

## **Análisis comportamental y electrofisiológico del aprendizaje inferencial de términos verbales nuevos: influencia del formato de presentación**

Pamela Lopes da Cunha<sup>1</sup>, Silvano Zanutto<sup>1,2</sup> y Alejandro Wainseboim<sup>1</sup>

<sup>1</sup>Laboratorio de Biología del Comportamiento, Instituto de Biología y Medicina Experimental-CONICET. Buenos Aires, Argentina

<sup>2</sup>Instituto de Ingeniería Biomédica, Facultad de Ingeniería, Universidad de Buenos Aires. Argentina

### **Resumen**

*Se estudió en adultos la influencia del formato de presentación en la capacidad de adquirir términos verbales nuevos mediante aprendizaje inferencial. Para ello se presentaron en computadora 70 escenas visuales distintas con figuras geométricas que realizaban distintos movimientos. Simultáneamente cada escena era descrita por una frase en un lenguaje artificial presentada en formato auditivo. Los participantes debían aprender cuáles términos denotaban cada uno de los movimientos presentados. Durante el testeo se presentaron 80 escenas con nuevas combinaciones de las figuras y movimientos del entrenamiento, simultáneamente con frases que las describían correctamente (condición "Correcta", 50% de ensayos) o que poseían un "verbo" incongruente con el movimiento de la escena (condición "Incorrecta", 50% de ensayos), mientras se registraba su actividad electroencefalográfica. Los participantes debían decidir en menos de 2 s, si la frase describía o no la escena. La mayoría de los participantes (67%) respondieron significativamente por encima del azar durante el testeo, siendo capaces de adquirir la estadística de co-ocurrencia entre términos y movimientos aún cuando la información lingüística fue presentada sólo en formato auditivo. Sin embargo, la ausencia de la información ortográfica (presente en protocolos anteriores) dificultó el proceso de asociación multimodal entre la escena visual y el término auditivo. El registro de los potenciales evocados mostró un componente*

---

Correspondencia con los autores: [awainseboim@mendoza-conicet.gob.ar](mailto:awainseboim@mendoza-conicet.gob.ar)

Artículo recibido: 30 de agosto de 2014

Artículo aceptado: 8 de noviembre de 2014

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

*de tipo ELAN de 50 a 500 ms posteriores a la aparición de la incongruencia entre “verbo” y movimiento. La generación de dicho potencial podría encontrarse relacionado con la detección de una incongruencia formal entre el término esperado y el observado.*

*Palabras clave:* Aprendizaje inferencial - formato de presentación - ELAN.

## **Abstract**

*We studied the influence of presentation format on the ability to acquire new verbal meanings by inferential learning in adults. 70 different visual scenes were presented by computer. Each scene consisted of geometrical figures that could perform different movements. Simultaneously, each scene was described by a sentence in an artificial language that was presented auditorily. Participants had to learn which non-word referred to each movement. During testing, 80 additional scenes with new combinations of the already presented figures and movements were presented simultaneously with sentences that either described the scene correctly (“Correct” condition, 50% of trials), or presented a “verb” that was incongruent with the movement of the scene (“Incorrect” condition, 50% of trials), while the electroencephalographic activity of the participant was recorded. Most participants (67%) responded significantly above chance during the test, showing that they were able to acquire the statistics of co-occurrence between “verbal” non-word and movements, even though linguistic information was presented only auditorily. Nevertheless, the absence of orthographic information (which was present in previous works) increased the difficulty of the process of multimodal association between visual scene and auditory term. Evoked potential recordings showed an ELAN type component 50 to 500 ms after presentation of the erroneous “verb”. The appearance of this component could be related to the detection of a formal incongruence between the expected and the presented non-word.*

*Key words:* Inferential learning - presentation format - ELAN.

## **1. Introducción**

La adquisición del lenguaje durante la infancia se produce por exposición al contexto lingüístico sin que al infante se le expliciten las reglas gramaticales del idioma al que se ve expuesto, ni la mayor parte del vocabulario adquirido. En adultos, esta capacidad no parecería estar desarrollada en el mismo grado, pues la adquisición de las reglas gramaticales de una lengua requiere de procesos de aprendizaje conscientes y explícitos. Las bases neuropsicológicas de esta diferencia de capacidad no se conocen con exactitud, y es posible que ambos grupos se basen en diferentes mecanismos de aprendizaje.

Estudios previos 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 & 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 & Gerken 1999; Marcus et al. 1999). A su vez, 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 los patrones de co-ocurrencia de palabras dentro de una gramática artificial puede 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).

A su vez la capacidad de analizar la estadística de co-ocurrencia entre estímulos visuales y auditivos, parece subyacer también al aprendizaje de significados lingüísticos nuevos en contextos que presentan potencial ambigüedad. Experimentos llevados a cabo por Yu y Smith (2007) presentaron por computadora sucesivos ensayos de entrenamiento, cada uno consistente en dos, tres o cuatro imágenes simultáneas de distintos objetos presentadas al mismo tiempo que se escuchaban una cantidad de términos inventados (no-palabras) igual al número de imágenes mostradas en ese momento. La tarea del sujeto consistía en poder aprender cuáles eran las no-palabras que denotaban cada una de las imágenes presentadas, indicando cada no-palabra, dadas las características del protocolo, sólo una de las imágenes presentadas, manteniéndose esta correspondencia unívoca a lo largo de todo el experimento. Cada uno de estos pares “no-palabra-imagen” se presentaba en 6 ensayos de manera intercalada con los demás pares. Durante el testeo, cada ensayo consistía en la presentación simultánea de cuatro de los objetos del entrenamiento junto con uno solo de los términos del entrenamiento, debiendo entonces elegir el participante cuál era el referente visual de la no-palabra escuchada en cada ensayo. Los resultados mostraron que si bien no es posible que los participantes puedan inferir el significado de los términos nuevos con un solo ensayo de entrenamiento, los sujetos sí fueron capaces de aprender los referentes visuales de los 18 términos inventados. Para lograrlo, los participantes debieron inferir el significado de cada término manteniendo entre ensayos la estadística de la co-ocurrencia entre imagen y estímulo auditivo. Experimentos similares realizados en infantes (12 y 14 meses de edad) mostraron que en este grupo, 10 pareos de imagen y no-palabra durante el entrenamiento bastaban para denotar un aprendizaje de los términos nuevos durante el testeo subsecuente (Smith & Yu 2008). Tanto en adultos como infantes, este tipo de aprendizaje inferencial se ha denominado aprendizaje estadístico trans-situacional, y se ha propuesto como uno de los fundamentos del aprendizaje de términos nuevos durante la adquisición del lenguaje materno en la infancia (Siskind 1996; Lacerda &

Sundberg 2006; Yu 2008; Yu & Smith 2007). Puesto que la adquisición de la lengua materna se da 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 & Sundberg 2006; Yu & Ballard 2007), este proceso parecería involucrar una asociación multimodal entre las distintas entradas sensoriales (principalmente visuales y auditivas) que procesa el individuo.

