

# Efectos de la complejidad y la presencia de subtítulos en textos expositivos sobre la comprensión lectora en estudiantes de pregrado\*

**Effects of the Complexity and Presence of Subtitles in Expository Texts on Undergraduate Students' Reading Comprehension**

**Efeitos da complexidade e a presença de legendas em textos expositivos sobre compreensão de leitura em estudantes de graduação**

**Geral Eduardo Mateus Ferro\*** Orcid: <https://orcid.org/000-0001-6166-2014>

**Lida Johanna Rincón Camacho\*\***

**Jenny Viviana Román González\*\*\***

\* Este artículo presenta los resultados de la primera fase del proyecto de investigación titulado "Estudio cognitivo sobre la relación entre la comprensión de lectura y la producción de textos académicos en estudiantes de pregrado" (DLE-457-17), financiado por el Centro de Investigaciones de la Universidad Pedagógica Nacional (CIUP-UPN).

\*\* Profesor investigador del Departamento de Lenguas de la Universidad Pedagógica Nacional. Director del grupo de investigación gipelec.  
**Correo electrónico:** [gmateus@pedagogica.edu.co](mailto:gmateus@pedagogica.edu.co);

\*\*\* Profesora investigadora del Departamento de Lenguas de la Universidad Pedagógica Nacional. Miembro del grupo de investigación Estilos Cognitivos.  
**Correo electrónico:** [ljrincon@pedagogica.edu.co](mailto:ljrincon@pedagogica.edu.co).

\*\*\*\* Profesora investigadora del Departamento de Lenguas de la Universidad Pedagógica Nacional. Miembro del grupo de investigación gipelec. [jvromang@pedagogica.edu.co](mailto:jvromang@pedagogica.edu.co).



## Para citar este artículo

Mateus, G., Rincón, L. y Román, J. (2019). Efectos de la complejidad y la presencia de subtítulos en textos expositivos sobre la comprensión lectora en estudiantes de pregrado. *Folios*, 50, 51-64.  
doi: 10.17227/Folios.50-10177

Artículo recibido  
2•05•2018

Artículo aprobado  
25•01•2019

## Resumen

En este artículo, se presentan los resultados de un estudio descriptivo comparativo realizado con 32 estudiantes de pregrado, cuyo objetivo era examinar la representación mental de un texto expositivo después de su lectura y la influencia del grado de complejidad y de la presencia o no de subtítulos sobre la comprensión de lectura. La diferencia en la complejidad de los textos se basa en el uso de léxico especializado. El principal resultado es que la representación mental de los estudiantes contiene más información proveniente de la base del texto cuando la complejidad es mayor; en tanto que en los textos menos complejos predomina la información correspondiente al modelo de situación y, en especial, a las inferencias asociativas.

### Palabras clave

comprensión de lectura; procesos cognitivos; base del texto; modelo de situación; representación mental

## Abstract

In this article, the results of a comparative descriptive study carried out with 32 undergraduate students are exposed, whose objective was to examine the mental representation of an expository text after its reading and the influence of the degree of complexity and the presence or absence of subtitles on reading comprehension. The difference in the complexity of the texts is based on the use of specialized lexicon. The main result is that the mental representation of the students contains more information coming from the text base when the complexity is greater; in the case of less complex texts, the information corresponding to the situation model predominates, especially the associative inferences.

### Key words

reading comprehension; cognitive processes; situation model; text base; mental representation

## Resumo

Neste artigo, apresentamos os resultados de um estudo descriptivo comparativo realizado com 32 estudantes de graduação, cujo objetivo foi examinar a representação mental de um texto expositivo após sua leitura e a influência do grau de complexidade e da presença ou não de legendas na compreensão de leitura. A diferença na complexidade dos textos é baseada no uso de léxico especializado. O principal resultado é que a representação mental dos alunos contém mais informações vindas da base do texto quando a complexidade é maior; enquanto nos textos menos complexos predomina a informação correspondente ao modelo de situação e, principalmente, às inferências associativas.

