

LA MENTE ALGEBRAICA vs EL CEREBRO ESTADÍSTICO

THE ALGEBRAIC MIND vs THE STATISTICAL BRAIN

Saray Ayala L.

Universidad Autónoma de Barcelona
Barcelona-España

RESUMEN

El debate entre Gary Marcus y Jeff Elman, en Marcus (1999, 1999a, 1999b, 1999c, 1999d, 2001) y Elman (1999, 1999a, 1999b, 1999c, 1999d) se enmarca dentro del debate general en Ciencia Cognitiva entre la hipótesis clásica (simbólica) y la hipótesis conexionista. Marcus critica lo que se denomina Conexionismo Eliminativista, que incluye los modelos conexionistas que eliminan el papel de los símbolos y las reglas explícitas. Según Marcus, los modelos conexionistas eliminativistas no logran generalizar del modo como hacen los humanos y, por tanto, el Conexionismo Eliminativista no constituye una buena explicación de la cognición humana. En este trabajo intentaré abordar el debate entre clásicos y conexionistas a partir de los experimentos llevados a cabo por Marcus en Marcus *et al.* (1999) y las críticas de Elman al mismo, e intentaré argumentar que Marcus no consigue una crítica consistente al Conexionismo Eliminativista y, asimismo, no logra ofrecer una buena defensa del Conexionismo Implementacional.

Palabras Clave: Conexionismo, Simbolismo, Reglas, Estadística.

ABSTRACT

The debate between Gary Marcus and Jeff Elman, in Marcus (1999, 1999a, 1999b, 1999c, 1999d, 2001) and Elman (1999, 1999a, 1999b, 1999c, 1999d) is situated in the general debate in Cognitive Science between the Classical (symbolic) hypothesis and the Connectionist one. Marcus is against what is called Eliminative Connectionism, which includes the connectionist models that eliminate the role of symbols and explicit rules. According to Marcus, eliminative connectionist models are not able to generalize in the way humans do. Thus, Eliminative Connectionism cannot furnish us with a plausible model of cognition. In this notice I shall approach the debate between classicists and connectionists as framed by Marcus in Marcus et al. (1999). I shall try to argue contra Marcus: his strategy against (Eliminative) Connectionism fails in getting a good argument; therefore he does not get a sound defence of Implementational Connectionism.

Key Words: *Connectionism, Symbolism, Rules, Statistic.*

INTRODUCCIÓN

En términos generales, hay dos posturas en Ciencia Cognitiva con respecto a la cuestión de cuál es la arquitectura correcta de la cognición. Según la postura clásica (simbólica), la cognición consiste en la manipulación (algebraica) de símbolos de acuerdo con reglas explícitas. Bajo el marco conexionista, de otra parte, el comportamiento de un sistema se considera dirigido por las regularidades del medio, sin identificar en el mismo, símbolos ni reglas que lo gobiernen. Según el Conexionismo, la actividad cognitiva es el resultado de las interacciones de unidades simples de procesamiento que operan en paralelo y no siguen reglas explícitas.¹ Abordaré el

¹ Para una introducción a la teoría conexionista el lector no familiarizado puede consultar McLeod, Plunkett. & Rolls. (1998). *Introduction to connectionist modelling of cognitive processes*. New York: Oxford University Press.

debate entre clasicistas y conexionistas a partir del trabajo de Gary Marcus (2001).

Marcus distingue dos formas de conexionismo: el eliminativista y el implementacional. El conexionismo eliminativista obedece a la caracterización del conexionismo ortodoxo, esto es, no identifica símbolos en el sistema y rechaza la necesidad de recurrir al empleo de reglas explícitas.² De acuerdo con la perspectiva implementacional, aquellos modelos conexionistas que logran conservar su equivalencia computacional con respecto a su contrapartida clásica consiguen hacerlo porque implementan un modelo clásico (gobernado por reglas). Esta versión implementacional del conexionismo es teóricamente un modelo clásico: el sistema cognitivo, analizado desde un nivel funcional de explicación, presenta una serie de símbolos manejados por reglas explícitas. Pero su implementación se lleva a cabo en una arquitectura distinta: una red de unidades simples operando en paralelo de acuerdo a unas reglas. Marcus sostiene que el único aporte del conexionismo al debate sobre cómo es la cognición es al nivel de la implementación (o realización) física, y no al nivel computacional o funcional.³ Esto es, que cuando una red conexionista logra dar cuenta de las tareas cognitivas requeridas lo hace porque está implementando un modelo clásico. Intentaré argumentar que Marcus no consigue una crítica consistente a los logros del conexionismo (eliminativista) y que, dada la propuesta y los resultados que éste ofrece, la arquitectura conexionista (eliminativista) puede ofrecernos un modelo plausible de la cognición.