En cuanto a la capacidad de asociar estímulos auditivos novedosos con imágenes que representen acciones, se ha visto que en infantes (de 4 a 8 meses; Marklund & Lacerda 2006), se encuentra 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 & Fasolo 2007).

En adultos, se ha visto que es posible adquirir significados verbales nuevos por medio de la presentación simultánea de acciones (movimientos) y formas ortográficas y fonológicas nuevas (Fernández Feijóo et al. 2008). En dicho estudio, a 20 participantes adultos sanos se les presentó por computadora 60 escenas de entrenamiento distintas, cada una compuesta por dos figuras geométricas, una de las cuales se encontraba inmóvil en el centro de la escena, mientras que la figura adicional realizaba uno de seis posibles movimientos en referencia a la figura inmóvil. De manera simultánea a la presentación de la escena se desplegaba en formato escrito (visual) y auditivo una frase palabra por palabra que describía en un idioma artificial la escena presentada. Durante la posterior fase de testeo a los sujetos se les presentaban 40 escenas adicionales con las mismas figuras y movimientos del entrenamiento pero en combinaciones novedosas. A diferencia del entrenamiento, durante el testeo sólo el 50% de las frases presentadas describía la escena mostrada en simultáneo. La tarea de los participantes consistía en responder en tiempos breves (<2 s) si la frase describía la escena mostrada (condición Correcta) o no (condición Incorrecta). En la condición Incorrecta, la incongruencia entre escena y frase se generaba al colocar un “verbo” que denotara un movimiento distinto al mostrado en la escena, si bien los participantes no eran informados de la naturaleza de la posible incongruencia entre escena y frase previamente.

Los resultados mostraron que 15 de los 20 participantes fueron capaces de responder significativamente por encima del azar en ambas condiciones experimentales. Por tanto, fueron capaces de inferir el significado de los términos verbales nuevos mediante el seguimiento de la co-ocurrencia entre términos y escenas del entrenamiento, si bien el aprendizaje no fue perfecto en ningún caso.

Posteriormente se realizó un segundo experimento con el objetivo de estudiar los correlatos neurobiológicos del aprendizaje inferencial de significados verbales nuevos (Lopes da Cunha et al. 2010). Se realizó entonces una modificación del protocolo de

Fernández-Feijóo et al (2008) aumentándose el número de estímulos utilizados. Se crearon 150 escenas visuales diferentes (mediante la combinatoria de 5 figuras geométricas y 6 movimientos distintos), y 150 frases describiendo cada una de las distintas escenas creadas. Durante el entrenamiento se presentaron 70 de las 150 escenas, junto con la correspondiente frase que la describía. La frase aparecía en formato audiovisual por debajo de cada escena, palabra por palabra de izquierda a derecha, y manteniéndose cada término en pantalla hasta la compleción de la frase. En la fase de testeo posterior, se presentaron las 80 escenas restantes (no vistas en el entrenamiento). Al igual que en Fernández-Feijóo et al. (2008), el 50% de las frases describía la escena presentada (“Correcta”) mientras que el restante 50% presentaba un término “verbal” que denotaba un movimiento distinto al visto en la escena simultánea (“Incorrecta”). Durante esta fase se registró la actividad electroencefalográfica (EEG) de los participantes con el objetivo de analizar la posible aparición de potenciales evocados relacionados con la aparición de la incongruencia entre término y movimiento.

Con el protocolo descripto, la mayoría de los individuos fueron capaces de aprender el significado de los 6 verbos presentados durante el entrenamiento (15 de 19). El análisis de los potenciales evocados mostró que la presentación de ensayos incorrectos generaba la aparición de una negatividad centro-parietal, aproximadamente entre los 400 ms y 700 ms luego de la presentación del término “verbal” incongruente con el movimiento de la escena. En experimentos con lenguaje natural, la aparición de una palabra semánticamente incongruente dentro de una frase genera una negatividad con topografía centroparietal que alcanza su máximo alrededor de los 400 ms luego de la aparición de la palabra incongruente (Kutas & Hillyard 1980; 1984; Neville 1985; Holcomb & Neville 1991; Hahne & Friederici 2002), denominándose por ello N400. La aparición de un componente de similares características en el trabajo de Lopes da Cunha et al. (2010) muestra que la incongruencia entre el movimiento mostrado en la escena y el “verbo” presentado en la frase generó un componente de origen cognitivo comparable al observado en protocolos psicolingüísticos clásicos. Validando por tanto el uso del protocolo experimental como herramienta para el análisis del aprendizaje de aspectos léxico-semánticos en condiciones controladas.

Los resultados de los dos experimentos descriptos mostraron que la mayoría de los participantes fueron capaces de inferir el significado de los términos “verbales” presentados. Para poder realizar dicha inferencia (asociando un cierto movimiento con un término particular), los sujetos debieron seguir la estadística de co-ocurrencia entre los movimientos y los términos presentados a lo largo de los ensayos de entrenamiento. Sin embargo, nos planteamos si la presentación audiovisual de los términos en ambos experimentos podría haber facilitado a los sujetos el aprendizaje de los significados nuevos. La influencia del formato de presentación en el nivel de aprendizaje logrado ha sido descripta en otros protocolos de psicología experimental. Por ejemplo, se ha observado que la inducción de memorias lingüísticas falsas disminuye con la presentación visual de los términos en lugar de su presentación

auditiva (Cleary & Greene 2002; Gallo et al. 2001; Kellogg 2001; Pierce et al. 2005; Smith, Hunt, & Gallagher, 2008). En contraposición al formato empleado en los experimentos de Fernández-Feijóo et al (2008) y Lopes da Cunha et al (2010), durante la adquisición de una primera lengua la información lingüística accesible al infante es exclusivamente de tipo auditiva. Por ello, en el presente trabajo nos planteamos estudiar si bajo condiciones experimentales más ecológicas (i.e. información lingüística en formato auditivo), sería posible en sujetos adultos sanos la adquisición de nuevos significados de tipo “verbal”, o si por el contrario, el formato de presentación auditivo disminuiría o incluso podría llegar a impedir el seguimiento de la estadística de co-ocurrencia entre los movimientos y los términos presentados.