### Palavras-chave

compreensão de leitura; processos cognitivos; base de texto; modelo de situação; representação mental

Los resultados sobre los procesos de comprensión de lectura en estudiantes de pregrado revelan la urgencia de pensar estrategias que favorezcan la formación en competencias del lenguaje en contextos académicos. Los puntajes en pruebas estandarizadas como las pruebas PISA<sup>1</sup> muestran que los estudiantes próximos a terminar la educación secundaria no están desarrollando las competencias esperadas en estos niveles. Por ejemplo, los más recientes resultados demostraron que el 47 % de los bachilleres colombianos está por debajo del nivel mínimo de la prueba de comprensión (Nivel 2).

Adicionalmente, los resultados de las pruebas Saber 11<sup>2</sup> del año 2016 muestran un desempeño promedio de 52,6 puntos en un rango posible entre 0 y 100 en lectura crítica. De acuerdo con los informes del ICFES, este valor ubica a los estudiantes en el umbral inferior del nivel de desempeño tres, caracterizado por “la capacidad de interpretar información de textos al inferir contenidos implícitos y reconocer estructuras, estrategias discursivas y juicios valorativos”; no obstante, no “se identifican la visión de mundo del escritor ni elementos paratextuales” (ICFES, 2017). Por otro lado, las pruebas Saber Pro<sup>3</sup> del mismo año reportan, para el caso

- 1 El programa pisa (Programme for International Student Assessment) se propone evaluar los sistemas educativos a nivel mundial, para lo cual aplican pruebas cada tres años en las que se examina el rendimiento de estudiantes de 15 años en áreas específicas de la formación escolar. Para el 2015, cerca de medio millón de estudiantes de 72 países tomaron la prueba, en la que se le evaluó en ciencias, matemáticas, lectura, problemas colaborativos y educación financiera. Para ampliar información sobre la prueba, se puede consultar <http://www.oecd.org/pisa>
- 2 El examen de Estado de la Educación media Saber 11° es una prueba estandarizada diseñada para los estudiantes que estén finalizando el grado undécimo, con el fin de obtener resultados oficiales que les permitan ingresar a la educación superior. También pueden presentarlo quienes ya hayan obtenido el título de bachiller o superado el examen de validación del bachillerato. Se compone de 5 pruebas: Lectura crítica, matemáticas, sociales, ciencias naturales e inglés. Su objetivo principal es comprobar el grado de desarrollo de las competencias de los estudiantes que están por finalizar el grado undécimo de la educación media. Para ampliar información sobre la prueba puede consultar <http://www2.icfes.gov.co/instituciones-educativas-y-secretarias/saber-11/informacion-de-la-prueba-saber11>
- 3 El examen de Estado de calidad de la educación superior, Saber Pro, es un instrumento estandarizado para la evaluación externa de la calidad de la educación superior. Se compone de 5 módulos que evalúan competencias genéricas: Lectura crítica, razonamiento cuantitativo, competencias ciudadanas, comunicación escrita

específico de los estudiantes de licenciaturas, un promedio de 142 puntos (rango entre 0 y 300), en la competencia genérica de lectura crítica, ubicándose por debajo de la media global ( $m = 150$ ). Estos datos corresponden al nivel de desempeño dos, que indica una comprensión del sentido global del texto, además de la realización de inferencias a partir de la información explícita del texto (ICFES, 2017).

Esto demuestra que, si bien en la universidad los estudiantes trabajan sus habilidades de lectura, no logran superar el umbral obtenido en el bachillerato; de hecho, hay una involución en el proceso, debido tal vez a la escasa formación en lectura académica que reciben los estudiantes, como lo demuestran algunas investigaciones recientes (González, Salazar y Peña, 2015).

