

Inferencias Espaciales en Comprensión de Texto: ¿Lingüísticas o Viso-Espaciales? Evidencia Cognitiva y de Neuroimagen

Débora I. Burín¹, Gastón Saux¹, Natalia Irrazabal¹, P. Lameiro², D. Saferstein², L. Ponto³ y R. Jorge³

¹ Facultad de Psicología, Universidad de Buenos Aires – CONICET

² Instituto Médico de Alta Tecnología, Buenos Aires

³ University of Iowa Hospital and Clinics, Iowa

Resumen

La comprensión de texto consiste en la construcción de una representación integrada del mismo. Las teorías clásicas suponen que ésta implica un formato verbal o proposicional. Una hipótesis más reciente asume que lo que se construye es una simulación, una representación multimodal, análoga a la experiencia de estar en la situación descrita por el texto. El objetivo del presente trabajo ha sido comparar las hipótesis verbal / proposicional y la hipótesis de simulación, tomando datos conductuales y de neuroimagen mediante SPECT. Doce adultos jóvenes leyeron narraciones que requerían inferencias espaciales, bajo el paradigma de congruencia / incongruencia, en dos condiciones: lectura para comprender las historias (Comprender), o para formarse una imagen mental viso-espacial (Imaginería). Se administró el radiotrazador ^{99m}Tc-ECD, en dos sesiones de 16 historias por cada condición, tomando los tiempos de lectura por oración; posteriormente se adquirieron neuroimágenes en cámara gamma. El tiempo de lectura para las oraciones que contenían una incongruencia espacial fue significativamente mayor que para las congruentes, lo cual supone que los lectores entendían la dimensión espacial del texto. La comparación de la activación cerebral Imaginería – Comprender mostró un cantidad significativamente mayor de voxels activos en zonas posteriores (giro lingual y fusiforme derecho, cuneus izquierdo), en tanto que la comparación Comprender – Imaginería activó en mayor medida zonas fronto-temporales. Esto sugiere que la comprensión de texto es una actividad principalmente lingüística (hipótesis verbal / proposicional); y que la actividad de zonas viso-espaciales del cerebro parece asociarse a la exigencia de formarse imágenes mentales.

Palabras clave: comprensión de texto – inferencias espaciales – SPECT – neuroimagen

Correspondencia con los autores: dburin@psi.uba.ar

Artículo recibido: 31 de marzo de 2011

Artículo aceptado: 15 de junio de 2011

Abstract

Text comprehension involves constructing an integrated representation of the ideas conveyed by it. Classical psycholinguistic theories assume that this construction takes on a verbal - propositional format. A recent hypothesis assumes that what is built is a simulation, a multimodal representation, analogous to the experience of being in the situation described in the text. The objective of this project was to compare the verbal / propositional and simulation hypotheses, employing convergent behavioral and neuroimaging SPECT data. Twelve young adults read stories that required inferences about space, under the consistency / inconsistency paradigm, in two conditions: reading stories to understand (Understand), or to form a mental visuo-spatial image (Imagery). The ^{99m}Tc-ECD radiotracer was administered in two sessions, in which 16 stories by each condition were read, measuring the reading time per sentence; thereafter neuroimages were acquired into the gamma camera. Reading times for sentences that contained a spatial inconsistency were significantly higher than for the consistent version, which implies that readers understood the spatial dimension of the text. Comparison of brain activation Imagery - Understand showed significantly more active voxels in posterior areas (right lingual and fusiform gyrus, left cuneus), while the comparison Understand - Imagery showed differences in fronto-temporal areas. These results suggest that reading comprehension is primarily a linguistic activity (verbal / propositional hypothesis); activity in visuo-spatial areas of the brain seems to be associated with the requirement to form mental images.

