

## Aprendizaje inferencial de términos verbales nuevos: un estudio de Potenciales Evocados

Pamela Lopes DaCunha<sup>1</sup>, Sergio Vernis<sup>1</sup>, Silvano Zanutto<sup>12</sup> y Alejandro Wainselboim<sup>1</sup>

<sup>1</sup> Laboratorio de Biología del Comportamiento, IByME-CONICET

<sup>2</sup> Instituto de Ingeniería Biomédica, Facultad de Ingeniería, UBA

### Resumen

*La capacidad de sujetos adultos para adquirir términos verbales nuevos mediante aprendizaje inferencial ha sido estudiada. Para ello se presentaron en computadora 70 escenas visuales diferentes consistentes en figuras geométricas que realizaban distintos posibles movimientos. De manera simultánea cada escena era descrita por una frase formada por no-palabras. Los participantes debían aprender cuáles eran los términos que denotaban cada uno de los movimientos presentados. En el testeo, 80 escenas con nuevas combinaciones de las figuras y movimientos del entrenamiento eran presentadas en simultáneo ya fuera con frases que las describían correctamente (condición "Correctas", 50% de ensayos) o que poseían un "verbo" que no describía el movimiento presentado en la escena (condición "Incorrectas", 0% de ensayos), mientras se registraba su actividad electroencefalográfica. Los participantes debían decidir en menos de 2 segundos, si la frase describía o no la escena presentada. 15 de los 19 participantes respondieron significativamente por encima del azar durante el testeo. El registro de los potenciales evocados mostró la aparición de una negatividad parieto-occipital en los 400 a 700 ms posteriores a la aparición del "verbo" incorrecto, de características compatibles con el N400. Los resultados muestran que en adultos es posible inferir el significado de términos verbales nuevos a partir de la co-ocurrencia entre contexto visual y lingüístico. La aparición de una negatividad de tipo N400 ante la incongruencia entre escena y "verbo", señala que la integración de la información contextual y lingüística se da de manera continua, apoyando la visión de una interpretación incremental de la frase.*

*Palabras clave:* aprendizaje inferencial - significados verbales - N400

---

Correspondencia con los autores: [awainselboim@ibyme.conicet.gov.ar](mailto:awainselboim@ibyme.conicet.gov.ar)

Artículo recibido: 29/09/2010

Artículo aceptado: 29/11/2010

<http://www.revneuropsi.com.ar>

ISSN: 1668-5415

## Abstract

*The capacity to acquire new verbal meanings by inferential learning was studied in adults. To do so, 70 different visual scenes were presented by computer. Each scene consisted of geometrical figures that could perform different possible movements. In simultaneous, each scene was described by a sentence formed by non-words. Participants had to learn which non-word referred to each presented 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 ("Correctas" condition, 50% of trials), or presented an erroneous "verb" ("Incorrectas" condition, 50% of trials), while the electroencephalographic activity of the participant was recorded. 15 of the 19 participants responded significantly above chance during the test. Evoked potential recordings showed the appearance of a parieto-occipital negativity 400 to 700 ms after presentation of the erroneous "verb", that was compatible with an N400. Results show that adults are able to infer new verbal meanings from the co-occurrence between visual and linguistic context. The appearance of an N400-like negativity with incongruence between scene and "verb" indicates that contextual and linguistic information are continuously integrated, supporting the vision of an incremental interpretation of sentences.*

*Key words:* inferential learning - verbal meaning - N400

## 1. Introducción

La adquisición del lenguaje durante la 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, ni la mayor parte del vocabulario adquirido. 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, 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 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). A su vez, el infante debe agrupar dichas unidades en las distintas categorías gramaticales que posee el idioma (sustantivo, verbo, pronombre, entre otros), 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 (por ejemplo: 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).

Por otra parte, 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 recientes llevados a cabo por Yu y Smith (2007) donde mostraron 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 sólo 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 y 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 y Sundberg, 2006; Yu 2008; Yu y 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 y Sundberg, 2006; Yu y 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 y 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 dónde la adquisición de sustantivos precede a la de verbos como clase gramatical (Gentner, 1978, 1982; D'Odorico y 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 fase posterior 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.

En el presente trabajo nos planteamos continuar con la línea anterior de trabajo, estudiando la neurofisiología del procesamiento de frases Correctas e Incorrectas mediante el registro de electroencefalografía (EEG) durante la fase de testeo. En particular, nos preguntamos si se generará una señal de potencial evocado (ERP) asociado a la aparición de una incongruencia entre el movimiento desplegado en la escena y el término verbal mostrado en simultáneo (incongruencia de tipo semántica). 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 y Hillyard, 1980, 1984; Neville, 1985; Holcomb y Neville, 1991; Hahne y Friederici, 2002), denominándose por ello N400. En el presente trabajo nos planteamos como hipótesis de trabajo que la aparición de una incongruencia entre verbo y movimiento de la escena generará un potencial de tipo N400.

