

## Adquisición de significados verbales nuevos por un proceso de asociación multimodal

María Eugenia Fernández Feijoo,<sup>1</sup> Sergio Vernis,<sup>1</sup> Silvano Zanutto,<sup>1,2</sup> Alejandro Wainselboim<sup>1</sup>

*1 Laboratorio de Biología del Comportamiento, IByME-CONICET*

*2 Instituto de Ingeniería Biomédica, Facultad de Ingeniería, UBA*

### Resumen

*Mientras que la adquisición del lenguaje durante la infancia se produce por exposición al contexto lingüístico, en adultos esta capacidad no parecería estar presente, pues la adquisición de las reglas gramaticales de una lengua requiere de procesos de aprendizaje explícitos (concientes). Se ha visto que infantes de 4 a 8 meses pueden asociar información visual con información auditiva de tipo lingüística con breves tiempos de presentación. La mayor dificultad en adultos por adquirir una segunda lengua por simple exposición al contexto lingüístico podría quizás en parte deberse a una dificultad en asociar entradas sensoriales múltiples con el estímulo lingüístico adecuado. Se estudió entonces qué capacidades poseen los adultos de asociar significados a formas verbales nuevas por co-ocurrencia entre acciones y formas lexicales nuevas. Se presentaron como entrenamiento 60 escenas visuales diferentes junto con frases formadas por no palabras que describían las escenas mostradas. Durante el testeó se presentaron 40 escenas nuevas junto con frases nuevas. En el 50% de los casos la frase describía correctamente la escena mostrada (frase Correcta), mientras que en el restante 50% el "verbo" no coincidía con la acción mostrada (frase Incorrecta). Los participantes decidían lo más rápidamente posible si la frase describía correctamente la escena o no. Los resultados mostraron que en adultos es posible adquirir significados verbales nuevos por co-ocurrencia entre acciones (movimientos) y formas lexicales nuevas. Diferencias de respuesta halladas entre tipos de frase durante el testeó permiten especular con respecto a los mecanismos involucrados en el proceso de decisión durante esta fase.*

*Palabras clave:* adquisición del lenguaje, asociación multimodal, significados verbales.

### Abstract

*Infant language acquisition develops by simple exposure to the linguistic context, while in adults acquisition of the grammatical rules of a language requires explicit (i.e. conscious) learning processes. It has been shown that 4 to 8 month old infants can associate visual information with auditory information that has a linguistic form with only brief exposure times. The greater difficulty in adults in acquiring a second language by simple exposure may be in part due to difficulties in associating multiple sensory inputs with the adequate linguistic stimulus. We studied therefore the capacity in adults to assign meaning to new verbal forms by co-occurrence between actions and new lexical labels. During training 60 different visual scenes were presented along*

---

Correspondencia con el autor: [wainselboim@dna.uba.ar](mailto:wainselboim@dna.uba.ar)

El presente trabajo fue financiado por la Agencia Nacional de Promoción Científica y Técnica (PICT 2485) y la Universidad de Buenos Aires (P014).

*with sentences formed by non-words that described the presented scenes. During the testing phase 40 new scenes were presented together with new sentences. In 50% of the cases the sentence described correctly the scene (Correct sentence), while in the remaining 50% the “verb” did not agree with the shown action (Incorrect sentence). Participants had to decide as fast and accurately as possible whether each sentence correctly described each visual scene. Results show that adults can assign new verbal meanings by co-occurrence between actions (movements) and new lexical labels. Differences in response patterns between both types of sentences during testing allow speculating about the mechanisms involved during the decision process in this phase.*

*Key words:* language acquisition, multimodal association, verbal meaning.