En ese sentido, se requiere adelantar investigaciones para determinar en detalle qué aspectos relacionados con la comprensión en estudiantes universitarios pueden incidir en la calidad de este proceso. Por esta razón, el estudio que se reporta se concentra en el análisis de la representación mental que surge del proceso lector, atendiendo a las coincidencias y diferencias que tienen lugar entre el texto efectivamente leído y la representación que el lector crea. En tanto, los textos expositivos son característicos del contexto académico, por lo que este es el tipo de texto que se emplea en la investigación. Particularmente se analizan los resultados de una prueba de recuerdo a partir de la teoría de los niveles de representación mental de Kintsch y Van Dijk (1978) y Van Dijk y Kintsch (1983).

e inglés. Además de los anteriores, hay 40 módulos asociados a temáticas y contenidos específicos que los estudiantes tienen la posibilidad de presentar de acuerdo con su área de formación profesional. La aplicación de los módulos de competencias genéricas que conforman el examen de Estado Saber Pro tienen como objetivo evaluar y proporcionar un reporte del grado de desarrollo de habilidades y conocimientos generales de estudiantes que han aprobado el 75 % de los créditos de sus respectivos programas de formación universitaria profesional. Para ampliar información sobre la prueba puede consultar <http://www2.icfes.gov.co/instituciones-educativas-y-secretarias/saber-pro/informacion-de-la-prueba-saber-pro>

## Los tres niveles de representación

El modelo de los tres niveles de representación (Kintsch y Van Dijk, 1978; Van Dijk y Kintsch, 1983) plantea que el propósito último del proceso de comprensión es acceder al significado del texto. En este proceso, el lector crea una representación mental (RM) compleja que integra la información proveniente del texto con la aportada por él mismo. La información añadida por el individuo proviene de su conocimiento previo y, en general, se puede identificar a través de las inferencias que construye.

Esta representación mental recoge, entonces, información variada y puede analizarse en tres niveles: *superficial*, *base del texto* y *modelo de situación*. El primer nivel guarda información que conserva aspectos formales del texto. En este caso, se almacena, entre otros elementos, el léxico literal así como aspectos sintácticos específicos en la memoria de corto plazo (MCP). El segundo nivel, *base del texto*, da cuenta de los significados construidos en la RM a partir de las proposiciones explícitas en el texto y las inferencias obligatorias (o inferencias puente, *bridging inferences*), necesarias para la conexión entre dichas proposiciones explícitas. En el tercer nivel, se integran las proposiciones del nivel de base del texto con las inferencias elaborativas y con el conocimiento previo del lector y los esquemas experienciales activados. La información correspondiente a este nivel de modelo de situación ha demostrado ser el de mayor persistencia en la memoria de largo plazo (MLP) (Kintsch, Welsch, Schmalhofer y Zimny, 1990).

Un análisis de las representaciones mentales de textos expositivos generadas por estudiantes de educación media en Bogotá, llevado a cabo con base en el análisis de los niveles de la representación mental de Kintsch y Van Dijk (1978), demostró que dichas representaciones integraban más información correspondiente al nivel modelo de situación (MDS) que al de base del texto (BT). En este sentido se estableció, también, que en el nivel modelo de situación, los conocimientos previos son una fuente de información más frecuente que

la retomada de la base del texto (Mateus, Castillo, Santiago y Rodríguez, 2012).

## Inferencias, su taxonomía y tipología

Lo presentado anteriormente evidencia la relevancia de las inferencias en la construcción de la representación mental de un texto, pues a través de ellas se establece un puente entre la información leída y los conocimientos previos del lector (León, 2003). Así, el lector debe llenar los vacíos que deja el texto, o integrar la información nueva con la ya existente en su memoria a largo plazo, lo cual genera nuevos conocimientos que brindan coherencia y sentido a lo que se lee (Santiago, Castillo y Mateus, 2014). Si bien ha habido dificultades en la configuración de un modelo teórico integrador que dé cuenta de los diversos tipos de inferencias, existe un modelo teórico cuyas tipificaciones interesan a los propósitos del presente estudio.