## 2. Método

### Sujetos

19 adultos sanos de entre 20 y 33 años de edad (media  $27,2 \pm 0,7$  años), participaron voluntariamente en el experimento. Todos los sujetos eran diestros, hablantes del español como primera lengua y tenían estudios terciarios en curso o completos. Los sujetos no tenían antecedentes de desordenes neurológicos o psiquiátricos, ni se encontraban bajo medicación alguna al momento del experimento.

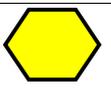
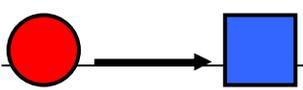
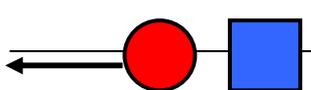
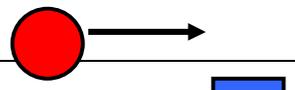
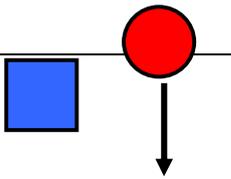
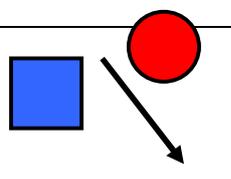
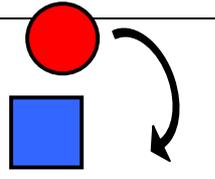
### Estímulos

Se realizó una modificación del protocolo de Fernández Feijóo et al. (2008) aumentándose el número de estímulos utilizados para estudiar en adultos la capacidad de asociar información visual de acciones con formas “verbales” nuevas y sus posibles correlatos electroencefalográficos. Para ello se crearon 150 escenas visuales diferentes (duración:  $3,34 \pm 0,01$  s), realizando todas las combinaciones posibles de 5 figuras geométricas y 6 movimientos distintos (Tabla 1). A su vez, se crearon 150 frases diferentes describiendo cada una de las distintas escenas creadas. 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, y el verbo refería al movimiento desplegado en la escena.

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

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

## Procedimiento

### *Entrenamiento*

Durante la realización del experimento cada participante se encontraba sentado en un cuarto aislado eléctrica y acústicamente, 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, y de manera simultánea, por debajo de cada escena, aparecía una frase que describía la escena mostrada (Fig. 1A). La frase aparecía en formato audiovisual palabra por palabra de izquierda a derecha, con un intervalo entre palabras de 350 ms. Cada palabra se mantenía en pantalla hasta el completamiento de la frase. En esta etapa se presentaron 70 ensayos diferentes, siendo cada ensayo presentado dos veces consecutivas. El intervalo entre ensayos era de 3000 ms.

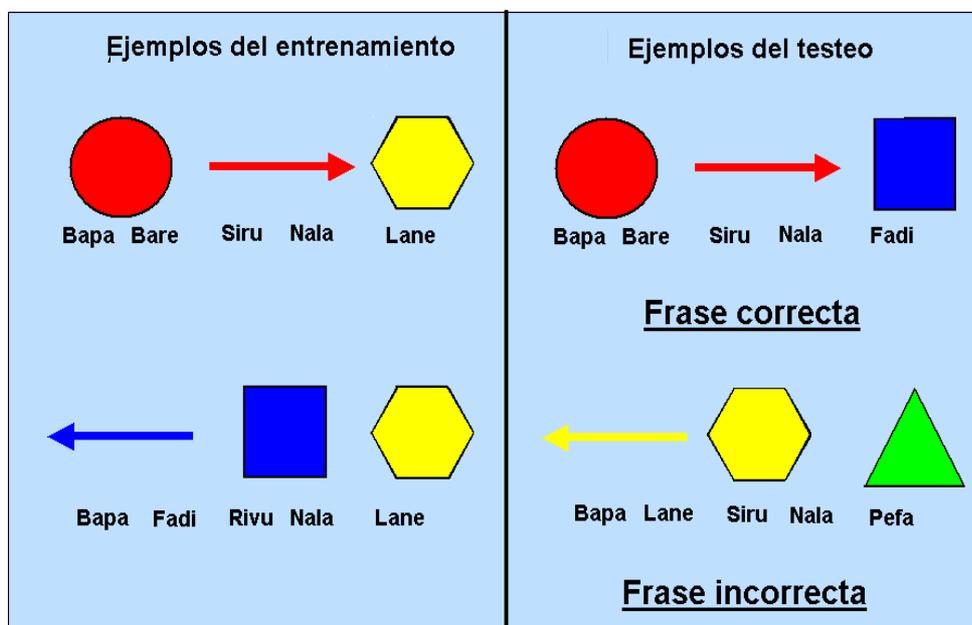
Para asegurar que los sujetos prestaran atención a los estímulos presentados, cada 4 a 7 frases se le preguntaba en formato visual si determinada palabra había aparecido en la frase anterior. Para responder debían apretar Ctrl Derecho para decir que "SI", y Ctrl Izquierdo para "NO". La duración total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

### *Testeo*

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, de manera similar al 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, Fig. 1B), mientras que en el restante 50% el "verbo" no correspondía con el movimiento mostrado en la escena visual (frase Incorrecta, Fig. 1B). Los sujetos tenían que decidir lo más rápidamente posible si la frase describía correctamente la escena o no, debiendo presionar en caso afirmativo Ctrl Derecho, y en caso negativo Ctrl Izquierdo. El tiempo límite para responder era 2000 ms luego de la aparición completa de cada frase. En total se presentaron las 80 escenas no vistas durante el entrenamiento, cada una de ellas una sola vez. Durante la fase de testeo se registró el electroencefalograma de los participantes.