## **Introducción**

La adquisición del lenguaje durante la temprana infancia se produce por exposición al contexto lingüístico sin necesidad de que al infante se le expliciten las reglas gramaticales del idioma al que se ve expuesto, en un proceso que parecería involucrar mecanismos de aprendizaje implícito (Reber, 1989). En el adulto, esta capacidad no parecería estar presente en el mismo grado, puesto que la adquisición de las reglas gramaticales de una lengua requiere de procesos de aprendizaje concientes y explícitos. Las bases neuropsicológicas de esta diferencia de capacidad no se conocen con exactitud hasta este momento. Diversos estudios han analizado la posibilidad de que la adquisición del lenguaje en infantes involucre el análisis estadístico de la señal lingüística (Aslin et al., 1998; Gómez y Gerken, 1999; Saffran et al., 1996). Los infantes pueden abstraer patrones de ordenamientos secuenciales per se y no sólo memorizar las secuencias presentadas durante el entrenamiento (Gómez y Gerken, 1999; Marcus et al., 1999). Adicionalmente, el infante debe agrupar dichas unidades en las distintas categorías gramaticales que posee el idioma (sustantivo, verbo, pronombre, etc), cabiendo la posibilidad que la categorización de los vocablos se lleve a cabo al menos parcialmente en función de las características formales del estímulo, como pueden ser marcadores fonológicos (e.g. patrones de acentuación, número de sílabas); o morfológicos (sufijos flexivos y derivativos). Experimentos llevados a cabo con gramáticas artificiales en adultos muestran que es posible abstraer una categoría lingüística en ausencia de contenido semántico si un porcentaje mayoritario de los miembros de dicha categoría comparten alguna característica formal común (Brooks et al., 1993; Frigo y MacDonald, 1998). Los patrones de co-ocurrencia de palabras dentro de una gramática artificial también parecen ser una fuente de información tomada en cuenta tanto por adultos como por niños para categorizar adecuadamente las principales clases gramaticales (Mintz, 2002; Monaghan et al., 2005).

Por otra parte, la adquisición de la lengua materna se realiza en un contexto ambiental caracterizado por una estrecha relación entre la interacción social y el estímulo lingüístico (Lacerda et al., 2004a; Lacerda et al., 2004b; Lacerda y Sundberg, 2006; Zukow-Goldring, 1990, 1996, 1997; Zukow-Goldring y Rader, 2001), por lo que este proceso parecería involucrar una asociación multimodal entre las distintas entradas sensoriales (principalmente visuales y auditivas) que procesa el individuo. Recientemente se ha comenzado a estudiar en infantes (3 a 12 meses de edad) las capacidades que poseen de asociar información visual del entorno con información auditiva de tipo lingüística, analizó si este mecanismo general de aprendizaje asociativo

es capaz de explicar la adquisición de términos lingüísticos que denoten objetos, cualidades físicas, o más recientemente, acciones (Koponen et al., 2003; Gustavsson et al., 2004, Klintfors y Lacerda, 2006; Marklund y Lacerda, 2006). La metodología utilizada involucra un primer período de exposición, durante el cual se le presentan al sujeto de manera simultánea un conjunto de imágenes en una pantalla de computadora durante un tiempo de 30 s, con el objetivo de medir el tiempo de fijación visual espontánea (medida de la preferencia) ante cada una de ellas. Posteriormente, durante el período de “entrenamiento”, cada imagen es presentada de forma individual de manera simultánea con frases cortas en formato auditivo. Las frases son formadas por 2 a 4 no-palabras compuestas de acuerdo a las reglas fonotácticas del idioma materno del sujeto, y grabadas previamente por el experimentador. Cada una de las frases empleadas incluye una no-palabra que denota ya sea un objeto (eg. “marioneta”), o una cualidad física como forma o color, presente en la imagen mostrada simultáneamente. Durante el testeo posterior se presentan nuevamente de manera simultánea las imágenes anteriores junto con nuevas frases que poseen una de las no-palabras vinculada en la fase de entrenamiento con una de las imágenes. Como medida del grado de asociación alcanzada por el infante entre imagen y no-palabra, se toma el tiempo que el sujeto fija la vista sobre la imagen que posee el objeto o cualidad física denotada dentro de la frase presentada simultáneamente. Un aumento significativo del tiempo de fijación visual con respecto al espontáneo para dicha imagen en el período de exposición inicial (previo al “entrenamiento”), es tomado como indicador de que el infante asoció el término novedoso con el objeto o cualidad física denotada. Los resultados obtenidos parecen mostrar que con tiempos de presentación audiovisual simultánea de entre 60 a 75 segundos el infante es capaz de asociar un término nuevo con el objeto o cualidad física correspondiente, postulándose que el proceso de adquisición de la lengua materna involucra mecanismos de aprendizaje de naturaleza implícita por medio de los cuales el infante correlaciona las entradas multisensoriales y lingüísticas presentes en la interacción con el adulto (Lacerda et al., 2004b). En cuanto al aprendizaje de términos que denoten acciones (“verbos”), se ha visto recientemente que la capacidad de asociar estímulos auditivos novedosos con imágenes que representen acciones, aumenta progresivamente con la edad en infantes de entre 4 y 8 meses (Marklund y Lacerda, 2006), encontrándose significativamente por debajo de la capacidad de asociar formas lingüísticas nuevas con objetos (“sustantivos”) o atributos sensoriales como forma o color (“adjetivos”). Estos resultados parecerían indicar una mayor dificultad implicada en el proceso de adquisición de los verbos en comparación con sustantivos y adjetivos, coincidiendo con resultados obtenidos en el análisis del orden de adquisición de los diferentes tipos de palabras en el lenguaje materno, en donde la adquisición de sustantivos precede a la de verbos como clase gramatical (Gentner, 1978, 1982, D’Odorico y Fasolo, 2007).