El modelo propone una taxonomía dicotómica: por un lado, están las inferencias *obligatorias* o *integrativas*, las cuales contribuyen a la coherencia local, se dan de modo automático y son necesarias para elaborar una representación coherente del texto (Just y Carpenter, 1987; Keefe y McDaniel, 1993; McKoon y Ratcliff, 1995); por otro lado, están las *inferencias elaborativas*, las cuales no son esenciales para la coherencia, pero sí para la integración del conocimiento previo (Van den Broek, 1994; Van den Broek, Fletcher y Ridsen, 1993; Trabasso y Magliano, 1996). Dentro de esta última tipología están las inferencias *explicativas* (que implican relaciones con elementos previos del texto y la creación de vínculos causales), las *asociativas* (que permiten la activación de información almacenada en el conocimiento previo, la cual se relaciona con lo que se lee) y las *predictivas* (que proyectan posibilidades, consecuencias y estados futuros con base en lo presentado en el texto).

Fincher-Kiefer (1994) integra esta taxonomía a los niveles de representación propuestos por Kintsch y Van Dijk (1978) y Van Dijk y Kintsch (1983), pues alude a un continuo en el que las inferencias van surgiendo según el nivel: en la estructura

superficial no se dan procesos inferenciales por enfocarse en la decodificación, mientras que en la base del texto se dan procesos inferenciales de tipo integrativo, limitados a la obligatoriedad que plantea la coherencia mediante las asociaciones entre conceptos. Por su parte, en el modelo de situación, dada su naturaleza interpretativa, se generan las inferencias elaborativas. El presente estudio parte de esta concepción integrativa, con el fin de caracterizar los procesos inferenciales de los participantes de acuerdo con el nivel en el que se presentan.

### **El efecto de la tipología del texto en el establecimiento de inferencias**

La investigación sobre el proceso de generación de inferencias en el marco de la comprensión lectora ha abierto un abanico de posibilidades dentro de las cuales cabe considerar la influencia que ejerce el tipo de texto leído. En términos generales, puede definirse una tipología textual como “una forma textual definida por sus características internas (estructurales y gramaticales)” (Adam, 1997, p. 4), cuya estructura ha sido establecida socialmente a través del uso, y que determina las condiciones de producción e interpretación de información en una situación comunicativa específica. Con el tiempo, estos tipos de texto comienzan a responder a patrones de organización general en función de contextos de comunicación recurrentes.

Dentro de estos patrones han emergido diversas clasificaciones de las tipologías textuales. Una de las más destacadas, debido a su carácter integrador, es la de Adam y Revaz (1996), la cual se basa en aspectos estructurales y funcionales para establecer la clasificación conformada por textos narrativos, descriptivos, argumentativos, explicativos y conversacionales. No obstante, la investigación se ha centrado en los narrativos y explicativos.

Los resultados obtenidos en diversos estudios evidencian que los textos narrativos facilitan, en general, la generación de un mayor número de inferencias elaborativas que los de tipo expositivo; de hecho, los textos narrativos se comprenden y se recuerdan mejor, permiten una mejor integración

de la información leída, así como un procesamiento global de la misma (Langer, 1990; Zaburchy y Ratner, 1992).

Dentro del repertorio de inferencias elaborativas, en la lectura de textos expositivos se identifica un mayor número de explicaciones y asociaciones que, como ya se expuso, implican establecer una relación causal antecedente del texto con información previamente leída o con el conocimiento previo del lector; lo cual genera una exigencia mayor para la memoria de trabajo. Por su parte, en los textos narrativos se genera un número mayor de predicciones, que relacionan causalmente aspectos del texto con información previa, pero de manera prospectiva; esta operación requiere un esfuerzo mayor para recuperar conocimiento proveniente de la memoria a largo plazo (León, Escudero y Van den Broek, 2003).