Como medidas comportamentales se tomó de cada frase si el participante respondía o no, acertaba o no y su latencia de respuesta medido desde la aparición del "verbo", por ser éste el estímulo clave en la aparición del error en la frase.

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. Posteriormente, 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 (Correctas-Incorrectas). 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 incluyó solamente los ensayos en los cuales los sujetos respondieron correctamente (aciertos). Las épocas fueron examinadas visualmente y excluidas del análisis si contenían artificios oculares o de otro tipo. De esta manera se excluyeron alrededor del 5% de las épocas.

### Estadística

Con el objetivo de establecer qué participantes respondieron de manera acertada por encima del nivel azaroso (50%) 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 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.

### 3. Resultados

#### Comportamentales

Los resultados obtenidos muestran que 15 de los 19 participantes (ca. 79%) respondieron significativamente por encima del azar durante la fase de testeo (test de G de bondad de ajuste,  $p < 0.05$ ), con lo cual 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 15 participantes que aprendió fue de  $87,00 \pm 2,54\%$ , siendo el número promedio de respuestas efectuadas  $62,40 \pm 4,26$ .

El porcentaje de respuestas efectuadas fue mayor para las frases Correctas que para las Incorrectas (Fig. 2), siendo esta diferencia significativa ( $t_{14} = -2,454$ ;  $p < 0,003$ ). El porcentaje de aciertos (Fig. 3) fue significativamente mayor para las frases Correctas comparadas con las Incorrectas ( $t_{14} = -3,610$ ;  $p < 0,005$ ). Respecto al tiempo de respuesta (Fig. 4), no fue significativamente diferente entre ambos tipos de frases ( $t_{14} = -1,287$ ;  $p = 0.219$ ) aunque se observó una tendencia a mayores tiempos de respuesta en las frases Correctas comparadas con las Incorrectas.