Dados los resultados descriptos en infantes y puesto que como fue mencionado anteriormente la adquisición de una primera lengua (L1) en infantes y una segunda lengua (L2) en adultos parecerían involucrar mecanismos de aprendizaje diferentes, es posible pensar que quizás las capacidades de asociación entre estímulos multimodales y de tipo lingüístico sean distintas en infantes y adultos. La mayor dificultad en adultos por adquirir una L2 por simple exposición al contexto lingüístico podría quizás en parte deberse a una dificultad en asociar entradas sensoriales de distinto tipo con el estímulo lingüístico adecuado. Se plantea por tanto estudiar qué capacidades poseen los adultos

de poder asociar significados a formas verbales nuevas por co-ocurrencia entre acciones y formas lexicales nuevas.

## **Metodología**

### ***Sujetos***

20 adultos sanos de entre 20 y 33 años de edad (media  $27,2 \pm 0,7$  años) de nivel educativo terciario participaron voluntariamente en el experimento.

### ***Estímulos***

Se realizó un protocolo para estudiar en adultos la capacidad de asociar información visual de acciones con formas “verbales” nuevas. Para ello se crearon 100 escenas visuales diferentes de  $3,34 \pm 0,01$  segundos de duración media con las distintas combinaciones posibles de movimientos y figuras geométricas de la Tabla 1. A su vez, se crearon 100 frases diferentes describiendo diferentes combinaciones de movimientos y figuras de la Tabla 1. Cada frase seguía el siguiente patrón:

BAPA-Sujeto-Verbo-NALA-Objeto

En donde tanto el sujeto como el objeto de cada frase referían a una de las figuras geométricas utilizadas.





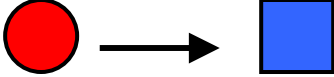
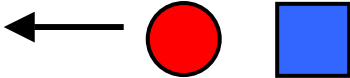
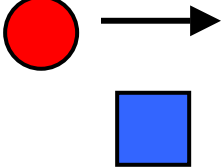
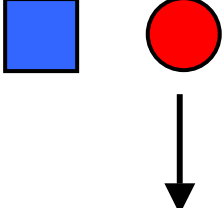
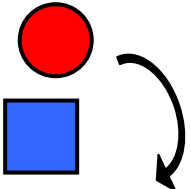
### ***Procedimiento***

#### ***Entrenamiento***

Durante la realización del experimento cada participante se encontraba sentado en un cuarto aislado, frente a un monitor de computadora de 17 pulgadas a una distancia de aproximadamente 0,5 m. La fase de entrenamiento comenzaba con una pantalla en donde se le indicaba al participante que el objetivo del experimento era que aprendiera el significado de las distintas palabras de un idioma inventado. Para ello se le presentarían diversas escenas junto con frases que las describían en el idioma desconocido. A partir de la pantalla inicial, cada ensayo consistía en la presentación en el centro de la pantalla de una escena visual diferente. De manera simultánea, por debajo de cada escena aparecía una frase que describía la escena mostrada. La frase aparecía palabra por palabra de izquierda a derecha, manteniéndose en pantalla cada palabra hasta la compleción de la frase. Cada frase además de aparecer en formato visual se presentaba simultáneamente en formato auditivo. En esta etapa se presentaron 60 ensayos diferentes, siendo cada ensayo presentado dos veces consecutivas. El intervalo entre ensayos era de 2000 ms.

Para asegurar que los sujetos prestaran atención a los estímulos presentados, cada 4 a 7 frases se le preguntaba en formato visual si determinada palabra había aparecido en la frase anterior. Para responder debían apretar Ctrl Derecho para “Sí”, y Ctrl Izquierdo para “No”. La duración total de la fase de entrenamiento era de aproximadamente 15 minutos.