Key words: Text comprehension – spatial inferences – SPECT – neuroimage

1. Introducción

La comprensión de texto es una tarea cognitiva compleja que involucra procesos lingüísticos de bajo nivel, como el procesamiento de grafemas, palabras y oraciones. Pero también implica procesos específicos del nivel de texto, como la realización de inferencias, que consiste en la activación de información no explícita para conectar las oraciones entre sí y con el conocimiento previo, a fin de construir una representación semántica integrada y coherente de lo que el texto refiere. Las inferencias resuelven los pronombres (por ej. “a ella”, “su”), los conectores de tiempo, lugar y otros adverbiales (“entonces”, “allí”), y mantienen actualizada la representación del texto en las dimensiones temporal, espacial, de estado emocional, y de relaciones causales entre sucesos (Graesser, Singer y Trabasso, 1994; van Dijk y Kintsch, 1983). En el caso de las inferencias espaciales, diferentes estudios conductuales han demostrado que los lectores actualizan la localización espacial de un personaje o de un objeto, y pueden detectar cambios implícitos o incongruencias en el escenario espacial (ver revisión de Rinck, 2005).

Las teorías clásicas de comprensión (Kintsch, 1988; van Dijk y Kintsch, 1983) suponen que la comprensión implica ir extrayendo del texto explícito proposiciones con formato verbal o abstracto, ir añadiendo proposiciones inferidas, y organizar una jerarquía organizada de proposiciones, esto es, la base de texto. Sin embargo,

modelos más recientes cuestionan que la comprensión involucre “una representación mental del texto”, sino que lo que se construye es una representación “de la situación descrita por el texto” (Zwaan y Radvansky, 1998). Este abordaje asume que lo que se construye es una simulación, una representación multimodal, análoga a la experiencia de estar en la situación descrita por el texto, una “simulación” de la experiencia real de percibir o “estar allí” (Glenberg y Kaschak, 2003; Zwaan, 2004). Específicamente, con respecto a las inferencias espaciales, la primera hipótesis supone que se derivan proposiciones sobre la localización, en tanto que la segunda afirma que se deriva un marco de naturaleza viso-espacial, en el cual se ubican los protagonistas y objetos.

La evidencia cognitiva y conductual no permite decidir entre ambas hipótesis. Por ejemplo, se ha hallado que el tiempo de respuesta para juicios sobre la localización de un objeto se relaciona con la distancia en la que se encuentra el protagonista de la historia (Franklin y Tversky, 1990; De Vega, Rodrigo y Zimmer, 1996; Rinck y Bower, 1995, 2000). No obstante, esto se ha demostrado en paradigmas que requieren que previamente el sujeto se aprenda la descripción de un escenario espacial (un mapa de una isla o de un edificio), induciendo así una representación gráfica. Este tipo de lectura, estratégica, se contrapone con una estrategia naturalista, donde el sujeto simplemente lee para comprender lo que pasa en una historia. El paradigma de consistencia, introducido por O’Brien y Albrecht (1992), permite estudiar las inferencias de forma implícita, sin forzar al sujeto a una representación viso-espacial. En este paradigma, los sujetos leen, frase a frase, una historia que contiene un personaje y dos espacios asociados a objetos distintos. En una frase crítica se muestra al personaje en una ubicación, interactuando con un objeto que pertenece a dicha localización (condición *consistente*) o no (condición *inconsistente*). En este paradigma, que no requiere aprendizaje previo de mapas ni juicios explícitos acerca del espacio, el tiempo de lectura para la frase inconsistente es típicamente mayor que para la consistente. Tal efecto en el tiempo de lectura implica que los lectores infirieron la localización espacial y los objetos asociados. Sin embargo, desde el punto de vista de la representación subyacente, no permite decidir si el mayor tiempo de lectura es resultado de una contradicción entre proposiciones de naturaleza verbal o de un procesamiento viso-espacial de la escena.

En una serie de experimentos previos hemos explorado el rol de los recursos de la memoria de trabajo verbal y viso-espacial en las inferencias espaciales en comprensión narrativa (Irrazabal, León y Burin, enviado). Las inferencias se realizan con los recursos de la memoria de trabajo (Kintsch, 1988; Zwaan y Radvansky, 1998). Hemos empleado el paradigma de consistencia (O’Brien y Albrecht, 1992) comparando las inferencias espaciales bajo instrucciones de lectura naturalista (“leer para comprender”) versus instrucciones de lectura estratégica (“leer para formarse una imagen mental”). En síntesis, los experimentos han mostrado que bajo lectura naturalista, las interferencias selectivas para el componente espacial de la memoria de trabajo no afectan los tiempos de lectura de la frase crítica ni la tasa de errores. En cambio, bajo instrucciones de imaginarse la historia, los tiempos de lectura sí se ven

afectados por una actividad espacial concurrente. Esto apoyaría la hipótesis proposicional.