Figura 2. Porcentaje de respuestas obtenidas para frases Correctas e Incorrectas

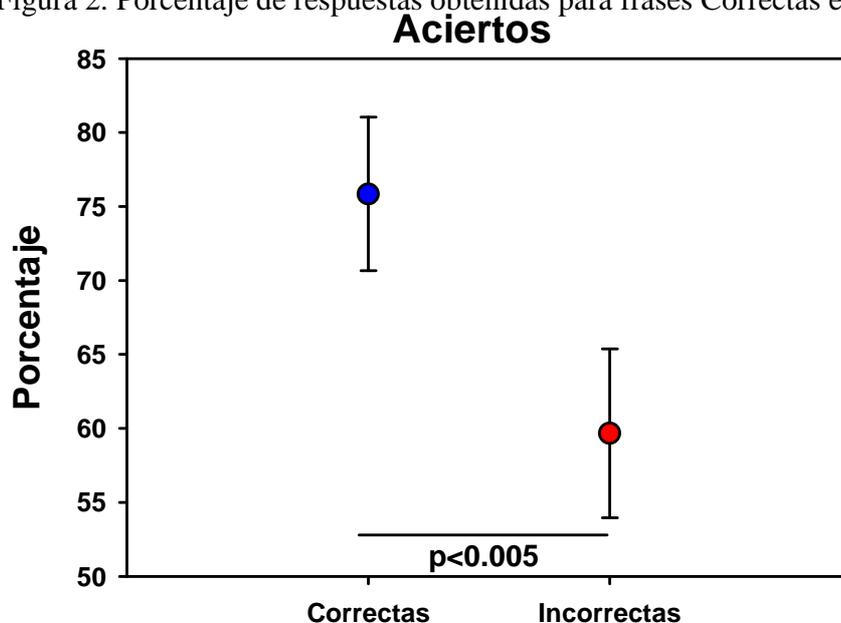


Figura 3. Porcentaje de aciertos ante frases Correctas e Incorrectas

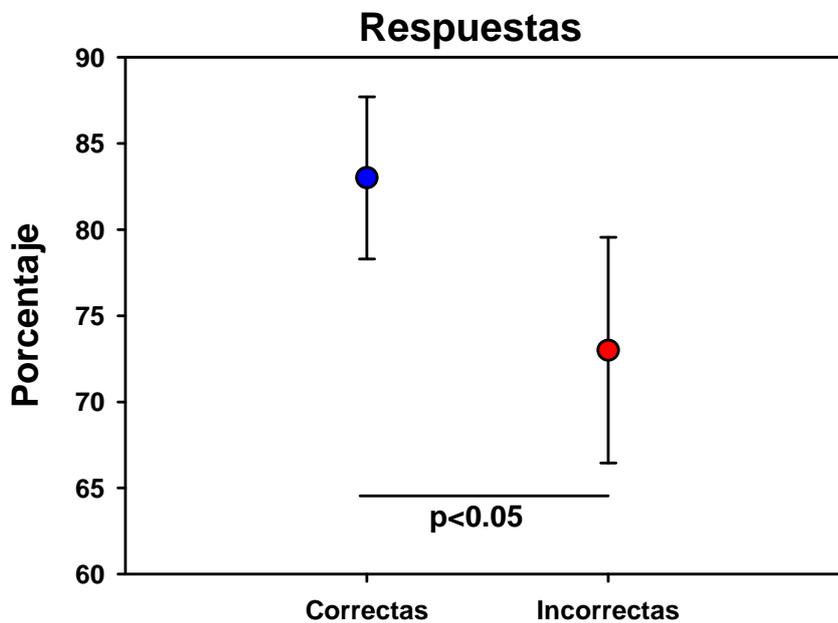
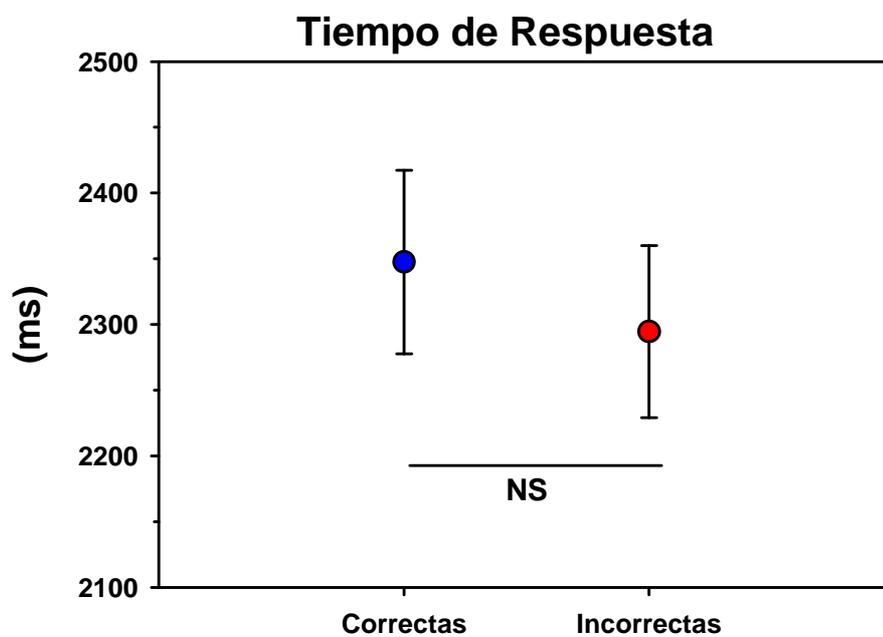