**Tabla 1. Estímulos utilizados**

Elemento/Acción	Palabra	Significado
	<b>Bapa</b>	<b>“El”</b>
	<b>Nala</b>	<b>“Hacia el”</b>
	<b>Bare</b>	
	<b>Lane</b>	
	<b>Fadi</b>	
	<b>Pefa</b>	
	<b>Siru</b>	<b>“Acercarse”</b>
	<b>Revu</b>	<b>“Alejarse”</b>
	<b>Fenu</b>	<b>“Pasar horizontalmente”</b>
	<b>Melu</b>	<b>“Pasar verticalmente”</b>
	<b>Jazu</b>	<b>“Girar alrededor”</b>

### *Testeo*

Una vez completado el entrenamiento, una pantalla indicaba que a partir de ese momento comenzaría el testeo con el objetivo de ver si el participante había aprendido el significado de las diferentes palabras. A continuación, y de manera similar a lo mostrado en el entrenamiento, cada ensayo consistía en la presentación de una escena visual nueva (no mostrada en el entrenamiento) junto con una frase nueva en formato audiovisual. Sin embargo, a diferencia del entrenamiento, sólo en el 50% de los casos la frase describía correctamente la escena mostrada (frase Correcta), mientras que en el restante 50% el “verbo” no correspondía con el movimiento mostrado en la escena visual (frase Incorrecta). Los sujetos debían decidir lo más rápidamente posible si la frase describía correctamente la escena o no. En caso afirmativo debían apretar Ctrl Derecho, y en caso negativo Ctrl Izquierdo. El tiempo límite para responder era 2000 ms luego de la aparición completa de cada frase.

Como medidas comportamentales se tomaron el número de respuestas acertadas y el tiempo de respuesta ante cada frase medido desde la aparición del verbo, por ser éste el estímulo clave en la aparición del error en la frase.

### *Estadística*

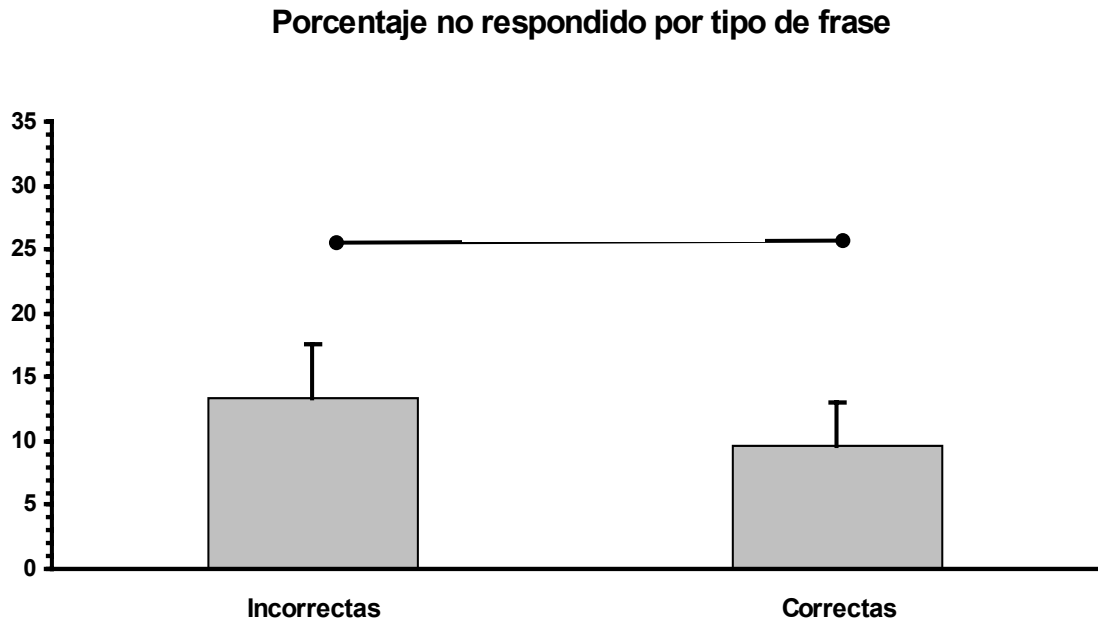
Con el objetivo de establecer qué participantes respondieron de manera acertada por encima del nivel azaroso (50%) se realizó un test de Heterogeneidad con el número de aciertos y desaciertos por individuo. Las respuestas de los individuos que no respondieron significativamente por encima del 50% fueron descartadas para los análisis comportamentales posteriores. Para el resto de los participantes se discriminó el porcentaje de frases Correctas e Incorrectas que no fueron respondidas, el porcentaje de respuestas acertadas ante los dos tipos de frases y el tiempo de respuesta promedio de los aciertos ante frases Correctas e Incorrectas. Los datos individuales fueron ingresados en un análisis poblacional realizándose un test de t pareado para cada una de las tres variables descriptas previamente.

