferentes niveles temporales. En este sentido, P. Petit (2007) ha propuesto que, en lugar de considerar la volición como un acto, la veamos como una consecuencia del hecho de estar generada por un determinado tipo de agentes. Esto significa que no es el estado consciente por sí mismo lo que garantiza que la acción sea voluntaria, sino una continuidad de procesos encadenados realizados bajo control de la atención consciente.

Desde esta perspectiva, el hecho de que una acción voluntaria pueda haberse iniciado antes de ser reconocida como tal no invalida su carácter de acción voluntaria. En el caso animal obviamente no podemos formular la cuestión en estos términos, pero sí podemos distinguir entre dos diferentes niveles de acciones realizadas en estado consciente; un nivel correspondería a las acciones meramente impulsivas, y otro nivel, en cambio, las que requieren un tipo de control en estado de atención consciente. A partir de esta distinción, la analogía sería, en el caso animal, que esta última forma de acción puede iniciarse antes de ser percibida como controlada por una atención consciente.

## V. CONCLUSIONES

Decir que una acción es voluntaria no significa que no esté sujeta a las leyes fundamentales de la física, sino que está causada por un agente autónomo intencional. Esto significa no sólo que la acción es causalmente atribuible al agente, como sistema autónomo, sino que además éste opera controlando intencionalmente un flujo de procesos agenciales de nivel inferior. Por ejemplo, cuando firmo un documento, no estoy especificando conscientemente todos los movimientos de mi mano; al contrario, en gran medida éstos son “automatismos”, es decir, procesos gobernados por niveles infraconscientes de mi SN. Lo que hace que mi firma sea voluntaria es que yo ejerzo un control global consciente sobre el proceso.

De manera análoga, un vertebrado superior ejerce un control de alto nivel en ciertas operaciones que requieren atención y construcción de “escenas” conscientes. Los niveles inferiores están controlados por la agencialidad consciente, pero sólo globalmente. Aunque si comparamos este ejemplo con el anterior hemos bajado un nivel en la escala, también aquí podemos hablar de niveles infraconscientes: éstos proporcionan formas estereotipadas de activación motora, que el control superior, consciente, ajusta en función de las circunstancias y regula especificando dirección y velocidad.

Es incuestionable que entre el tipo de agencialidad de un animal para el que la consciencia cubre sólo el espacio temporal de un presente corto (una escena consciente construida sobre la historia de sus respuestas cargadas de valores) y la de los humanos, que somos capaces de prever las consecuencias a largo plazo de nuestros actos, hay una profunda diferencia. Pero la agencialidad autónoma de tales animales constituye el origen de la nuestra.

## AGRADECIMIENTOS

Este trabajo ha sido financiado con las ayudas de Grupos Consolidados del Gobierno Vasco IT-250-07, y los Proyectos de Investigación HUM2005-02449 y BFU2006-01951/BMC del MEC, y de los fondos Feder de la Unión Europea.

## REFERENCIAS

- ABOITIZ, F. 1992: "The evolutionary origin of the mammalian cerebral cortex." *Biological Research* 25(1): 41-49.
- BARANDIARAN, X. & MORENO, A. 2006: "On What Makes Certain Dynamical Systems Cognitive: A Minimally Cognitive Organization Program." *Adaptive Behavior*. 14(2): 171-185.
- BARANDIARAN, X. 2008: *Mental Life. A naturalized approach to the autonomy of cognitive agents*. Tesis Doctoral. Universidad del País Vasco.
- BROWN, T. S. 1911: "The intrinsic factors in the act of progression in the mammal." *Proc. Royal Soc. London, Series B*, 84: 308-319.
- COTERILL, R. 2001: "Evolution, cognition and Consciousness" *Journal of Consciousness Studies*, 8(2): 3-17.
- CRICK, F. & KOCH, C. 1990: "Towards a neurobiological theory of consciousness." *Seminars in the Neurosciences* 2: 263-275.
- CHRISTENSEN, W. 2007: "Evolutionary origins of volition," En Spurrett, D., Kincaid, H., Ross, D. & Stephens, L. (eds.) *Distributed Cognition and the Will: Individual Volition and Social Context*, Cambridge, Mass.: MIT Press, 255-288.
- DAMASIO, A. 1994: *Descartes' Error. Emotion, Reason and the Human Brain*. New York: G. P. Putnam's Sons. Trad. española.
- DAMASIO, A. 1999: *The feeling of what happens*. New York: Harcourt Brace and Co.
- DAMASIO, A. 2003: *Looking for Spinoza*. Harcourt: Brace and Co. Trad. española.
- DIEGUEZ, A. 2005: "Representación, cognición y evolución." *Contrastes*, suplemento 10: 15-38.
- DI PAOLO, E. 2003: "Organismically-inspired robotics: homeostatic adaptation and natural teleology beyond the closed sensorimotor loop." In Murase, K., Asakura, T. (Eds.), *Dynamical Systems Approach to Embodiment and Sociality*. Adelaida: Advanced Knowledge International, 19-42.
- EDELMAN, G. M. 1992: *Bright Air, Brilliant Fire: On The Matter of the Mind*. New York: Basic Books.