### **La complejidad del texto y el efecto de los subtítulos como variables para la comprensión**

Usualmente, los estudios sobre comprensión de lectura han examinado los efectos de variables asociadas al lector, al contexto de la tarea y al texto. Particularmente, este estudio se centra en aspectos relacionados con algunas características de los textos, a saber, su *complejidad*, directamente vinculada con el grado de dificultad en su comprensión, y el *efecto de la presencia o no de subtítulos*.

La complejidad del texto se examina desde la facilidad/dificultad que el lector tiene para recuperar la representación léxica de una palabra, expresada en una fijación más o menos prolongada en la misma (Cutler, 1983; Rayner y Duffy, 1986). Parte de los factores que hacen que el procesamiento sea complejo son la frecuencia de uso, la familiaridad y el significado de las palabras (Klare, 1968). Este estudio se concentra en el efecto de la frecuencia léxica para determinar su nivel de complejidad. Esta variable determina que a mayor exposición a una palabra, mejor será la recuperación de su representación léxica y semántica, dicho de otro modo, el procesamiento de una palabra más

frecuente será más eficaz que el de una menos frecuente (Arya, Hiebert y Pearson, 2011).

Los estudios sobre frecuencia de uso han demostrado que esta explica en buena parte la activación del significado de una palabra por parte del lector, proceso que es fundamental en la comprensión de lectura (Arya, Hiebert y Pearson, 2011; Fajardo, Hernández y González, 2012). En esta medida, un texto que posea un amplio grupo de palabras de baja frecuencia de uso, como en el caso de textos científicos, representa un desafío para el lector ya que las palabras especializadas tienen un grado de precisión elevado, lo que deja en desventaja a quien no conozca su significado y que no logre inferirlo por el mismo contexto (Arya, Hiebert y Pearson, 2011). Al respecto, se ha explicado que, al no estar automatizado el proceso de acceso al léxico para las palabras de baja frecuencia, la integración de este significado en un procesamiento de nivel superior (por ejemplo, construir la representación global de un texto) se dificulta y la comprensión de lectura es escasa (Fajardo, Hernández y González, 2012).

Otra variable que es comúnmente analizada en estudios de recuerdo libre es el efecto del título (y subtítulo) en la comprensión, estos estudios pretenden determinar en qué medida la comprensión de textos se modifica con la presencia/ausencia de título, o su posición en el texto (antepuesto/pospuesto) (Dooling y Lachman, 1971; Migueles y García, 1991; Sellán, Sanz y Valle, 1995). En particular, Dooling y Lachman (1971) demostraron que quienes leen un texto con título al comienzo, obtienen resultados significativamente mejores que quienes leen sin este. Adicionalmente, Sellán et al. (1995) encontraron que la aparición del título al finalizar la lectura produce mejores respuestas a las preguntas puntuales que se formulan para medir la comprensión. De acuerdo con estos hallazgos, la mejora en los resultados de comprensión no se debería al conocimiento general que tienen los lectores sobre el tema de los textos, sino más bien, a la información que aportan el título y los paratextos para la explicación del tema. Un impacto positivo de esta variable se evidencia

en la desambiguación de textos, especialmente en tareas con textos experimentales (Bernárdez, 1995; Bransford y Johnson, 1972), y en el apoyo que brindan los subtítulos en la formación de la macroestructura durante la lectura de un texto. La presencia de los mismos supondría la facilitación de claves semánticas para la generación del esquema mental jerárquico de los significados (Van Dijk, 1980; Kinstch, 1998).

Bajo este marco teórico, este estudio se propone responder a la pregunta ¿cuál es el efecto de la complejidad de textos expositivos (con y sin subtítulos) sobre los niveles de representación mental en una tarea de recuerdo? Adicionalmente, los datos obtenidos también permiten responder algunos interrogantes específicos relacionados con las características de las representaciones mentales de los estudiantes, a saber: ¿cómo es la distribución de la información (proposiciones) en función de los niveles BT y MDS? ¿Qué tipo de inferencias predominan en el mds? ¿Las proposiciones “recordadas” en el mds son consistentes o inconsistentes frente a la base del texto?