## **Resultados**

Los resultados obtenidos muestran que 15 de los 20 participantes (75%) respondieron significativamente por encima del azar durante la fase de testeo (test de Heterogeneidad,  $p < 0.05$ ), demostrando de esta manera que la mayoría de los individuos fueron capaces de aprender el significado de los 5 verbos presentados durante el entrenamiento. El porcentaje de aciertos para el grupo de 15 participantes que aprendió fue de  $90,17 \pm 2,78\%$ , siendo el número promedio de respuestas efectuadas  $35,33 \pm 1,49$ .

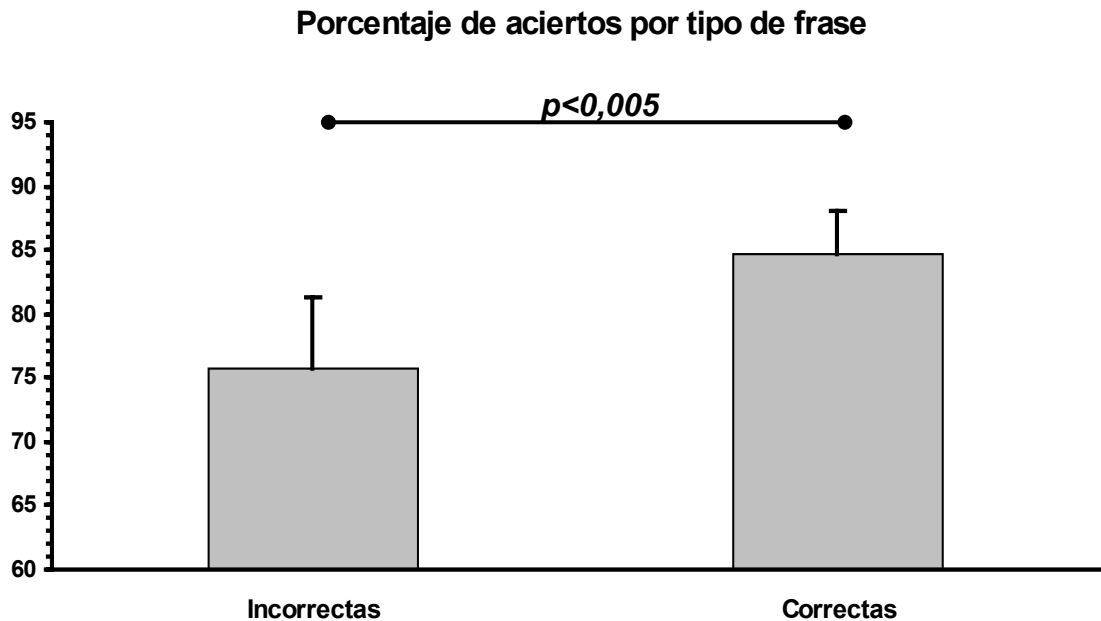
El porcentaje de respuestas efectuadas fue mayor para las frases Correctas que para las Incorrectas (Fig. 1), siendo esta diferencia marginalmente significativa ( $t=2.13$ ,  $p=0.52$ ).