² Por simplicidad, podemos adoptar la definición que Marcus ofrece de regla en términos de “relaciones abstractas abiertas para las cuales podemos sustituir valores arbitrarios” (1999, p. 77) (“open-ended abstract relationships for which we can substitute arbitrary items”).

³ Véase Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: Henry Holt & Company; respecto a los tres niveles de explicación de un sistema.

Permítanme empezar resumiendo brevemente un experimento reciente realizado por Gary Marcus.⁴ En una serie de experimentos,⁵ Marcus y colaboradores (1999),⁶ examinaron las capacidades de generalización de niños de siete meses sobre un corpus de secuencias de tres sílabas. En la fase de habituación, los niños se familiarizaron con secuencias de tres sílabas que seguían o la gramática ABB (e.g., “ga ti ti”) o la gramática AAB (e.g., “ga ga ti”). En la fase de evaluación, los niños fueron expuestos a secuencias formadas por sílabas nuevas, tales como “wo fe fe” o “wo wo fe”, y sólo la mitad de estas secuencias eran consecuentes con la gramática a la que los niños habían sido expuestos. Marcus y colaboradores encontraron que los niños mostraban preferencia por las secuencias que no eran consecuentes con la gramática con la que habían sido familiarizados.⁷ La interpretación de estos resultados por parte de Marcus y colaboradores fue afirmar la existencia de una operación algebraica que guiaba el comportamiento de los niños. Ya que la estructura superficial de las secuencias era distinta en la fase de habituación (e.g. “ga ga ti”) y en la de evaluación (e.g. wo wo fe”), la preferencia de los niños por los ejemplares no consistentes parecía indicar que atendieron a la estructura abstracta que subyace a las secuencias.

Además, Marcus realizó una red recurrente simple (SRN)⁸ y argumentó, para refutar la hipótesis de trabajo del

⁴ *et al.*

⁵ Marcus *et al.* realizaron tres grupos separados de experimentos con diferentes grupos de datos, con intención de dar cuenta de las posibles críticas desde el conexionismo y obtener así resultados más sólidos. Para nuestros intereses, podemos centrarnos en el último de esos experimentos (que se explica antes) e ignorar los otros.

⁶ Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science* (283).

⁷ Esta preferencia consistía en el hecho de que los niños miraban durante más tiempo a la luz que acompañaba a la presentación de las oraciones que no seguían la gramática de la fase de habituación.

⁸ Una SRN es una red con flujo de información hacia adelante estándar (feedforward) ayudada de una capa de contexto que se usa para copiar el patrón de

conexionismo de acuerdo con la cual el comportamiento de los niños debería ser explicado en términos puramente estadísticos, que un modelo tal no podría dar cuenta de los datos obtenidos en su experimento.

En una respuesta reciente a Marcus y colaboradores, Jeff Elman (1999a) sostiene que sus experimentos no implican que la mente haga uso de reglas. Para sostener su afirmación empíricamente, Elman realizó una simulación conexionista que lograba capturar los resultados obtenidos por aquellos. Para hacer frente a los datos y preservar la adecuación computacional de su modelo, Elman tuvo en consideración algo que Marcus había ignorado en su simulación: la experiencia anterior de los niños. A los siete meses los niños han oído ya aproximadamente 6 millones de palabras, que incluyen los estímulos que Marcus utiliza en la fase de evaluación de su experimento. La evidencia indica que los niños son sensibles a las regularidades contenidas en los estímulos percibidos, y gracias a la cantidad de datos a los que han sido expuestos pueden aprender a distinguir estímulos similares de estímulos diferentes, una capacidad que resulta ser esencial para realizar la tarea en cuestión. De este modo, antes de las fases de habituación y evaluación, Elman expuso la SRN a un grupo de 120 sílabas que corresponderían a los estímulos a los que los niños habían sido expuestos desde su nacimiento. En esta fase de preentrenamiento, la SRN dominó la relación igual/diferente haciendo uso de un algoritmo extensamente utilizado conocido como *retropropagación*. Bajo esta técnica de aprendizaje, las conexiones entre las unidades de procesamiento son ajustadas gradualmente para reducir la diferencia entre la respuesta actual de la

activación que la red adquiere en el instante t de procesamiento. En el instante $t+1$ la red puede entonces procesar la información de dos fuentes: la información que viene del medio vía la capa input y la que previamente se almacenó en la capa de contexto. De este modo, una SRN logra una memoria a corto plazo que le permite llegar a ser competente en el procesamiento de un input que se extiende en el tiempo (e.g., datos lingüísticos).