## 2. Método








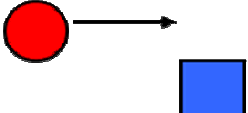
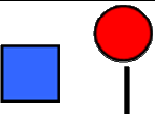
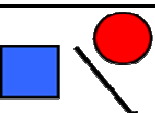
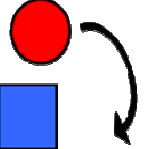
### *Participantes*

24 adultos sanos entre 20 y 35 años de edad (media  $25,7 \pm 0,9$  años), participaron voluntariamente en el experimento. Todos los sujetos eran diestros, hablantes nativos de Español Rioplatense y poseían estudios terciarios en curso o completos. Los sujetos no tenían antecedentes de desórdenes neurológicos o psiquiátricos, ni se encontraban bajo medicación alguna al momento del experimento.

### *Materiales*

Se utilizaron 150 escenas visuales diferentes (duración:  $3,34 \pm 0,01$  s), por medio de la combinación de 5 figuras geométricas y 6 movimientos distintos (Tabla 1), donde una de las figuras se movía en relación a otra inmóvil (Lopes da Cunha et al. 2010). A su vez, se utilizaron 150 frases diferentes (en un lenguaje artificial), que describían cada una de las distintas escenas. Las frases seguían el formato *BAPA-Sujeto-Verbo-NALA-Objeto*, en donde el “sujeto” refería a la figura en movimiento, el “objeto” hacía referencia a la figura inmóvil, y el “verbo” refería al movimiento desplegado en la escena.

**Tabla 1.** Estímulos utilizados para la confección de las frases de entrenamiento y testeo.

<i>ELEMENTO/ACCIÓN</i>	<i>PALABRA</i>	<i>SIGNIFICADO</i>
	<i>Bapa</i>	<i>“El”</i>
	<i>Nala</i>	<i>“Hacia el”</i>
	<i>Bare</i>	
	<i>Lane</i>	
	<i>Fadi</i>	
	<i>Mere</i>	
	<i>Pefa</i>	
	<i>Siru</i>	<i>“Acercarse”</i>
	<i>Revu</i>	<i>“Alejarse”</i>
	<i>Fenu</i>	<i>“Pasar horizontalmente”</i>
	<i>Melu</i>	<i>“Pasar verticalmente”</i>
	<i>Naru</i>	<i>“Pasar en diagonal”</i>
	<i>Jazu</i>	<i>“Girar alrededor”</i>

*Procedimiento*

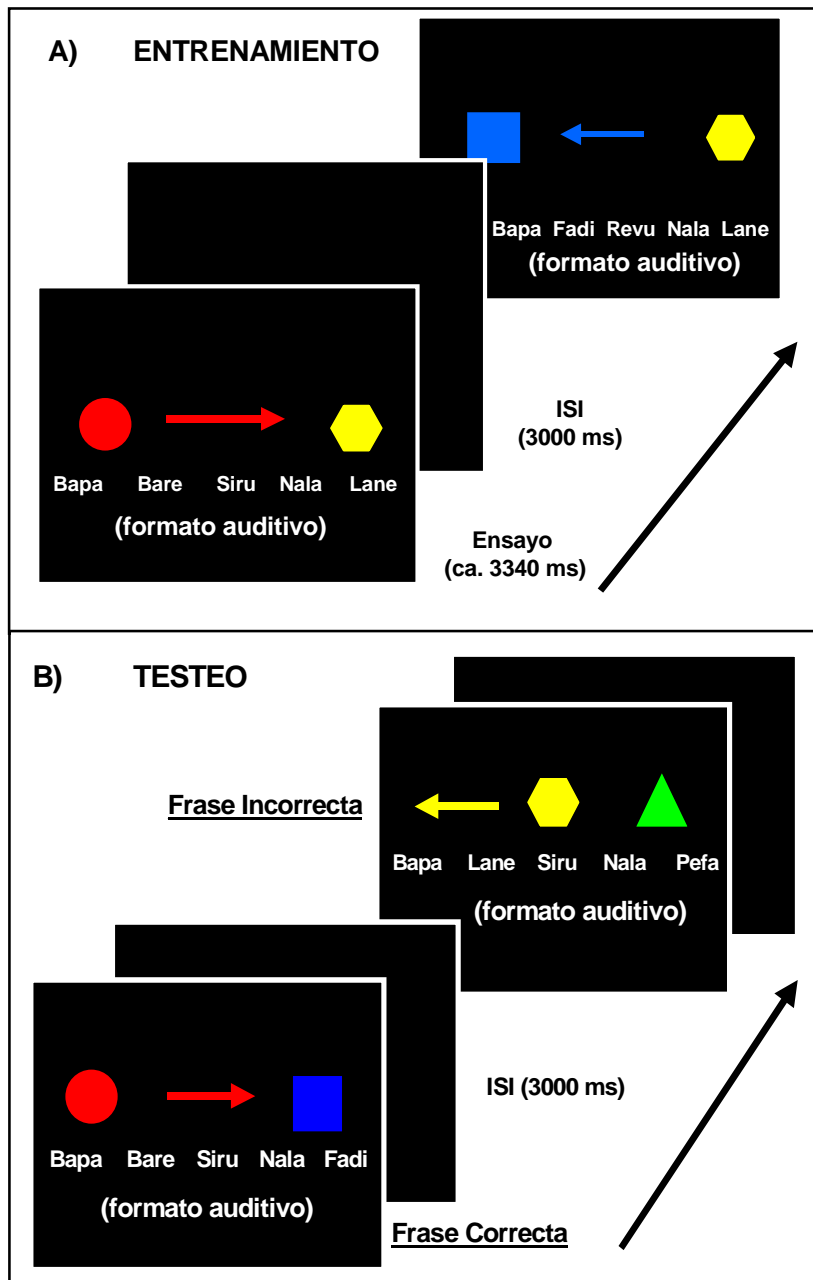
Entrenamiento. El experimento se llevó a cabo en un cuarto aislado eléctrica y

acústicamente, frente a un monitor de computadora de 17 pulgadas a una distancia de aproximadamente 0,5 m. Una pantalla inicial le indicaba al participante que el objetivo del experimento era que aprendiera el significado de las distintas palabras de un lenguaje artificial. Para ello se le presentarían distintas escenas junto con frases que las describían en dicho lenguaje. Luego de la pantalla inicial, cada ensayo consistió en la presentación en el centro de la pantalla de una escena visual distinta, y simultáneamente una frase describiendo dicha escena (Figura 1A). Con el objetivo de presentar una situación experimental más ecológica, i.e. de mayor similitud con la situación de adquisición del lenguaje materno durante la infancia, las frases se presentaron en formato auditivo, siendo el intervalo entre palabras de 400 ms. En esta etapa se presentaron 70 ensayos diferentes. El intervalo entre ensayos fue de 3000 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 fue de aproximadamente 18 minutos.