**Figura 1. Porcentaje de frases incorrectas y correctas que no fueron respondidas durante la fase de testeo.**

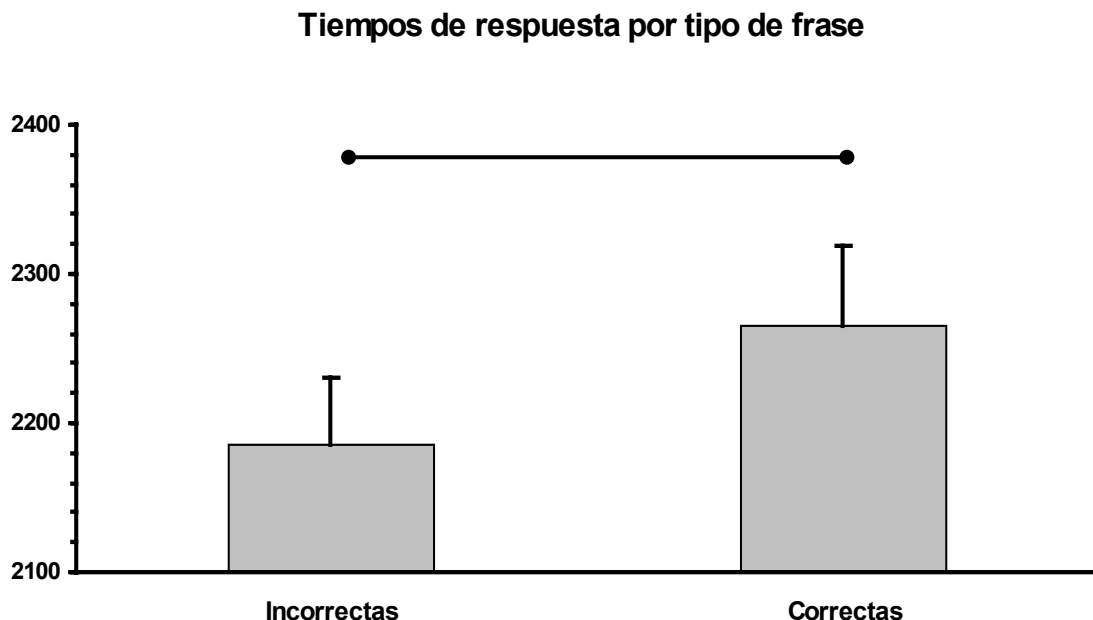


Por su parte, tanto el porcentaje de aciertos (Fig. 2) como el tiempo de respuesta (Fig. 3) fueron significativamente mayores para las frases Correctas comparadas con las Incorrectas ( $t=-3,47$ ,  $p<0,005$  y  $t=-3,09$ ,  $p<0,01$  respectivamente).

**Figura 2. Porcentaje de aciertos durante la fase de testeo discriminado por tipo de frase.**



**Figura 3. Tiempo de respuesta de los aciertos durante la fase de testeo discriminado por tipo de frase.**



## Discusión

La mayoría de los estudios sobre adquisición de lenguaje se han centrado en el análisis de los procesos de aprendizaje involucrados en la adquisición de las reglas gramaticales que subyacen al uso de un lenguaje. Mediante el empleo de gramáticas artificiales con reglas combinatorias arbitrarias entre los elementos constituyentes del “idioma” se ha podido ver que tanto infantes como adultos son capaces de aprender dichas reglas por medio de un tipo de aprendizaje que se ha denominado “implícito” por no existir conciencia por parte del individuo de los mecanismos involucrados durante el proceso de aprendizaje (Knowlton y Squire, 1994, 1996; Reber, 1967, 1976, 1989). Operativamente el sujeto es capaz de tomar decisiones en función del aprendizaje logrado, pero no es capaz de explicar qué es exactamente lo que aprendió (Beery y Dienes, 1993; Ellis, 1994). Contrariamente, el aprendizaje explícito se refiere a aquél en el cual el sujeto tiene conciencia de al menos parte de los mecanismos implicados durante el proceso de aprendizaje y es por tanto capaz de verbalizar en qué consistió este proceso (Ellis, 1994).

La adquisición de vocabulario nuevo y los mecanismos de aprendizaje involucrados en este proceso ha recibido considerablemente menos atención en comparación con el punto anterior. La distinción entre aprendizaje implícito y explícito no parece ser tan clara en este caso como en el aprendizaje de reglas gramaticales. Estudios recientes muestran que en infantes de menos de 1 año de edad es posible asociar objetos y cualidades como color y forma a formas fonológicas nuevas al presentarse de manera pareada las imágenes con el estímulo auditivo (Koponen et al., 2003; Gustavsson et al., 2004, Klintfors y Lacerda, 2006). Se asume que los mecanismos involucrados son de



naturaleza implícita por asumirse que no existe conciencia por parte del infante del proceso involucrado en este aprendizaje. Los resultados del presente trabajo muestran que en sujetos adultos es posible adquirir significados verbales nuevos por medio de la presentación simultánea de acciones (movimientos) y formas ortográficas y fonológicas nuevas. A diferencia de los infantes, en los participantes adultos el proceso de aprendizaje fue conciente, por lo que correspondería a la definición clásica de explícito. Sin embargo, en el aprendizaje de una L2 en adultos, los procesos habituales de adquisición de vocabulario involucran normalmente hacer directamente explícito el significado de la palabra a aprender, ya sea mediante definiciones o directamente por medio de traducciones a la L1. En el presente estudio los participantes debían inferir el significado de los diferentes términos a través de asociar los lexemas con los diferentes aspectos de las escenas presentadas, tales como formas y movimientos. En este sentido, el tipo de aprendizaje involucrado en la adquisición de vocabulario en el adulto más similar al aquí presentado, es el denominado aprendizaje incidental (Huckin y Cody, 1999). En este caso, el significado de la palabra desconocida se produce por un proceso de inferencia a partir de la información disponible en el texto dentro del cual se encuentra el término nuevo. Este tipo de adquisición de vocabulario parece involucrar tanto procesos concientes (explícitos) como inconscientes (implícitos) por parte del sujeto (Rieder, 2003).