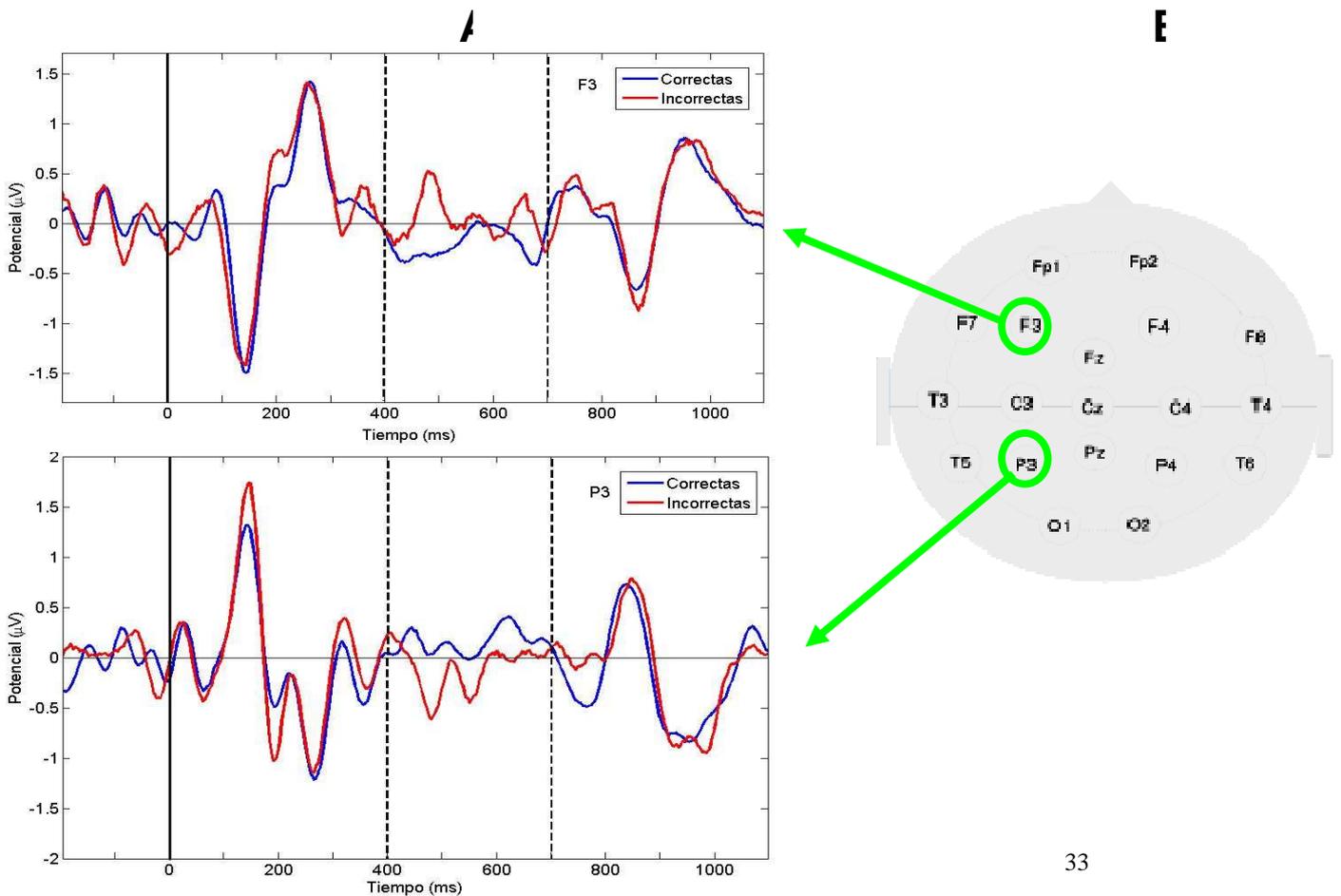
Figura 4. Tiempo de respuesta promedio de los aciertos ante frases Correctas e Incorrectas



**EEG**

La inspección visual del Gran Promedio, considerando sólo los ensayos acertados mostró diferencias entre la presentación de frases Correctas e Incorrectas en los trazados de los canales tanto frontales como parieto-occipitales aproximadamente entre los 400 ms y 700 ms luego de la presentación del “verbo” (Fig. 5. A). A partir de este resultado se definieron dos ventanas de interés para el análisis estadístico: 0-400 ms y 400-700 ms luego de la aparición del “verbo”. Para cada ventana temporal se calculó el voltaje promedio por condición y participante para cada uno de los 19 electrodos de registro utilizados (Fig. 5. B), los cuales fueron ingresados en un ANOVA de medidas repetidas con los factores Electrodo (Fp1 a Pz) y Condición (Correctas e Incorrectas). El ajuste de Greenhouse-Geisser fue aplicado para corregir las violaciones a la esfericidad.

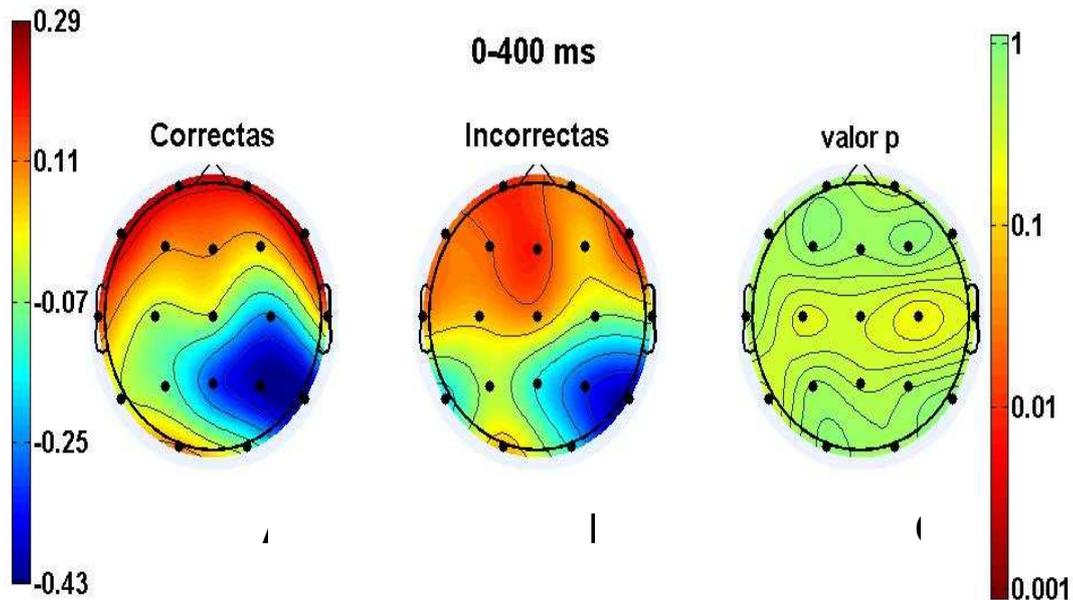
Figura 5. A. Gran Promedio para las condiciones Correctas e Incorrectas para F3 y P3  
B. Distribución topográfica de los 19 electrodos utilizados para el registro EEG