El *objetivo* del presente trabajo ha sido comparar la hipótesis verbal / proposicional y la hipótesis analógica / de simulación, tomando datos conductuales y de neuroimagen mediante SPECT, en un experimento donde los participantes leían narraciones que elicitan inferencias espaciales mediante el paradigma de congruencia / incongruencia, bajo dos condiciones: lectura para comprender las historias o lectura para formarse una imagen.

- Hipótesis verbal / proposicional: Teoría clásica de comprensión (Kintsch, 1988; van Dijk y Kintsch, 1983). Supone que la comprensión implica ir extrayendo del texto explícito proposiciones con formato verbal o abstracto, ir añadiendo proposiciones inferidas y organizar una jerarquía organizada de proposiciones. Las inferencias espaciales consistirían en proposiciones acerca de la localización.

- Hipótesis analógica / de simulación: Se construye una representación multimodal, una simulación, análoga a la experiencia de estar en la situación descrita por el texto (Glenberg y Kaschak, 2003; Zwaan, 2004). Las inferencias espaciales consistirían en una representación viso-espacial en la cual se ubican protagonistas y objetos.

Las investigaciones en pacientes con lesión cerebral y con neuroimágenes en sujetos normales han delineado circuitos diferentes para lenguaje y para percepción, imágenes, y navegación viso-espacial. La comprensión de lenguaje de alto nivel se asocia con una “red extendida de lenguaje” que comprende áreas temporales y frontales (Ferstl, Neumann, Bogler, y von Cramon, 2008). En particular, la capacidad de generar inferencias y detectar incongruencias semánticas se ha relacionado con activación temporal anterior y medial, frontal medial y cingulado (ver Ferstl et al., 2008; Mason y Just, 2004). Por su parte, la representación de objetos y escenas viso-espaciales comprendería un circuito occipito-temporal, incluyendo áreas específicas del giro fusiforme (Aguirre, Zarahn, y D’Esposito, 1998; Kosslyn y Thompson, 2003; O’Craven y Kanwisher, 2000), en tanto que otro circuito fronto-parietal se encargaría de la navegación, acciones y transformaciones en la dimensión espacial (Aguirre, Zarahn y D’Esposito, 1998; Kosslyn y Thompson, 2003; Lattini-Corazzini, Nesa, Ceccaldi, Guedj, Thinus-Blanc, Cauda, Dagata y Péruch ., 2010). Aunque existen muchos estudios de neuroimagen de la comprensión de palabra y oraciones, pocos se han ocupado del procesamiento de texto; según un reciente meta-análisis, sólo dos se han ocupado de las inferencias (Ferstl, Neumann, Bogler y von Cramon, 2008), los estudios de Ferstl, Rinck y von Cramon (2005). En un estudio con resonancia magnética funcional, empleando el paradigma de consistencia, encontraron que, además de las áreas consideradas como la “red extensa del lenguaje” (corteza perisilviana, temporal anterior, temporal ventro-medial, prefrontal lateral, todo a predominio izquierdo), las inferencias emocionales activaban la amígdala y corteza prefrontal ventro-medial. Esto daría apoyo a la idea de que la comprensión implica la activación de las zonas del cerebro que se encargan de la experiencia real de aquello descrito por el texto.

En nuestro experimento apelaremos al conocimiento de las redes del lenguaje y de la percepción viso-espacial en el cerebro para contrastar dos condiciones de lectura en las que se generan las inferencias espaciales: formando una imagen mental versus lectura naturalista. La comparación de la activación en ambas condiciones permite contrastar las hipótesis verbal / proposicional con la hipótesis analógica / de simulación, tomando datos conductuales y de neuroimagen mediante SPECT, en sujetos normales, mientras leen narraciones que elicitán inferencias espaciales implícitas. La *hipótesis verbal/ proposicional* predice activación distinta en ambas condiciones de lectura: mayor activación fronto-temporal para la lectura naturalista, y en zonas modales posteriores para la condición de formarse una imagen.

2. Método

Sujetos

12 adultos jóvenes (edad media = 26.25, *s.d.* = 4.58; 8 mujeres, 4 varones), alumnos de grado y posgrado de carreras humanísticas o sociales. Voluntarios con consentimiento informado.