Testeo. Una vez finalizado el entrenamiento, una pantalla indicaba que a partir de ese momento comenzaría el testeo para probar si el participante había aprendido el significado de las distintas palabras. A continuación, y de manera similar al entrenamiento, cada ensayo consistió en la presentación de una escena visual nueva (no mostrada en el entrenamiento) junto con una frase nueva en formato auditivo. A diferencia del entrenamiento, sólo el 50% de las frases describía correctamente la escena mostrada (frase Correcta, Figura 1B), mientras que en el restante 50% el “verbo” no correspondía con el movimiento mostrado en la escena visual (frase Incorrecta, Figura 1B). Los participantes debían responder lo más rápidamente posible si la frase describía correctamente la escena, presionando en caso afirmativo Ctrl Derecho, y en caso negativo Ctrl Izquierdo. El tiempo límite para responder fue de 2000 ms luego de la presentación completa de cada frase. En total se presentaron las 80 escenas no vistas durante el entrenamiento. Se registró ante cada frase si el participante respondía o no (% de respuestas), acertaba o no (% de aciertos) y su latencia de respuesta medido desde la aparición del “verbo” (Tiempo de respuesta). Durante esta fase se registró el electroencefalograma de los participantes.





**Figura 1.** Ejemplos del protocolo de entrenamiento (A); y testeo (B).

Registro EEG. El electroencefalograma fue registrado mediante un equipo Akonic Bio-PC con una tasa de muestreo de 256 HZ, a partir de 19 electrodos montados de estaño (Sistema Internacional 10/20, referencia biauricular, ElectroCap Inc.). La resistencia de los electrodos fue mantenida por debajo de 10 kΩ durante todos los registros. El análisis y procesamiento de los registros fue realizado mediante el

paquete de análisis de datos electroencefalográfico EEGLAB (Delorme & Makeig 2004).

Se analizaron solamente los registros de los individuos que superaron el criterio de aprendizaje establecido *a priori* (ver más abajo). Cada registro fue filtrado de 1 a 20 Hz y re-referenciado a la referencia promedio de los 19 electrodos. El registro fue luego separado en épocas por condición experimental (Correcta/Incorrecta). Cada época abarcó 200 ms (línea de base) y 1800 ms posteriores a la aparición del “verbo” de cada frase. Para el posterior análisis se incluyeron solamente los ensayos en los cuales los sujetos respondieron correctamente (aciertos). Las épocas fueron procesadas mediante Análisis de Componentes Independientes (ICA) con el objeto de identificar y eliminar los componentes de artificios oculares (parpadeos y movimientos) que contaminaran los registros. Las épocas que poseían artificios de origen no ocular fueron eliminadas del análisis subsecuente. Dichas épocas correspondieron a aproximadamente 6% del total registrado.

#### *Análisis de datos*

Para establecer cuáles participantes respondieron acertadamente por encima del nivel azaroso (50%), mostrando así haber aprendido los términos “verbales” del lenguaje artificial, se realizó un test de G de bondad de ajuste (Zar 1999) con el número de aciertos y desaciertos por individuo. Las respuestas de los individuos que no respondieron acertadamente por encima del 50% fueron descartadas para los análisis comportamentales posteriores y para el estudio del registro EEG.

Para el resto de los participantes se discriminó por tipo de frase (Correcta vs Incorrecta) el porcentaje de respuestas, el porcentaje de aciertos y el tiempo de respuesta de los aciertos. Los datos individuales fueron ingresados en un análisis poblacional realizándose un test de t pareado (de muestras dependientes) para cada una de las tres variables.

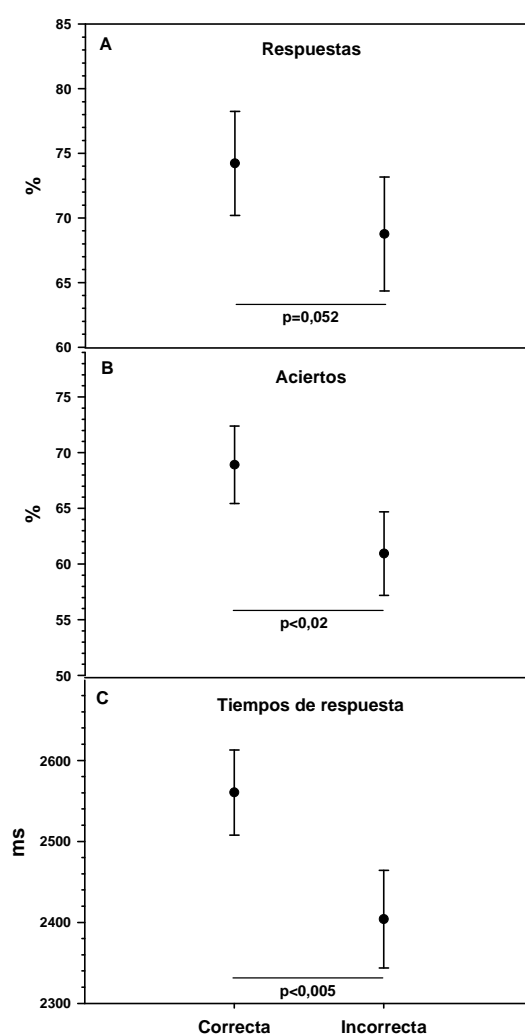
Para el análisis de los potenciales evocados se escogió la ventana de interés mediante inspección visual del Gran Promedio por condición experimental (Correcta vs Incorrecta). Por sujeto se calculó luego el promedio de dicha ventana para cada condición y electrodo, realizándose un ANOVA de medidas repetidas con factor intrasujeto Electrodo (Fp1 a Pz) y Condición (Correcta vs Incorrecta). El ajuste de Greenhouse-Geisser fue aplicado para corregir las violaciones a la esfericidad, y se utilizó el Test de Fisher para las comparaciones *post-hoc*.