red y la correcta deseada para cada unidad output (el valor es proporcionado por un “maestro”). Este tipo de aprendizaje se considera supervisado, debido a que estamos facilitándole a la red una medida de error.

La SRN de Elman fue entonces entrenada y evaluada sobre un conjunto de secuencias idéntico a aquel utilizado en el experimento de Marcus y colaboradores. El punto clave es que el proceso llevado a cabo por la red parece ser explicable sobre la base de las regularidades estadísticas inherentes en el corpus de datos: una vez que la red conoce la relación igual/diferente, puede codificar información sobre secuencias de sílabas y leer su estructura gramatical abstracta fuera de las relaciones particulares entre pares de sílabas iguales y diferentes (e.g., A y A (iguales), A y B (diferentes)).

Sin embargo, de acuerdo con Marcus, aunque el modelo de Elman es computacionalmente adecuado, fracasa en su intento de demostrar que el comportamiento de los niños está guiado exclusivamente por las regularidades estadísticas. Marcus considera que la simulación de Elman consigue generalizar del modo requerido porque hace uso de una regla que está escondida en el algoritmo de aprendizaje que utiliza:

[Crucialmente, una] parte del sistema que [Elman propone] sí implementa una regla: el maestro externo... En particular, el maestro externo incorpora una regla universalmente abierta del tipo, *para todas las sílabas x, y, si x = y entonces output = 1, en caso contrario, output = 0* (Marcus, 1999, p. 289).⁹

El argumento de Marcus es que aunque el comportamiento de la red al nivel del microprocesamiento puede ser explicado en términos de los postulados que la teoría conexio-

⁹ [Crucially, a] part of the system that [Elman proposes] does implement a rule: the external teacher... In particular, the external teacher incorporates a universally open-ended rule of the sort, *for all syllables x, y, if x = y then output 1 else output 0*.

nista típicamente utiliza –estados en el espacio oculto, etc.– la red logra ser competente gracias a haber aprendido previamente la relación igual/diferente en un estilo clásico. Esto es, por medio del uso de una regla explícita que le dice a la red que responda con un output = 1 si las dos sílabas que acaba de captar son ejemplares del mismo tipo representacional, y con un output = 0 si pertenecen a diferentes tipos representacionales.¹⁰

En este punto hay al menos dos opciones que podemos considerar. Por una parte, podemos seguir a Marcus y asumir que el “maestro externo” utilizado en los algoritmos de aprendizaje supervisado incorpora una regla. Por otra parte, podemos argumentar que las técnicas de aprendizaje supervisado no incorporan reglas en absoluto. Manteniendo esta última opción, seguiré una segunda y más prometedora línea de discusión.¹¹ Argumentaré ahora que comparando formas de aprendizaje supervisado con otras técnicas de aprendizaje ampliamente utilizadas con SRNs –a saber, aprendizaje auto-supervisado– el teórico clásico está forzado a elegir entre dos

¹⁰ Podemos responder a Marcus aquí, en defensa de la postura conexionista, diciendo que esas reglas gramaticales que él afirma que subyacen a la ejecución de los niños (y a la simulación de Elman) crean otras regularidades estadísticas (regularidades que Marcus no reconoce, apelando a que las secuencias en la fase de evaluación estaban compuestas por sílabas nuevas). “AAB, por ejemplo, indicaba que una sílaba iría seguida por una instancia de la misma sílaba y después por una sílaba diferente. Así, en la fase de preentrenamiento, el niño fue expuesto a una regularidad estadística que dominaba las secuencias de eventos perceptualmente iguales y diferentes” (Elman & Seidenberg, 1999b) (“AAB, for example, indicated that a syllable would be followed by another instance of the same syllable and then a different syllable. Thus, in the pretraining phase, the infant was exposed to a statistical regularity governing sequences of perceptually similar and different events”).