La *n* final fue de 11 sujetos ya que uno de los participantes fue excluido de la muestra debido a sus valores extremos en activación a lo largo de las condiciones.

Materiales y procedimiento

- *Instrucciones de Comprender vs Imagen*: En la primera sesión, leían las historias con la consigna de “leer para comprender”. Al día siguiente, leían las historias “para formarse una imagen mental, como fotos o una película en tu mente”.

- *Congruencia vs. Incongruencia*: permite verificar la realización de inferencias espaciales en los tiempos de lectura, sin forzar un juicio explícito o el aprendizaje de un mapa o imagen.

Se emplearon 32 historias experimentales (16 por condición) adaptadas de De Vega (1995), que tenían la siguiente estructura (ver Figura 1):

Figura 1. Esquema de Estructura de un Texto y Texto de Ejemplo

Una oración que introduce un protagonista	<i>La señorita Julia vive en una típica casa inglesa</i>
Dos oraciones (4 líneas) describiendo 2 escenarios: adentro / afuera o arriba / abajo, cada uno asociado con un objeto.	<i>Delante tiene un jardín muy bien cuidado, con hermosas petunias de todos los colores. Dentro de la casa hay una colección de relojes que heredó de su abuelo.</i>

El protagonista va de un lugar a otro	<i>La señorita Julia acaba de entrar en su casa, camina con su bastón.</i>
El protagonista interactúa con un objeto asociado a ese lugar (consistente), o asociado al otro lugar (inconsistente).	<i>Como de costumbre, contempló sus relojes / sus petunias con orgullo.</i>
Pregunta	<i>¿El jardín está al fondo de la casa?</i>

Los sujetos participaron en dos sesiones, en dos días consecutivos, a la misma hora para cada sujeto. En cada sesión completaron una de las condiciones de lectura y, dentro de ella, se aleatorizó el orden de presentación de las historias. Los participantes leían las historias en una computadora, frase a frase, pasando a la siguiente apretando una tecla. Se medía el tiempo de lectura de la oración crítica (congruente / incongruente) y los aciertos en la pregunta de comprensión.

- *SPECT*: Se administró a cada sujeto el radiotrazador 30 mCi 99mTc-ECD por sesión. Luego, se adquirieron las imágenes en una cámara gamma de cabezal dual Picker Prism XP 2000 con colimador de alta resolución.

3. Resultados

El análisis de varianza para medidas repetidas mostró un efecto significativo de la congruencia ($F 1, 10 = 8.85, p = .014, \eta^2 = 0.47$) y de las instrucciones ($F 1,10 = 14.69, p = 0.003, \eta^2 = 0.59$), sobre los tiempos de lectura para la oración crítica, pero no efectos significativos de interacción. La Tabla 1 muestra que la lectura de la oración crítica en condición de “leer para formarse imágenes” (Imagen) requirió más tiempo que en condición “leer para comprender” (Comprender), y que las oraciones incongruentes se leyeron más lento en ambas condiciones.

Los resultados de neuroimagen fueron analizados con el paquete *SPM8* (FIL Methods Group, 2010). Las imágenes de *SPECT* fueron realineadas, co-registradas y normalizadas espacialmente, suavizadas con filtro de 8mm. La cuenta de *voxels* fue normalizada usando escalamiento proporcional. Se realizaron contrastes t para comparar la activación en las condiciones Comprender – Imagen, e Imagen - Comprender. Se encontraron diferencias significativas en los análisis de volumen total de *voxels*, con el umbral de activación en $p < 0.001$ y un tamaño mínimo del *cluster* de 50 *voxels* contiguos. Se ubicaron las regiones con el *software Talairach Client* v. 2.4.2 (Lancaster et al., 1997, 2000). La Tabla 2 y la Figura 2 muestran las zonas de activación diferencial.