### **3. Resultados**

#### *Comportamentales*

16 de los 24 participantes respondieron significativamente por encima del azar durante la fase de testeo (test G de bondad de ajuste,  $p < 0,05$ ), mostrando que la mayoría de los individuos fueron capaces de aprender el significado de los 6 verbos

presentados durante el entrenamiento. El porcentaje de aciertos para el grupo de 16 participantes que aprendió fue de  $86,41 \pm 2,49\%$  sobre el total respondido, siendo el promedio de respuestas efectuadas de un  $75,47 \pm 3,61\%$ . El porcentaje de respuestas efectuadas por el grupo de 16 participantes que aprendió fue mayor para la condición Correcta que para la Incorrecta (Figura 2A), siendo esta diferencia marginalmente significativa ( $t_{15} = -2,283$ ;  $p = 0,052$ ). El porcentaje de aciertos (Figura 2B) fue significativamente mayor para la condición Correcta comparada con la Incorrecta ( $t_{15} = -2,630$ ;  $p < 0,02$ ). Respecto al tiempo de respuesta (Figura 2C) fue significativamente diferente entre ambos tipos de frases ( $t_{15} = -3,764$ ;  $p < 0,005$ ) resultando mayores tiempos de respuesta en la condición Correcta vs. Incorrecta.



**Figura 2.** Porcentaje de respuestas (A), aciertos (B) y Tiempos de respuesta para la condición Correcta e Incorrecta

### **EEG**

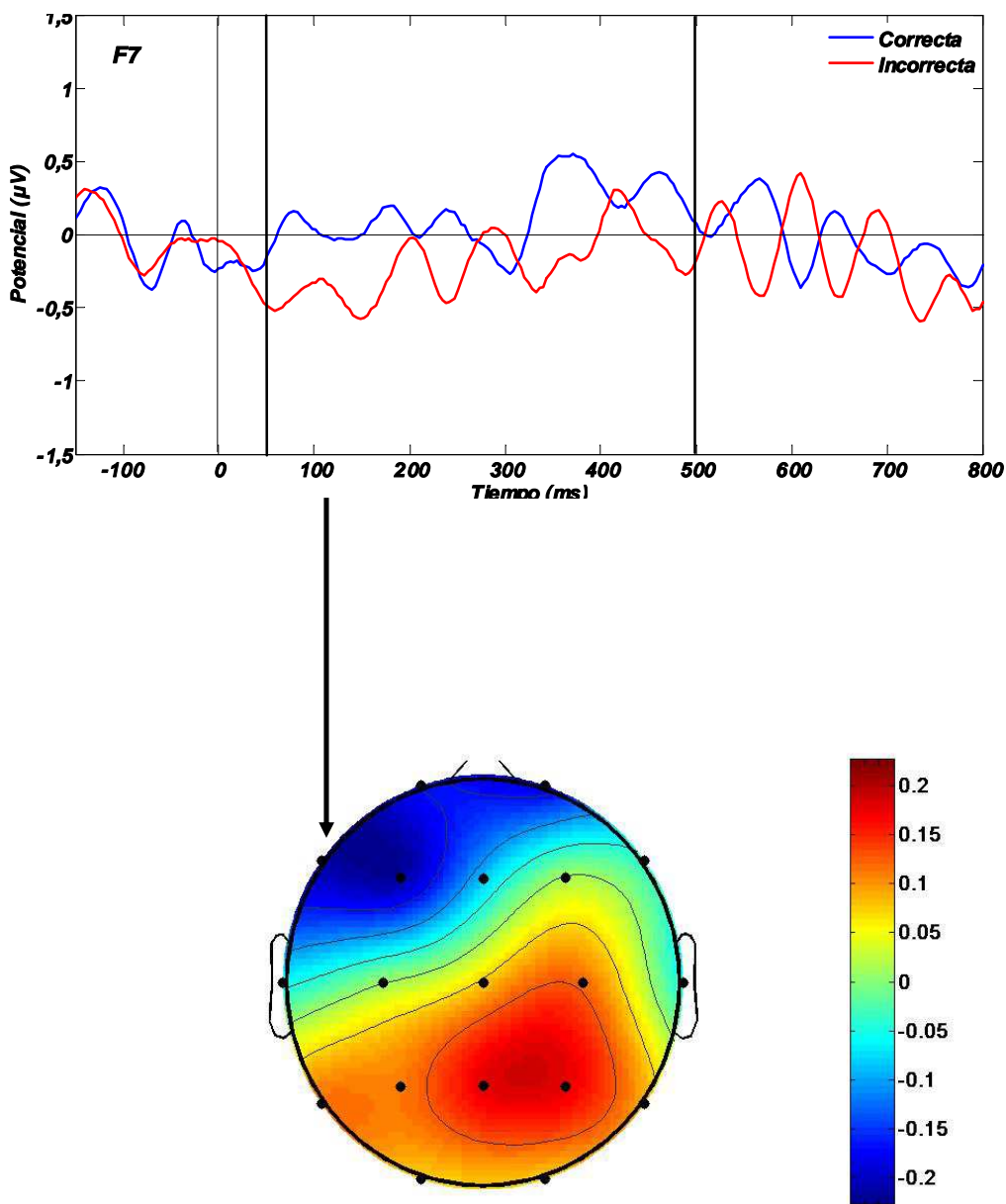
La inspección visual del Gran Promedio, considerando solo los ensayos acertados, mostró diferencias entre condiciones entre los 50-500 ms post-“verbo” principalmente en los trazados de los canales frontales (Figura 3, panel superior). El ANOVA de medidas repetidas para dicha ventana mostró una interacción significativa entre Electrodo y Condición ( $F_{18,252} = 3,365, p < 0,05$ ), viéndose que la falta de correlación entre el movimiento observado y el verbo presentado en la condición Incorrecta, generaba (respecto de la condición Correcta) una negatividad frontal lateralizada a la izquierda y una positividad centroparietal lateralizada a la derecha. Las comparaciones *post-hoc* revelaron diferencias distribuidas en los electrodos F7, T3 y los ubicados en la zona central ( $p < 0,05$ ). Por latencia y topografía, el potencial podía equipararse a un componente de tipo ELAN (Figura 3, panel inferior).