0-400 ms

El análisis de los datos de los primeros 400 ms posteriores a la aparición del verbo no mostró efectos principales de Condición ( $F_{1,14} = 0,141$ ,  $p = 0,713$ ), ni una interacción significativa entre Electrodo y Condición ( $F_{18,252} = 0,694$ ,  $p = 0,562$ ). Por tanto, no se hallaron diferencias de potencial entre condiciones experimentales para esta primera ventana temporal (Fig. 6 A y B). Con el objeto de visualizar los resultados estadísticos, se realizaron comparaciones post-hoc para cada electrodo entre condiciones (Fig. 6, derecha), aunque la ausencia de significancia en la interacción Electrodo y Condición no lo hiciera necesario.

Figura 6. Topografía del potencial promedio (en  $\mu V$ ) durante los primeros 400 ms posteriores a la aparición del verbo ante Frases Correctas e Incorrectas, y valores de p para cada electrodo en las comparaciones post-hoc (entre condiciones)

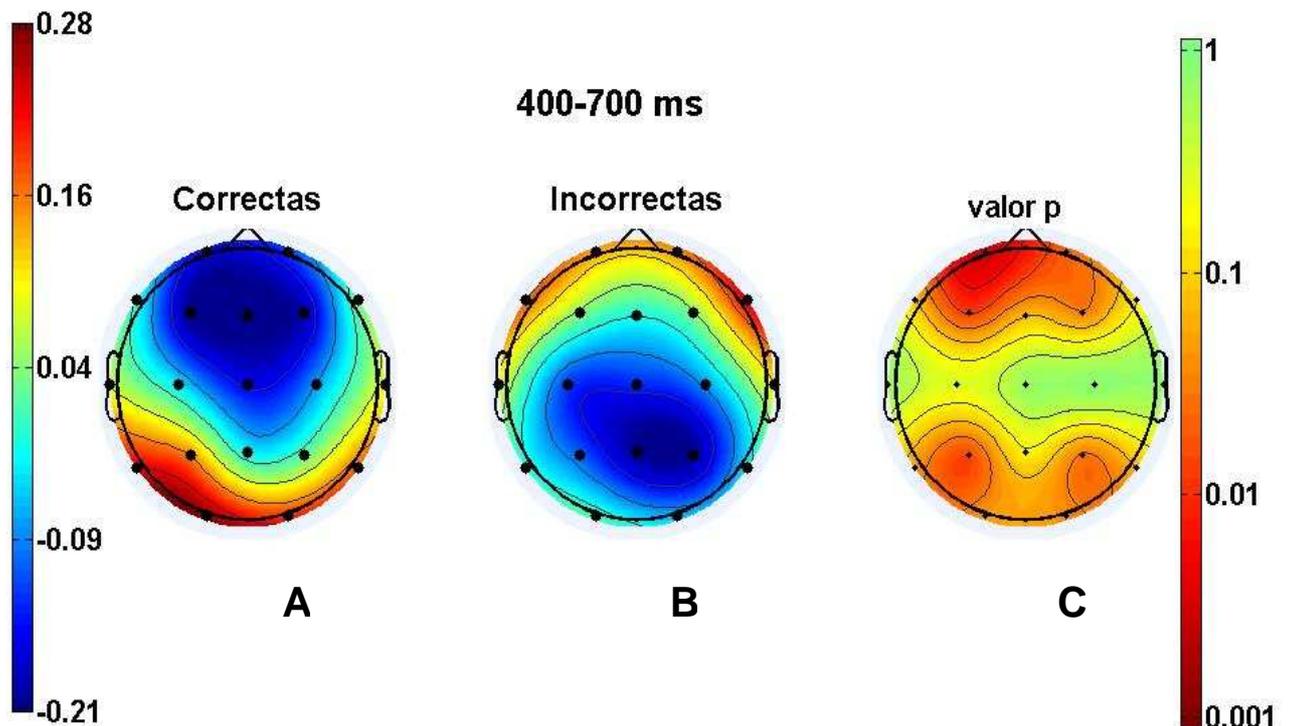


400-700 ms.