- EDELMAN, G. M. & TONONI, G. 2000: *A universe of consciousness. How matter becomes imagination*. New York: Basic Books. Trad. española.
- EMMECHE, C. 2000: "From Robotics and Cybernetic Vehicles to Autonomous Systems; the organism lost and found?" *Communication and Cognition - Artificial Intelligence*, 17, (3-4):159-187.
- ENGEL, A.K. & SINGER, W. 2001: "Temporal binding and the neural correlates of sensory awareness." *Trends in Cognitive Sciences* 5(1):16-25.
- GERHART, J & KIRSCHNER, M. 1997: *Cells, Embryos and Evolution*. London: Blackwell Science.
- KAUFFMAN, S. 2000: *Investigations*. Oxford: Oxford University Press.
- KEIJZER, F. 2006: "Differentiating animality from agency: Towards a foundation for cognition." En R. Sun & N. Miyake (Eds.) *Proceedings of CogSci/ICCS 2006*, Alpha NJ: Sheridan Printing: 1593-1598.
- LEWIS, M.D. 2005: "Bridging emotion theory and neurobiology through dynamic systems modeling." *Behavioral and Brain Sciences* 28: 169-194.
- LIBET, B., GLEASON, C. A., WRIGHT, E. W., AND PEARL, D. K. 1983: "Time of conscious intention to act in relation to onset of cerebral activity" *Brain* 106: 623-642.
- LLINÁS, R., RIBARY, U., CONTRERAS, D. & PEDROARENA, C. 1998: "The neuronal basis for consciousness." *Philosophical Transactions of the Royal Society of London, Series B*, 353:1841-1849.
- MAES, P. 1994: "Modeling Adaptive Autonomous Systems" *Artificial Life* 1: 135-162.
- MORENO, A. & LASA, A. 2003: "From basic adaptivity to early mind: the origin and evolution of cognitive capacities. *Evolution and Cognition*, 9 (1): 12-24.
- MORENO, A. & ETXEBERRIA, A. 2005: "Agency in natural and artificial systems" *Artificial Life* 11(1-2): 161-176.
- MOSS, L. 2006: "Redundancy, Plasticity, and Detachment: the Implications of Comparative Genomics for Evolutionary Thinking." *Philosophy of Science* 73: 930-946.
- NICOLIS, G. & PRIGOGINE, Y. 1977: *Self-organization in Non-equilibrium Systems*. New York: Wiley.
- PETIT, P. 2007: "Neuroscience and Agent-Control." En Spurrett, D., Kincaid, H., Ross, D. & Stephens, L. (eds.) *Distributed Cognition and the Will: Individual Volition and Social Context*, Cambridge, Mass.: MIT Press: 77-91.
- SHEPHERD, G. M. 1994: *Neurobiology*. 3rd Ed. New York: Oxford University Press.
- SHERRINGTON, C. S. 1947: *The Integrative Action Of The Nervous System* (2nd ed.). New Haven: Yale University Press.

- SMITHERS, T. 1997: "Autonomy in Robots and Other Agents." *Brain and Cognition* 34. 88-106.
- TONONI, G. & EDELMAN, G. M. 1998: "Consciousness and Complexity". *Science* 262: 1846-1851.
- WEGNER, D.M. 2002: *The Illusion of Conscious Will*. Cambridge: MIT Press.