## **4. Discusión**

### **Comportamentales**

En el presente trabajo analizamos si se conservaba en adultos la capacidad de adquirir el significado de nuevos términos “verbales” mediante aprendizaje inferencial al presentar la información lingüística relevante en formato exclusivamente auditivo. En otras palabras, nos interesaba estudiar si los sujetos serían capaces de adquirir la estadística de co-ocurrencia entre términos y escenas en el caso de no poder leer los términos nuevos. Para ello, realizamos una modificación del protocolo experimental utilizado en trabajos anteriores (Fernández Feijóo et al. 2008; Lopes da Cunha et al. 2010), presentando las frases del lenguaje artificial sólo en formato auditivo. De esta manera intentamos presentar una situación experimental que tuviera mayor similitud con la situación de adquisición del lenguaje materno durante la infancia.

Los resultados mostraron que aún bajo este nuevo formato de presentación de los estímulos, la mayoría de los participantes (67%) fueron capaces de inferir el significado de los términos verbales presentados. Para poder realizar dicha inferencia, asociando cada movimiento con un término particular, los sujetos debieron seguir la estadística de co-ocurrencia de los movimientos y términos presentados entre ensayos.



**Figura 3.** Panel superior: Gran Promedio para las condiciones Correcta e Incorrecta para el canal F7. Las líneas punteadas indican la ventana temporal analizada (50 a 500 ms). Panel inferior. Topografía de la onda de diferencia (Incorrecta-Correcta) en la ventana temporal de 50 a 500 ms posteriores a la aparición del término “verbal”. Escala en µV.

Sin embargo, la comparación de los resultados del presente trabajo con los obtenidos en Lopes da Cunha et al (2010), muestran que el formato de presentación influyó en el porcentaje de individuos que fueron capaces de asociar cada término con su respectivo movimiento (67% vs. 79% respectivamente). Por tanto, la presentación exclusivamente auditiva de los términos dificultó el aprendizaje en comparación con la presentación audiovisual. La influencia del formato de presentación en el nivel de aprendizaje logrado ha sido previamente descrito. Se ha observado que la inducción de memorias lingüísticas falsas disminuye con la presentación visual de los términos en lugar de su presentación auditiva (Cleary & Greene 2002; Gallo et al. 2001; Kellogg 2001; Pierce et al. 2005; Smith et al., 2008). Mientras la presentación auditiva del estímulo lingüístico activa necesariamente sólo la información fonológica del término, la presentación audiovisual activa tanto la información fonológica como la ortográfica de dicho término. Esta información (ortográfica) adicional podría facilitar el reconocimiento del término presentado, contribuyendo a disminuir la generación de memorias falsas (Gallo et al., 2001; Kellogg 2001).

#### **Resultados por tipo de frase (Correcta vs Incorrecta)**

El análisis de los porcentajes de respuesta por condición mostró una disminución marginalmente significativa de las respuestas efectuadas para los ensayos con frase Incorrecta vs. Correcta. El menor número de respuestas obtenidas en la condición Incorrecta indica que los participantes excedieron el tiempo límite para responder un mayor número de veces bajo esta condición que bajo la condición Correcta. Puesto que en cada ensayo la aparición de la escena precedía a la aparición del lexema “verbal” que denotaba el movimiento mostrado (ca.  $1350 \pm 30$  ms), es posible que la visualización de dicho movimiento llevara a la activación de uno o más candidatos lexicales de manera previa a la aparición de la forma lexical *per se*. La coincidencia entre el lexema presentado y alguna de las formas 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. En este último caso, es posible que la falta de coincidencia entre lo presentado y lo esperado dispare procesos de recomparación entre la/s forma/s lexical/es activada/s, y el lexema efectivamente presentado en la frase, aumentando de esta manera el tiempo de decisión durante los ensayos con frase Incorrecta.

Con el objetivo de analizar la posible influencia del formato de presentación del estímulo lingüístico en esta variable, se comparó el presente resultado con el obtenido en Lopes da Cunha et al. (2010). Para ello, se ingresaron los valores de porcentaje de respuesta en un ANOVA de medidas repetidas con el factor intrasujeto “Condición” (Correcta e Incorrecta) e intersujeto “Formato” (Audiovisual, Auditivo). La interacción Condición x Formato no resultó significativa ( $F_{1,29} = 0,94$ , NS), sin embargo, el análisis mostró un efecto principal de “Condición” ( $F_{1,29} = 11,04$ ,  $p < 0,005$ ), por ser los ensayos con frase Incorrecta respondidos menos veces que en la condición Correcta, independientemente del formato de presentación de la frase. A su vez, si bien no hubo un efecto principal de “Formato” ( $F_{1,29} = 0,96$ , NS), se observó

que con el formato de presentación Auditivo los porcentajes de respuesta descendieron respecto del Audiovisual. Este patrón se dio tanto para la condición Correcta (Auditivo: 74,12%; Audiovisual: 83,00%) como en la condición Incorrecta (Auditivo: 68,75%; Audiovisual: 73,00%). Este resultado indica que la presentación exclusivamente auditiva de las frases favoreció que los participantes excedieran el tiempo límite para responder un mayor número de veces. Esto podría estar indicando una mayor incertidumbre al momento de responder, quizás como consecuencia de un menor nivel de aprendizaje logrado bajo este formato de presentación respecto del audiovisual.

El análisis del porcentaje de aciertos por condición mostró un descenso significativo en la condición Incorrecta vs. Correcta. Esto indica que la probabilidad de que los participantes decidieran que una frase incorrecta era correcta fue mayor a la probabilidad de que decidieran que una frase correcta era incorrecta. Dado que el grado de aprendizaje no fue perfecto (porcentaje de aciertos menor al 100%), es posible que la aparición del movimiento en la escena activara más de un candidato lexical con una alta probabilidad de que uno de ellos fuera el correcto y uno o más fueran incorrectos. Durante la presentación de una frase correcta, habría entonces una alta probabilidad de coincidencia entre uno de los candidatos activados y el presentado, siendo por tanto alta la probabilidad de que el individuo responda acertadamente. 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 podría decidir que la frase es correcta aún cuando no lo fuera. De esta manera, existiría una mayor probabilidad de dar como correcta una frase incorrecta que viceversa. El mismo patrón se observó en Lopes da Cunha et al. (2010), por lo que al igual que con el porcentaje de respuestas, se analizó la posible influencia del formato de presentación de las frases en esta variable, ingresando los resultados obtenidos en el presente trabajo y en Lopes da Cunha et al (2010) en un ANOVA de medidas repetidas con factor intrasujeto "Condición" e intersujeto "Formato". Nuevamente se halló un efecto principal de "Condición" ( $F_{1,29} = 20,39$ ,  $p < 0,0001$ ), puesto que los ensayos con frase Incorrecta fueron respondidos menos acertadamente que con frase Correcta, independientemente del formato de presentación de la frase. A su vez, si bien no hubo un efecto principal de "Formato" ( $F_{1,29} = 0,23$ , NS), se observó que bajo el formato Auditivo los aciertos fueron menores en la condición Correcta (Auditivo: 68,91%; Audiovisual: 75,83%), pero no en la Incorrecta (Auditivo: 60,93%; Audiovisual: 59,67%). Este patrón de resultados podría estar relacionado con el grado de dificultad para lograr el aprendizaje de los términos verbales. Así, al aumentar la dificultad de la tarea de aprendizaje por las características de presentación de la información lingüística (formato Auditivo), disminuiría el grado de asociación entre cada término verbal y el movimiento denotado por el mismo. En el caso de la condición Correcta, si la aparición de una escena activa más de un candidato lexical, el menor grado de asociación logrado por el lexema correcto con el movimiento disminuiría la probabilidad de responder que la frase es efectivamente Correcta, disminuyendo así el