El ANOVA de medidas repetidas en la ventana de 400 a 700 ms posteriores a la aparición del “verbo” no mostró un efecto principal de Condición ( $F_{1, 14} = 0,011$ ;  $p = 0,917$ ), pero sí una interacción significativa entre Electrodo y Condición ( $F_{18, 252} = 5,016$ ;  $p < 0,005$ ). La distribución topográfica de cada condición experimental mostró que mientras en la condición Correctas la presentación de frases que concordaban con la escena presentada elicita una negatividad fronto-central (Fig. 7.A), la incongruencia entre el movimiento de la escena y el verbo presentado en simultáneo

elicitaba una negatividad centro-parietal característica de un componente de tipo N400 (Fig. 7.B). Las comparaciones post hoc revelaron diferencias distribuidas en los electrodos frontales y parieto-occipitales (Fig. 7.C).

Figura 7. Topografía del potencial promedio (en  $\mu V$ ) en la ventana temporal de 400 a 700 ms posteriores a la aparición del “verbo” en Frases Correctas e Incorrectas, y valores de p de las comparaciones post-hoc para cada electrodo (entre condiciones)



#### 4. Discusión

##### *Resultados Comportamentales*

Los resultados obtenidos en el presente trabajo replican y extienden los obtenidos por Fernández Feijóo et al. (2008). Desde el punto de vista comportamental, un porcentaje mayoritario de los participantes (15 de 19) fueron capaces de inferir el significado de los 6 términos verbales nuevos mediante el seguimiento de la co-ocurrencia entre la aparición de los distintos términos y la información contextual dada por las diferentes escenas mostradas durante el entrenamiento, aún en ausencia total de conocimiento previo del “idioma” presentado.

El patrón de respuestas obtenido durante el testeó mostró diferencias cuantitativas entre las dos condiciones experimentales del mismo tipo que las obtenidas en

Fernández Feijóo et al. (2008), aún cuando en el presente trabajo se duplicó el número de ensayos de testeo en comparación con el mencionado trabajo. Así, tanto el porcentaje de respuestas como el de aciertos fueron significativamente mayores para las frases Correctas comparado con las Incorrectas. En el primer caso, el menor número de respuestas obtenidas en la condición Incorrectas indica que los participantes excedieron el tiempo límite para responder un mayor número de veces bajo esta condición que en los ensayos de Correctas. Puesto que en cada ensayo la aparición de la escena precedía a la aparición del lexema “verbal” que denotaba el movimiento mostrado ( $1335 \pm 27$  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 de Incorrectas.

Por su parte, el porcentaje de aciertos fue significativamente mayor para la condición Correctas en comparación con Incorrectas. 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, 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 sea. De esta manera, existiría una mayor probabilidad de dar como correcta una frase incorrecta que viceversa.

El análisis de los tiempos de respuesta no arrojó diferencias significativas entre condiciones, al contrario de lo observado en Fernández Feijóo et al. (2008). Sin embargo, aun en ausencia de significancia estadística, el patrón observado fue el mismo que en el mencionado estudio, puesto que los participantes tardaron en promedio 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, este resultado 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. En

el presente trabajo los tiempos de reacción fueron mayores para ambas condiciones experimentales en comparación con Fernández Feijóo et al. (2008), habiendo disminuido por otra parte la diferencia entre condiciones experimentales (ca. 28 ms). Es posible que el aumento en los tiempos de reacción en ambos tipos de frase en el presente trabajo sea consecuencia de un aumento en el nivel de cansancio de los participantes debido a la mayor extensión y cantidad de ensayos (80 de testeo) del presente experimento en comparación con el protocolo de Fernández Feijóo et al. (2008) en donde se presentaron 40 ensayos de testeo. Un aumento del nivel de cansancio podría haber disminuido también las diferencias entre condiciones experimentales.

### *EEG*

Los resultados obtenidos mostraron que en los primeros 400 ms posteriores a la aparición del término verbal, no existieron diferencias entre condiciones experimentales en los potenciales registrados. Sin embargo, entre los 400 y los 700 ms se observó una clara diferenciación en los potenciales generados ante las dos condiciones presentadas en el testeo. Mientras que en la condición Correctas se obtuvo una negatividad fronto-central, en los ensayos de Incorrectas se generó una negatividad parieto-occipital de tipo N400.

La negatividad frontal observada en los ensayos de Correctas, podría estar relacionada con procesos atencionales dirigidos específicamente hacia el movimiento desplegado en la escena, disparados por la aparición del término verbal en la frase. La latencia y ubicación frontal del componente negativo encontrado es similar al componente ADAN (Attention Driven Anterior Negativity), el cual se elicitaba típicamente ante la aparición de un estímulo clave (“cue”) que señala la aparición posterior de un estímulo relevante para la tarea (Nobre et al., 2000; Hopf y Mangun, 2000; Eimer et al., 2002; Seiss et al., 2007). Si bien este componente ha sido descrito típicamente como de distribución contralateral a la posición en la que se espera la aparición del estímulo relevante, es posible que en el presente trabajo, dada la ubicación siempre central en la pantalla de la computadora del movimiento desplegado en cada ensayo, generara una distribución más bilateral del componente. Por otra parte, la generación de un componente de tipo N400 en el presente experimento ante la aparición del término que denotaba el movimiento en los ensayos de la condición Incorrectas, muestra que la incongruencia entre el movimiento mostrado en la escena y el “verbo” presentado en la frase genera un componente de origen cognitivo comparable al observado en protocolos psicolingüísticos clásicos.