Los resultados obtenidos durante la fase de testeo mostraron varias diferencias cuantitativas en el comportamiento de los participantes entre los dos tipos de frases presentadas. En primer término respondieron menos veces (ya sea acertada como desacertadamente) ante las frases incorrectas en comparación con las correctas. Dado que existía un tiempo límite para efectuar la respuesta (2000 ms post culminación de la frase) el resultado obtenido parece indicar que los participantes tardaron más en decidir qué respuesta efectuar cuyo la frase presentada era incorrecta. La aparición de la escena visual, con el movimiento correspondiente, precedía en alrededor de  $1335 \pm 27$  ms a la aparición de la forma lexical correspondiente al "verbo". Es posible pensar entonces que la aparición de cada escena lleve a la activación de uno o más potenciales candidatos lexicales de manera previa a la aparición de la forma lexical per se. De esta manera, la coincidencia entre el lexema presentado y la/s forma/s activadas llevaría a una decisión por parte del individuo de que la frase es correcta, mientras que la no coincidencia entre lo presentado y lo esperado (lexemas activados) llevaría a la decisión de que la frase es incorrecta. Dentro de este esquema general es posible pensar que en este último caso, ante la falta de coincidencia entre lo presentado y lo esperado, pueda haber un tiempo adicional de decisión que esté relacionado con mecanismos de recomparación entre la/s forma/s activada/s y la presentada. Esta hipótesis necesitaría ser corroborada mediante estudios adicionales.

El porcentaje de respuestas acertadas fue significativamente mayor para las frases correctas en comparación con las incorrectas. Este resultado indica que la probabilidad de que los sujetos decidieran que una frase incorrecta era correcta fue mayor a la probabilidad de que decidieran que una frase correcta era incorrecta. Como fue mencionado previamente, es posible que la aparición de la escena visual active más de un candidato lexical puesto que el grado de aprendizaje logrado no fue perfecto. En este caso, dada la activación de más de un lexema candidato, habría una alta probabilidad de que uno de ellos fuera el correcto y uno o más fueran incorrectos. En el caso de la presentación de una frase correcta, habría una alta probabilidad de coincidencia entre uno de los candidatos activados y el presentado, siendo por tanto alta la probabilidad de

que el sujeto responda acertadamente que la frase presentada es correcta. Sin embargo, en el caso de la presentación de una frase incorrecta, si el lexema presentado en la frase coincide con uno de los candidatos lexicales incorrectos ya activados, el sujeto decidiría que la frase es correcta aún cuando no lo sea. De esta manera, existiría una mayor probabilidad de dar como correcta una frase incorrecta que viceversa.

En cuanto a los tiempos de respuesta, los resultados mostraron que los participantes tardaron significativamente más en responder acertadamente que una frase era correcta. Dado que a los participantes no se les instruyó que los errores de las frases estarían en el verbo, el resultado obtenido podría estar relacionado con que durante la presentación de una frase correcta, el sujeto debería esperar necesariamente hasta la compleción de la frase hasta poder decidir que no se presentaban errores, mientras que durante la presentación de una frase incorrecta la aparición del error en la posición del verbo podría disparar el proceso de decisión antes que en el caso de las frases correctas.

La naturaleza de las diferencias en patrones de respuesta halladas entre los dos tipos de frases permiten especular con respecto a los mecanismos involucrados en el proceso de decisión durante la fase de testeo. La corroboración de las hipótesis planteadas requerirá de estudios adicionales y de metodologías que permitan analizar con mayor detalle los procesos subyacentes al aprendizaje logrado y a los procesos de decisión durante el testeo. En este sentido el análisis de las señales eléctricas originadas en la actividad cerebral cortical (electroencefalograma o EEG) durante la realización de las tareas experimentales planteadas puede ser una valiosa herramienta de estudio dada su alta resolución temporal. Este tipo de análisis se encuentra actualmente en curso en nuestro grupo.

## Bibliografía

- Aslin, R., Saffran, J., y Newport, E. (1998). Computation of conditional probability statistics by 8-months-old infants. *Psychological Science* 9(4), 321-324.
- Berry, D.C., y Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. New Jersey: Lawrence Erlbaum Associates.
- Brooks, P.B., Braine, M.D.S., Catalano, L., Brody, R.E., y Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language* 32, 79-95.
- D'Odorico, L., y Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language* 34, 891-907.
- Ellis, N.C. (1994). Introduction: implicit and explicit language learning – an overview. En N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-31). London: Academic Press.
- Frijo, L., y McDonald, J.L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language* 39, 218-245.
- Gentner, D. (1978). On relational meaning : The acquisition of verb meaning. *Child Development* 49, 988-998.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. En S. Kuczaj (Ed.), *Language development (2), Language, thought and culture* (pp. 301-334). New Jersey: Lawrence Erlbaum Associates.