Tabla 1. Media (MeanReadTime) y S.D. de los Tiempos de Lectura (en Milisegundos) para la Oración Crítica, y Media (MeanCorrect) del Número de Respuestas Correctas (sobre Ocho Total) en la Pregunta de Comprensión, según Congruencia y Tipo de Instrucciones de Lectura

	Mean ReadTime	S.D.	Mean Correct	S.D.
Imagery-Consistent	3720.20	912.19	6.73	1.272
Imagery-Inconsistent	4220.01	1355.01	6.91	0.944
Understand-Consistent	2967.452	940.49	6.27	0.647
Understand-Inconsistent	3379.69	962.78	6.55	1.128

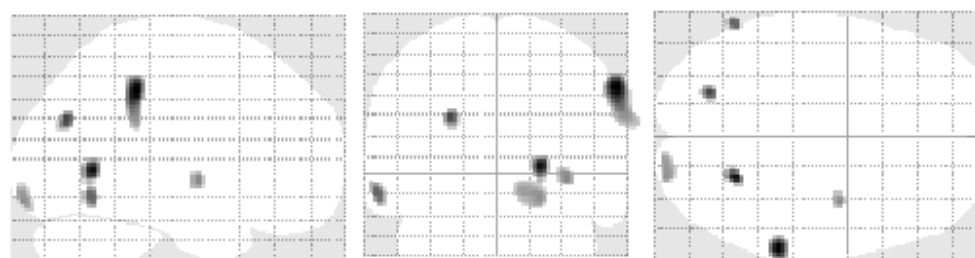
Tabla 2. Localización de Regiones de Picos Máximos de Aumento en el Flujo Sanguíneo Cerebral Regional (Normalizados), en las Condiciones de Lectura de Textos Espaciales bajo la Condición de Imágenes Mentales, en Comparación a la Lectura para Comprender (*Imagery* – *Understand*) y en la Comparación Inversa (*Understand* – *Imagery*)

Anatomical Location of Peak Voxels	Size (n voxels)	Coordinates			Z score
		x	y	z	
Imagery - Understand					
R Lingual Gyrus	104	22	-58	4	3.26
L Cuneus	123	-24	-72	30	3.09
R Claustrum sub lobar	72	36	-4	-2	2.89
R Fusiform Gyrus	226	22	-94	-12	2.86
R Middle Occipital Gyrus	66 *	40	-78	16	2.67
Understand - Imagery					
L Superior Temporal Gyrus	211	-60	14	-6	3.86
R Medial Frontal Gyrus	345	12	36	40	3.73
R Insula	75	42	18	4	3.57
R Temporal Sub-gyrus (medial)	274	30	4	-26	3.06
L Cerebellar Tonsil	232	-14	-50	-38	3.03

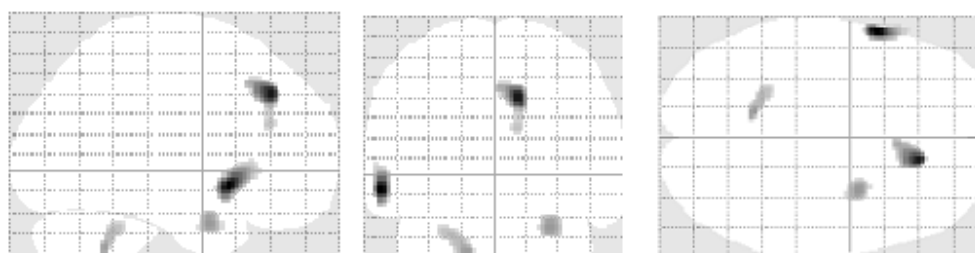
Nota: Resultados significativos dada una combinación de umbral de Z ($p < 0.001$, a una cola, sin corrección), y un umbral de tamaño mínimo para el cluster (50 vóxeles contiguos). La localización

anatómica corresponde al atlas de Talairach y Tournoux con sus coordenadas en mm (L: izquierda, D: derecha). * $p < 0.00$

Figura 2. Esquema de Vistas Sagital, Coronal y Axial, de las Localizaciones de Regiones de Picos Máximos de Aumento en el Flujo Sanguíneo (Normalizados), en las Condiciones de Lectura de Textos Espaciales bajo la Condición de Imágenes Mentales, en Comparación a la Lectura para Comprender (*Imagery – Understand*) y en la Comparación Inversa (*Understand – Imagery*)



Imagery – Understand



Understand – Imagery.