¹¹ Hay que apuntar que Marcus no niega el hecho de que los humanos explotamos fuentes estadísticas. Él sólo afirma que las herramientas algebraicas son necesarias (aunque no suficientes) para explicar la arquitectura de la cognición. De este modo, el lector notará que optar por la primera opción nos llevaría a una posición en el debate donde ambos contendientes están de acuerdo en que el aprendizaje supervisado implica alguna forma de comportamiento gobernado por reglas, y que el flujo de información en la red es explicado parcialmente en términos conexionistas. Tal postura favorecería seguramente al clásico, quien no está obligado a mostrar que la cognición es exclusivamente clásica.

extremos: a) “panreglismo” (“pan-rulism”) (i.e., postular reglas en cualquier forma de aprendizaje), o b) negar la existencia de reglas en los algoritmos de aprendizaje supervisado.

La red de Elman anteriormente considerada, fue pre-entrenada sobre una tarea de categorización para distinguir entre sílabas iguales y diferentes. Las SRNs, sin embargo, son típicamente entrenadas sobre tareas de predicción donde la red debe predecir el siguiente ejemplar de una secuencia.¹² Aunque entrenadas por retropropagación, las SRNs entrenadas sobre tareas de predicción se denominan autosupervisadas, como opuestas al tipo de supervisión antes mencionado, que puede ser caracterizado como completamente supervisado, ya que el *output* deseado no es proporcionado externamente por el neuromodelador, sino que consiste en el siguiente *input* en la secuencia de datos.¹³ En resumen, el punto clave para nuestros intereses es que la SRN entrenada sobre una tarea de predicción no requiere un “maestro externo”. La principal diferencia entre las dos formas de entrenamiento dirigido por error podría ser caracterizada en términos de la fuente de la señal: cuando la fuente es proporcionada naturalmente por el régimen de aprendizaje, el aprendizaje es sólo autosupervisado. Cuando somos nosotros, y no el medio, quienes proporcionamos la señal, el aprendizaje se considera totalmente supervisado.

La relación de esta distinción con el debate sobre la arquitectura de la cognición es la siguiente: por un lado, Marcus acepta que el flujo de información desde la capa *input* a la capa *output* de la red es explicado localmente, examinando las interacciones de las unidades simples de procesamiento. A este nivel de explicación, sólo la estadística importa. Sin embargo, si el conexionismo implementacional está en lo correcto, como

¹² Ejemplos bien conocidos en la literatura conexionista son Elman (1990, 1998).

¹³ El ejemplar *b* que alimenta a la red en el instante $t+1$ de procesamiento es el objetivo para el instante t de procesamiento, donde el output actual es aquel producido por el ejemplar *a* en el instante t .

mantienen los clásicos, el carácter reglado del comportamiento (“rule-governed” behaviour) debe venir de la señal externa. El simpatizante del conexionismo puede atacar poniendo énfasis bien en el término “regla” o bien en el término “gobernado” de la ecuación “gobernado por reglas” (“rule-governed”). La segunda estrategia constituye una defensa ante la postura clásica y ha sido atendida en numerosas ocasiones.¹⁴ Sugiero, sin embargo, poner el énfasis en el término “regla” de la ecuación y elaborar un ataque específico contra la estrategia de Marcus, y no sólo contra los clásicos en general. Como Elman y Seidenberg (1999a) señalan, en una red con aprendizaje supervisado dirigido por error, el algoritmo es tan sólo: “un modo simple y directo para hacer que el modelo codifique las propiedades de los estímulos que por hipótesis subyacen a la tarea de categorización de los niños” (p. 288).¹⁵ (“a simple and direct way to cause the model to encode statistical properties of the stimuli that we hypothesize underlie babies’ categorization performance”).

La medida de error ofrece al neuromodelador una forma de codificar aquellas regularidades estadísticas contenidas en el corpus de datos que, bajo la hipótesis de trabajo conexionista, subyacen a la ejecución del sujeto. Que la tarea sea de categorización o de predicción no debería marcar diferencia alguna, así como no parece realizarlo el nivel de supervisión del sistema: las redes con aprendizaje supervisado hacen lo mismo que las denominadas autosupervisadas, de modo que el origen de la señal maestra (venga de fuera, en el caso de las supervisadas, o esté integrada en el input que recibe la red, en el caso de las autosupervisadas) no puede ser lo que determine la existencia de reglas. El párrafo citado mantiene lo mismo: independiente-

¹⁴ Se trata de establecer una distinción entre comportamiento descrito mediante reglas (rule-described behaviour) y comportamiento gobernado por reglas (rule-governed behaviour), y argumentar que las redes conexionistas son mecanismos cuyo comportamiento es susceptible de ser descrito mediante reglas, pero en los que las reglas no tienen poder causal.