número de aciertos para esta condición.

Al igual que en Lopes da Cunha et al. (2010), los tiempos de respuesta por condición experimental mostraron un descenso significativo para la condición Incorrecta respecto de la Correcta. Dado que a los participantes no se les instruyó que los errores de las frases estarían en el verbo, este resultado podría estar relacionado con que durante la presentación de una frase Correcta, el sujeto debía esperar necesariamente hasta la compleción de la frase para 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 una frase Correcta. Al analizar los resultados del presente trabajo y de Lopes da Cunha et al. (2010), en función tanto de la condición experimental como del formato de presentación de la frase, se halló un efecto principal de "Condición" ( $F_{1,29} = 12,79$ ,  $p < 0,005$ ), ya que en ambos formatos de presentación de frase el tiempo de respuesta fue mayor en la condición Correcta. A su vez se halló un efecto marginalmente significativo de "Formato" ( $F_{1,29} = 3,79$ ,  $p = 0,06$ ) debido a que los tiempos de respuesta fueron mayores cuando las frases fueron presentadas de manera auditiva tanto en frases de tipo Correcta (Auditivo: 2560,36 ms; Audiovisual: 2347,46 ms), como ante frases de tipo Incorrecta (Auditivo: 2403,85 ms; Audiovisual: 2294,51 ms). Este análisis parecería mostrar que la presentación audiovisual de la frase facilitó en cierta medida la respuesta de los participantes en comparación con la presentación en formato auditivo. Como fue mencionado anteriormente, es posible que la información visual sumada a la auditiva facilite tanto el aprendizaje del término presentado (Gallo et al., 2001; Kellogg 2001), como su reconocimiento al momento del testeo, lo cual podría contribuir a disminuir el tiempo de decisión de los individuos. Una disminución en el tiempo de decisión de los individuos debería también verse reflejado en un aumento en el porcentaje de respuestas, puesto que cuanto menor el tiempo de decisión del participante, mayor la probabilidad de responder dentro del tiempo límite permitido. En este sentido, si bien el porcentaje de respuestas obtenidas entre ambos formatos de presentación no difirió significativamente entre sí, el patrón observado concuerda con lo esperado. Así bajo el formato audiovisual se obtuvieron menores tiempos y mayores porcentajes de respuesta mientras que bajo el formato auditivo los tiempos aumentaron y los porcentajes de respuesta disminuyeron.

### **EEG**

En el presente trabajo, se halló una negatividad frontal izquierda en la ventana de 50 a 500 ms luego de la presentación de un término verbal incorrecto. La distribución y latencia del componente hallado son compatibles con los descriptos para el componente ELAN (Friederici et al. 1993; Friederici 1995; Friederici et al. 1996; Hahne & Friederici 1999; Isel et al. 2007; Lau et al. 2006; Neville et al. 1991; Rossi et al. 2005), aunque su duración total (450 ms) fue mayor a la habitualmente hallada. Este componente ha sido tradicionalmente asociado a la aparición de una incongruencia de clase gramatical entre el término esperado y el observado. Así, la



aparición de una preposición en una posición inesperada dentro de la oración genera un ELAN con un máximo hacia los 125 ms posteriores a su presentación (Neville et al. 1991). De manera similar, estudios hechos en alemán muestran que la aparición de un participio en una posición no esperada genera una negatividad frontal izquierda temprana. Este tipo de resultados ha llevado a postular la hipótesis de que este componente señala la aparición de un término de clase gramatical distinta a la esperada (Friederici 2002).

Estudios más recientes han hallado que el potencial ELAN se encuentra también modulado por el grado de expectativa de un término generada por el contexto oracional previo (Lau et al. 2006). Por otra parte, se ha visto que el ELAN puede ser elicitado no sólo por términos de una clase gramatical inesperada, sino también por estructuras de frase inesperadas (Bradley & Hestvik 2010). En este caso, la aparición de la negatividad anterior izquierda depende de que la estructura inesperada comience con una palabra claramente indicativa de su tipo oracional. Por ejemplo, la aparición de una frase nominal inesperada generará un ELAN sólo en el caso de que comience con un artículo, puesto que un artículo señala inequívocamente el comienzo de una frase nominal. Este tipo de resultado sumado a la ventana temporal temprana en la que se observa el potencial (típica de los procesamientos sensoriales), han llevado a postular la Hipótesis Sensorial del ELAN (Dikker et al., 2009). De acuerdo a esta hipótesis el componente ELAN es un componente sensorial que se elicitaba ante la aparición de una incongruencia entre la forma esperada y la observada, sin involucrar ningún tipo de procesamiento sintáctico temprano por parte del individuo.

En tareas de procesamiento de oraciones en lenguaje natural, experimentos realizados con magnetoencefalografía han hallado que ciertos componentes sensoriales tempranos como el M100 son modulados por las características formales de los términos presentados (Dikker et al. 2010). En el trabajo de Dikker et al. (2010), la aparición de un término sustantival inesperado generaba siempre un M100, pero cuando ciertas características formales se relacionaban con una alta probabilidad con la categoría sustantival inesperada, el componente sensorial aumentaba aún más en amplitud.