- Gomez, R., y Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135.
- Gustavsson, L., Sundberg, U., Klintfors, E., Marklund, E., Lagerkvist, L., y Lacerda, F. (2004). Integration of audio-visual information in 8-months-old infants. En L. Berthouze, H. Kozima, C.G. Prince, G. Syini, G. Stojanov, G. Metta, y C. Balkenius (Eds.). *Proceedings of the Fourth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (pp.143-144). Genova.
- Huckin, T., y Coady, J. (1999). Incidental vocabulary acquisition in a second language: a review. *Studies in Second Language Acquisition* 21, 181-193.
- Klintfors, E. y Lacerda, F. (2006). Potential relevance of audio-visual integration in mammals for computational modeling. *Ninth International Conference on Spoken Language Processing* (pp. 1403-1406). Pittsburgh.
- Knowlton, B.J., y Squire, L.R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 79–91.
- Knowlton, B.J., y Squire, L.R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 169–181.
- Koponen, E., Gustavsson, L, y Lacerda, F. (2003). Effects of linguistic variance on sound-meaning connections in early stages of language acquisition. 15th ICPHS (pp. 1975-1978). Barcelona.
- Lacerda, F., y Sundberg, U. (2006). An Ecological Theory of Language Acquisition. *Revista de Estudos linguísticos da Universidade do Porto* 1, 53-106.
- Lacerda, F., Klintfors, E., Gustavsson, L., Lagerkvist, L., Marklund, L. y Sundberg, U. (2004a). Ecological Theory of Language Acquisition. En L. Berthouze, H. Kozima, C.G. Prince, G. Syini, G. Stojanov, G. Metta, y C. Balkenius (Eds.). *Proceedings of the Fourth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (pp.147-148). Genova.
- Lacerda, F., Marklund, E., Lagerkvist, L., Gustavsson, L., Klintfors, E. y Sundberg U. (2004b). On the linguistic implications of context-bound adult-infant interactions. En L. Berthouze, H. Kozima, C.G. Prince, G. Syini, G. Stojanov, G. Metta, y C. Balkenius (Eds.). *Proceedings of the Fourth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (pp.149-150). Genova.
- Marcus G., Vijayan, S., Byi Rao, S., Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80.
- Marklund, E., y Lacerda, F. (2006). Infants' ability to extract verbs from continuous speech. *Ninth International Conference on Spoken Language Processing* (pp. 1403-1406). Pittsburgh.
- Mintz, T.H. (2002). Category induction from distributional cues in an artificial language. *Memory and Cognition* 30: 678-686.
- Monaghan, P., Chater, N., y Christiansen, M.H. (2005). The differential contribution of phonological and distributional cues in grammatical categorisation. *Cognition*, 96, 143-182.
- Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, 81, 115–119.

- Reber, A.S. (1976). Implicit learning of synthetic language. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 88-94.
- Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219-235.
- Rieder, A.G. (2003). Implicit and explicit learning in incidental vocabulary acquisition. *Views* 12(2), 24-39.
- Saffran, J.R., Aslin, R.N., Newport, E.L. (1996). Statistical learning by eight-month-old infants. *Science* 274, 1926-1928.
- Zukow, P.G. (1990). Socio-perceptual bases for the emergence of language: An alternative to innatist approaches. En C. Dent, y P.G. Zukow (Eds.), *The idea of innateness: Effects on language y communication research. Developmental Psychobiology*, 23, 705- 726.
- Zukow-Goldring, P.G., y Rader, N. (2001). Perceiving referring actions. *Developmental Science*, 4, 28-30.
- Zukow-Goldring, P. (1996) Sensitive caregivers foster the comprehension of speech: When gestures speak louder than words. *Early Development and Parenting*, 5 (4), 195-211.
- Zukow-Goldring, P. (1997) A social ecological realist approach to the emergence of the lexicon: Educating attention to amodal invariants in gesture and speech. En C. Dent-Read y P. Zukow-Goldring (Eds.), *Evolving explanations of development: Ecological approaches to organism-environment systems* (pp. 199-250). Washington, D. C.: American Psychological Association.