4. Discusión

El patrón de resultados conductuales replicó el efecto de consistencia: el tiempo de lectura para las oraciones que contenían una incongruencia espacial fue significativamente mayor que para las congruentes. Esto supone que los sujetos realizaban la inferencia espacial. Los tiempos de lectura fueron mucho mayores bajo la instrucción de formarse una imagen mental, lo cual apoyaría la idea de que generaban una imagen mental para la comprensión.

El análisis de las neuroimágenes mostró que la resta de la activación de las condiciones Imagen – Comprender arrojó un saldo de *vóxeles* activos a predominio de hemisferio derecho, alrededor del giro fusiforme, giro lingual, y occipital medial, así como en el cúneo izquierdo. Todas estas áreas corresponden a la vía de procesamiento de objetos y escenas visuales. Por otro lado, la comparación de *vóxeles* activos en la resta Comprender - Imagen mostró activación significativamente mayor

en zonas mediales fronto-temporales derechas, y en el polo temporal izquierdo, lo cual converge con la “red extendida de lenguaje” de estudios de comprensión semántica de alto nivel.

Estos resultados apoyarían la hipótesis lingüística-proposicional, en la medida en que muestran que la comprensión narrativa bajo instrucciones de “leer para comprender” implica principalmente zonas de lenguaje y semánticas, en tanto que cuando los lectores usan estratégicamente la capacidad de formarse imágenes mentales se reclutan recursos cognitivo-neurales de representación viso-espacial.

Dado el número pequeño de participantes y el uso de *SPECT*, con sus limitaciones de resolución espacial y temporal, se debe considerar el conjunto de resultados como exploratorio, a investigar en el futuro con muestras mayores y técnicas con mayor definición.

Bibliografía

Aguirre, G. K., Zarahn, E. y D'Esposito, M. (1998). Neural components of topographical representation. *Proceedings of the National Academy of Sciences of the USA*, 95, 839–846.

De Vega, M. (1995). Backward updating of mental models during continuous reading of narratives. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 373-386.

Ferstl, E. C., Neumann, J., Bogler, C. y von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, 29, 581-593.

FIL Methods Group (2010). SPM8 Manual. London: Wellcome Trust Centre for Neuroimaging. Retrieved 4/14/2010 from <http://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf>

Glenberg, A. M. y Kaschak, M. P. (2003). The body's contribution to language. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, V43 (pp. 93-126). New York: Academic Press.

Graesser, A. C., Singer, M. y Trabasso, T (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371-395.

Irrazábal, N., León, J. A. y Burin, D. I. (no publicado). *Verbal and Spatial Working Memory in Spatial Inferences during Narrative Comprehension*.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A Construction-Integration model. *Psychological Review*, 95, 163-182.

Kosslyn, S. M. y Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129, 723– 746.

Lancaster J.L, Rainey L.H, Summerlin J.L, Freitas C.S, Fox P.T, Evans A.E, Toga A.W y Mazziotta J.C. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Hum Brain Mapp* 5:238 –242.

Lancaster, J.L; Woldorff, M.; Parsons, L.M; Liotti, M.; Freitas, C.S; Rainey, L.; Kochunov, P.V; Nickerson, D.; Mikiten, S. y Fox, P.T. (2000). Automated Talairach Atlas Labels For Functional Brain Mapping. *Human Brain Mapping* 10:120 –131.

Lattini-Corazzini, L., Nesa, M. P., Ceccaldi, M., Guedj, E., Thinus-Blanc, C., Cauda, C., Dagata, F. y Péruch, F. P. (2010). Route and survey processing of topographical memory during navigation. *Psychological Research*, e-pub, retrieved on April 28, 2010, DOI: 10.1007/s00426-010-0276-5.

Mason, R. A. y Just, M. A. (2004). How the brain processes causal inferences in text: A theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychological Science*, 15, 1-7.

O'Craven, K. y Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12, 1013–1023

Van Dijk, T. A. y Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic.

Zwaan, R. (2004). The immersed experience: toward an embodied theory of language comprehension. In B.H. Ross. (Ed.). *The psychology of language and motivation*, 44. New York: Academic Press.

Zwaan, R y Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*. 123, 162-185.

La investigación fue financiada por el Consejo Nacional de Investigaciones Científicas y Técnicas – CONICET - y por la Secretaría de Ciencia y Técnica, Universidad de Buenos Aires - UBACYT.