Dados los antecedentes descriptos, creemos que en el presente trabajo la aparición de un ELAN en los ensayos con frase Incorrecta, podría estar basado en la detección de una incongruencia formal entre el término esperado y el observado. Por ejemplo, si el movimiento observado en la escena de testeó presentada se correspondía con “jazu” (movimiento circular), la aparición de otro término inesperado (i.e. incorrecto) generaría un potencial de tipo ELAN por la falta de congruencia formal entre ambos términos.

Cabe mencionar que en el presente trabajo esperábamos que en los ensayos de tipo Incorrecta, la incongruencia entre el movimiento de la escena y el término “verbal elicitara un potencial de tipo N400, de manera similar a lo obtenido en Lopes da Cunha et al. (2010). La ausencia de este componente en el presente trabajo podría quizás deberse a un enmascaramiento por efecto del fuerte componente de tipo ELAN elicitado ante la aparición de los términos verbales incorrectos en una ventana

temporal similar a la del N400. Hipotetizamos que ante la ausencia de la información lingüística visual, facilitadora del aprendizaje, una estrategia compensadora podría haber sido prestar particular atención a las características formales auditivas de los términos presentados, para así poder realizar la tarea de aprendizaje satisfactoriamente. Un alto grado de atención al aspecto formal de los términos presentados, podría haber aumentado el efecto de generación del componente ELAN, y disminuido el del N400.

## 5. Conclusiones

En el presente trabajo se observó que sujetos adultos fueron capaces de adquirir la estadística de co-ocurrencia entre términos nuevos y movimientos de las escenas aún cuando la información lingüística era presentada sólo en formato auditivo. De esta manera intentamos presentar una situación experimental que tuviera mayor similitud con la situación de adquisición del lenguaje materno durante la infancia. La comparación del patrón de respuestas obtenidas con las halladas cuando la información lingüística fue presentada de manera audiovisual, mostró que la ausencia de la información ortográfica dificultó el proceso de asociación multimodal entre la escena (visual) y el término (auditivo). La generación de un potencial de tipo ELAN ante la aparición de una frase Incorrecta, podría estar relacionado con la detección de una incongruencia formal entre el término esperado y el observado. La ausencia de un componente de tipo N400 podría deberse a una estrategia compensadora por parte de los participantes, que ante la ausencia de la información ortográfica de los términos presentados, prestaran particular atención a las características formales auditivas de los términos, aumentando así el efecto de generación del componente ELAN, y disminuyendo el del N400.

## Bibliografía

- Aslin, R., Saffran J. & Newport, E. (1998). Computation of conditional probability statistics by 8-months-old infants. *Psychological Science*, 9(4), 321-324.
- Bradley, E.D. & Hestvik A. (2010) Testing the sensory hypothesis of the early left anterior negativity with auditory stimuli. *LSA Annual Meeting*, Baltimore, Maryland
- Cleary, A.M., & Greene, R. L. (2002). Paradoxical effects of presentation modality on false memory. *Memory*, 10, 55– 61.
- D'Odorico, L., & Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language*, 34, 891-907.
- Delorme, A. & Makeig S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134, 9-21

- Dikker, S., Rabagliati, H., & Pykkänen, L. (2009). *Sensitivity to syntax in visual cortex. Cognition*, 110(3), 293-321.
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pykkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, 21(5), 629-634.
- Fernández-Feijóo, M. E.; Vernis, S.; Zanutto, S. & Wainseboim, A. 2008. Adquisición de significados verbales nuevos por un proceso de asociación multimodal. *Revista Argentina de Neuropsicología*, 12, 1-12.
- Friederici, A.D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, 50, 259-281.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78-84.
- Friederici, A. D., Hahne, A. & Mecklinger, A. (1996). The temporal structure of syntactic parsing: Early and late effects elicited by syntactic anomalies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1219-1248.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1(3), 183-192.
- Gallo, D. A., McDermott, K. B., Percer, J. M. & Roediger, H. L. (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 339-353.
- 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. & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109-135.
- Hahne, A. & Friederici, A.D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, 13, 339-356.
- Hahne, A. & Friederici, A.D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11(2), 194-205.
- Holcomb, P. J. & Neville, H. J. (1991). Natural speech processing: An analysis using event related brain potentials. *Psychobiology*, 19, 286-300.
- Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological Psychology*, 74, 337-346 .
- Kellogg, R.T. (2001). Presentation modality and mode of recall in verbal false memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 913-919.

- Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, 207, 203-205.
- Kutas, M. & Hillyard, S., (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161-163.
- Lacerda, F. & 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. & 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*, 147-148. Genova.
- Lacerda, F., Marklund, E., Lagerkvist, L., Gustavsson, L., Klintfors, E. & 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*, 149-150. Genova.
- Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98, 74-88.
- Lopes da Cunha, P., Vernis, S., Zanutto, S., Wainelboim, A. (2010). Aprendizaje inferencial de términos verbales nuevos: un estudio de Potenciales Evocados. *Revista Argentina de Neuropsicología*, 16, 23-40.
- Marcus, G., Vijayan, S., Byi Rao, S. & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77-80.
- Marklund, E. & Lacerda, F. (2006). Infants' ability to extract verbs from continuous speech. *Ninth International Conference on Spoken Language Processing*, 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. & Christiansen, M.H. (2005). The differential contribution of phonological and distributional cues in grammatical categorisation. *Cognition*, 96, 143-182.
- Neville, H.J. (1985). Brain potentials reflect meaning in language. *Trends in Cognitive Neuroscience*, 3, 91-92.
- Neville, H., Nicol, J.L., Barss, A., Forster, K.I. & Garrett, M.F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3(2), 151-165.
- Pierce, B. H., Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2005). The modality effect in false recognition: Evidence for test-based monitoring. *Memory and Cognition*, 33, 1407-1413.
- Rossi, S., Gugler, M.F., Hahne, A., & Friederici, A.D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, 384, 293-321.
- Saffran, J.R., Aslin, R.N. & Newport E.L. (1996). Statistical learning by eight-month-old

infants. *Science* 274, 1926–1928.

Siskind, J.M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61, 39–61.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558-1568.

Smith, R. E., Hunt, R. R., & Gallagher, M. P. (2008). The effect of study modality on false recognition. *Memory and Cognition*, 36, 1439–1449.

Yu, C. (2008). A Statistical Associative Account of Vocabulary Growth in Early Word Learning. *Language Learning and Development*, 4(1), 32–62.

Yu, C. & Ballard, D. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, 70, 2149–2165

Yu, C. & Smith, L.B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18(5), 414-420.

Zar, J.H. (1999). *Biostatistical analysis*. New Jersey: Prentice Hall.