La aparición de un potencial de tipo N400 en los ensayos de Incorrectas apoya, por su parte la visión de una interpretación incremental de la frase con una integración semántica temprana de la información contextual (Sedivy et al., 1999; Knoeferle et al., 2009). Knoeferle et al. (2009) realizaron un experimento con registro EEG en el cual presentaron por medio de una computadora 160 dibujos diferentes en el que cada uno representaba una acción particular (por ejemplo, una gimnasta aplaudiendo a una periodista). Luego de la exposición a cada imagen (por al menos 3s), ésta

desaparecía, y se presentaba palabra por palabra una frase escrita (en inglés) que describía la imagen vista previamente o que presentaba un error en el verbo con respecto a la acción vista en la imagen previa. Los resultados mostraron la generación de un componente de tipo N400 ante la aparición de un verbo incongruente con la acción presentada en la imagen previa. Los autores interpretaron sus resultados como indicativos de un proceso incremental de interpretación de la información semántica de la frase, que integra y verifica de manera continua la información contextual y lingüística.

A diferencia de Knoeferle et al. (2009), en dónde la información contextual (pictórica) de la acción precedía a la lingüística (verbo) en alrededor de 4600 ms, en el presente trabajo ambas fuentes de información (escena y frase) se presentaron de manera simultánea, precediendo la aparición del movimiento a la del “verbo” en alrededor de 1330 ms. Por tanto, la generación de un N400 ante la aparición de los lexemas verbales incongruentes con el movimiento de la escena en este caso indica que el proceso de integración y verificación de la información lingüística con la contextual es aún más rápido que lo pensable por medio de paradigmas de presentación de imagen y frase de manera no simultánea (Wassenaar y Hagoort, 2007; Knoeferle et al., 2009).

Los resultados obtenidos en el presente trabajo podrían complementarse a futuro con datos de tiempos de fijación visual sobre cada uno de los diferentes componentes de la escena durante la aparición secuencial de la frase. Sería esperable que la aparición de una incongruencia entre el movimiento de la escena y el lexema verbal aumentara el tiempo de fijación sobre ambos, en comparación con los ensayos con escenas y frases congruentes. Por otra parte, los resultados obtenidos hasta el momento muestran que en adultos es posible inferir el significado de términos verbales nuevos mediante el seguimiento explícito de la co-ocurrencia entre término y contexto visual. Nos planteamos como próximo paso estudiar si es posible la adquisición de términos verbales nuevos mediante aprendizaje de tipo implícito.

## 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.
- D’Odorico, L., y Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language*, 34, 891-907.
- Eimer, M, Van Velzen, J. y Driver, J. (2002). Crossmodal interactions between audition, touch and vision in endogenous spatial attention: ERP evidence on preparatory states and sensory modulations. *Journal of Cognitive Neuroscience*, 14, 254–271.
- Fernández Feijóo, M. E.; Vernis, S.; Zanutto, S.; Wainelboim, A. 2008. “Adquisición de significados verbales nuevos por un proceso de asociación multimodal”. *Revista Argentina de Neuropsicología*, 12, 1-12.

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.

Hahne, A. y Friederici, A.D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, 13, 339-356.

Holcomb, P. J. y Neville, H. J. (1991). Natural speech processing: An analysis using event related brain potentials. *Psychobiology*, 19, 286-300.

Hopf, J.M. y Mangun, G.R. (2000). Shifting visual attention in space: An electrophysiological analysis using high spatial resolution mapping. *Clinical Neurophysiology*, 111, 1241–1257.

Knoeferle, P., Urbach, T.P. y Kutas, M. (2009). Is incremental semantic interpretation related to end-of-sentence verification?: Evidence from correlation analyses. En N.A. Taatgen y H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society*, 1127-1132.

Kutas, M. y Hillyard, S., (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161–163.

Kutas, M., y Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, 207, 203-205.

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*, 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.

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. y 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.

Nobre, A.C., Sebestyen, G.N. y Miniussi, C. (2000). The dynamics of shifting visuospatial attention revealed by event related brain potentials. *Neuropsychologia*, 38, 964-974.

Saffran J.R., Aslin R.N., Newport E.L. (1996). Statistical learning by eight-month-old infants. *Science* 274, 1926-1928.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-148.

Seiss, E., Gherri, E., Eardley, A.F. y Eimer, M. (2007). Do ERP components triggered during attentional orienting represent supramodal attentional control? *Psychophysiology*, 44(6), 987-990.

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

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

Wassenaar, M., y Hagoort, P. (2007). Thematic role assignment in patients with Broca's aphasia: Sentence-picture matching electrified. *Neuropsychologia*, 45(4), 716-740.

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

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

Yu, C. y 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.

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