¹⁵ Elman & Seidenberg, 1999a, p. 288.

mente del origen de la señal, el aprendizaje dirigido por error es un “modo simple y directo para hacer que el modelo codifique las *propiedades estadísticas* de los estímulos”.¹⁶

Si las consideraciones anteriores son correctas y el origen de la señal de error no es lo que está en juego, Marcus parece quedar obligado a mantener uno de los siguientes enunciados:

1. Los modelos autosupervisados son sistemas gobernados por reglas.
2. Los modelos totalmente supervisados no son sistemas gobernados por reglas.

La opción (1) viene a ser una versión fuerte de lo que antes he apodado “panreglismo” (“pan-rulism”): la opinión de que todos los tipos de algoritmos de aprendizaje dirigido por error incorporan reglas clásicas.¹⁷ (Des)afortunadamente, Marcus no estaría de acuerdo con el enunciado (1). En un debate diferente, Marcus (2001) afirma que las SRNs de predicción de Elman no logran manifestar una forma fuerte de sistematicidad, una característica esencial de la cognición.¹⁸ El problema esta vez para el conexionismo surge de su incapacidad para generalizar correctamente en dominios particulares. Sin embargo, cuando Elman (1998) propone un modelo que consigue generalizar, Marcus no argumenta que la red está implementando

¹⁶ *Idem*. Las cursivas son del autor.

¹⁷ Alguien podría argüir que otras formas de aprendizaje sin error dirigido (e.g., aprendizaje Hebbiano) pueden acabar con el debate –el lector habrá notado que dado que el aprendizaje Hebbiano no requiere el empleo de un “maestro externo”, el asunto de si la señal de error incorpora una regla o no simplemente no surge–. Sin embargo, las formas de aprendizaje biológicamente plausibles deben incorporar una medida de error, y ya que el aprendizaje Hebbiano funciona exclusivamente sobre la base de correlaciones en el corpus de datos, no será por lo tanto de ayuda en, digamos, tareas complejas de categorización. El “panreglismo” concede que una forma u otra de aprendizaje dirigido por error debe estar presente. (Cfr. O’Reilly & Munakata, 2000).

¹⁸ Cfr. Fodor, J. A. & Pylyshyn, Z., 1988, pp. 3-71.

una regla. Marcus sólo se queja cuando Elman cambia de una tarea de predicción a una de categorización.

En base a las consideraciones anteriores, Marcus, en mi opinión, se ve forzado a optar por el enunciado (2). A mi entender, Marcus interpreta el hecho de que el *output* deseado (*target output*) no venga del medio (sino del “maestro externo”), como la causa de la emergencia de un comportamiento reglado. He argumentado que a menos que opte por el “panreglismo”, no hay modo de defender que el comportamiento de la red esté gobernado por reglas. La presencia de un maestro “externo” provoca cambios en el espacio de activación de la red. Sin embargo, las microactivaciones en el espacio representacional no se interpretan en términos de valores arbitrarios que implementan variables abstractas en un estilo clásico, simplemente por el hecho de que la señal de error tenga un origen particular.

Se podría decir, en contra de mi estrategia argumental, que es posible que Marcus no acuse a las redes con aprendizaje autosupervisado de implementar una regla justamente porque no las considera exitosas en sus tareas. Marcus tiene en cuenta ciertas redes autosupervisadas que no logran generalizar del modo requerido, y quizás por esa razón no entra a discutir si implementan reglas o no. Una posible respuesta a esto sería que, independientemente de que las redes con aprendizaje autosupervisado logren dar cuenta de las tareas que se le asignan, y dejando a un lado la discutible justicia de la afirmación de Marcus de que tales redes no generalizan del modo requerido¹⁹, el solo hecho de que Marcus considere que en el algoritmo de aprendizaje se esconde una regla es ya motivo de posible crítica, pues desde el conexionismo se sostiene que tales algoritmos no afectan al tipo de procesamiento que la red

¹⁹ Elman (1990) y Elman (1998), son ejemplos de que una SRN autosupervisada da cuenta con éxito de tareas de predicción generalizando palabras a contextos nuevos.

ejecuta (un procesamiento basado en los datos recibidos únicamente) sino sólo al output resultante.²⁰

Otra posible réplica contra mi trabajo podría ser respecto al tipo de reglas que se ha considerado. Desde el principio se habla de reglas explícitas y se hace uso de la definición que da Marcus, a saber: “open-ended abstract relationships for which we can substitute arbitrary items” (p. 77).²¹

Cabe apuntar que no se han tenido en cuenta otro tipo de reglas (reglas implícitas), y por tanto la conclusión sólo puede ser aplicada al tipo considerado (esto es, se ha defendido que las redes con aprendizaje supervisado, así como las autosupervisadas no implementan reglas explícitas). Así, es posible replicar que las redes de las que se ha defendido que no implementan reglas (explícitas), sí hacen uso de otras reglas con poder causal sobre el comportamiento de la red: reglas implícitas, o hard-wired.

Ante ello, mi contrarréplica sería alegar que la crítica de Marcus a Elman parece consistir en afirmar que las redes con aprendizaje supervisado hacen uso de una regla incorporada en el algoritmo de aprendizaje que le dicta lo que tiene que hacer en cada caso. El tipo de regla al que parece estar refiriéndose Marcus es una regla explícita, y es por eso que lo pertinente era centrarse en ese tipo de reglas, respondiendo a Marcus en el mismo marco en el que él se situaba. Por tanto, la conclusión, efectivamente, sólo se aplica a reglas explícitas, pero a lo largo del trabajo se ha dedicado importante atención a subrayar que la clave (tanto para la crítica a Marcus como en el debate general sobre implementación de reglas) es cómo la red procesa, de qué hace uso para responder lo que responde, cuáles son los datos que utiliza. Y se ha dicho que lo único que

²⁰ Hay que tener en cuenta que este “control sobre el output” no es global, no se controla la respuesta total de la red, sino que se va midiendo el error de cada unidad output, error que se retropropaga y del que la red aprende.

²¹ Marcus, 1999, p. 77.

tiene poder causal en ella son los inputs recibidos y la medida de error retropropagada. Por tanto, también se ha intentado extender la conclusión respecto a las reglas en general y no sólo a las reglas explícitas, si bien no dando argumento concluyente alguno ni centrando el debate en tal marco.

REFERENCIAS

- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Elman, J. L. (1998). Generalization, simple recurrent networks, and the emergence of structure. En M.A. Gernsbacher & S.J. Derry (Eds.), *Proceedings of the twentieth annual conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Elman, J. L. (1999). *Generalization, rules, and neural networks: A simulation of Marcus*. (et al. 1999).
- Elman, J. L. & Seidenberg, M. S. (1999a). Networks are not 'hidden rules'. *Trends in Cognitive Science*, 3 (8), 288-289.
- Elman, J. L. & Seidenberg, M. S. (1999b). *Do infants learn grammar with algebra or statistics?*, aparecido como carta en *Science*, 284, 433. Abril 16.
- Elman, J. L. & Seidenberg, M. S. (1999c). Reply to Marcus. *Science*. Abril 17.
- Elman, J. L. (1999d). *Commentary on The Algebraic Mind, by Gary Marcus*. Cambridge: MIT Press.
- Fodor, J. A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.
- Marcus, G. F. (1998). Rethinking Eliminative Connectionism. *Cognitive Psychology*, 37 (3).
- Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, 283.
- Marcus, G. F. (1999a). Connectionism: with or without rules? Response to J.L. McClelland and D.C. Plaut. *Trends in Cognitive Sciences*, 3, 168-170.
- Marcus, G. F. (1999b). Rule learning by seven-month-old infants and neural networks. Response to Altmann and Dienes. *Science*, 284, 875a.
- Marcus, G. F. (1999c). Reply to Seidenberg and Elman, *Trends in Cognitive Sciences*, 3, 289.

- Marcus, G. F. (1999d). *Response to Jeff Elman's. Commentary on The Algebraic Mind*. Cambridge: MIT Press.
- Marcus, G. F. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. Cambridge: MIT Press.
- McLeod, Plunkett & Rolls (1998). *Introduction to Connectionist Modelling of Cognitive Processes*. New York: Oxford University Press.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: Henry & Company.
- O'Reilly, R. C. & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience*. Cambridge: MIT Press.

saray.ayala@uab.es