## Metodología

El estudio siguió un diseño descriptivo comparativo, con metodología mixta, cuyo objetivo fue examinar los procesos de comprensión de lectura de 32 estudiantes de pregrado, en los programas de licenciatura en Español y Lenguas Extranjeras y en Español e Inglés de la Universidad Pedagógica Nacional, a través de una prueba de recuerdo libre. Específicamente, el estudio se propuso analizar los efectos de la complejidad sobre algunas medidas de recuerdo.

Complejidad	Subtítulos	
	SI	NO
Menor	01	02
Mayor	03	04

**Figura 1.** Esquema de diseño del estudio

**Fuente:** elaboración propia.

Para la prueba se seleccionaron dos textos de carácter expositivo, cada uno con dos versiones: con

y sin subtítulos. En ese orden de ideas, la muestra se dividió en cuatro grupos, según las características del texto a leer (ver figura 1). O1 y O2 recibieron el texto “Los grupos y la comunicación” (menor complejidad), el primero con subtítulos y el segundo sin ellos. O3 y O4 recibieron el texto “Efectos de un andamiaje metacognitivo y adaptativo sobre el aprendizaje y la autorregulación en cursos de educación virtual” (mayor complejidad), con sus respectivas variaciones según presencia o ausencia de subtítulos. En la medida en que el estudio hacía parte del desarrollo regular de una clase de pregrado, los textos trataban temas que hacían parte de la metodología de la clase que cursaban los estudiantes de la muestra.

El estudio tiene tres variables, a saber: *la complejidad del texto original*, definida por la presencia de léxico especializado: el texto 2 presenta mayor dificultad de acceso al léxico que el texto 1; la segunda variable es *el subtítulo*, cuyos indicadores corresponden a la presencia o ausencia de subtítulos en el cuerpo del texto; y la tercera es *el nivel de representación mental*, que tiene como indicadores el tipo de información (proposiciones) correspondiente a la base del texto (BT) y al modelo de situación (MDS).

### Participantes

Para la realización de la prueba de recuerdo libre fueron seleccionados 32 estudiantes de primer semestre de las licenciaturas ya mencionadas (21 mujeres y 11 hombres) con edades entre los 17 y 22 años, que fueron informados de la naturaleza y propósitos de la prueba y manifestaron su consentimiento por escrito.

### Instrumentos diseñados para la recolección de la información

Para la aplicación de la prueba de recuerdo libre se contó con dos instrumentos: textos para lectura y consigna de tarea de escritura. Para la fase de sistematización se empleó la matriz de registro y categorización de Mateus et al. (2012).

### Textos para la lectura

Para la prueba se seleccionaron dos textos de carácter expositivo. El primero está compuesto por 109 proposiciones (777 palabras) y su tema corresponde a la comunicación humana y la conformación de grupos. El texto 2 cuenta con 92 proposiciones (750 palabras) y el tema es la educación virtual y la autorregulación del aprendizaje. Los textos fueron seleccionados con el fin de establecer diferencias en términos de la complejidad de su contenido, según la frecuencia de uso del léxico<sup>4</sup> presente en los textos. El índice de frecuencia se examinó a través del Corpus de Referencia del Español Actual (CREA)<sup>5</sup>. Los resultados señalan que el texto 1 presenta un promedio de índice de frecuencia mayor ( $m= 37837,430$ ), en comparación con el texto 2 ( $m= 27461,612$ ). La diferencia entre los promedios, correspondiente a 10375,818, indica que la complejidad del texto 1 es menor que la del texto 2.

Para garantizar que los textos seleccionados fueran efectivamente diferentes en tema y grado de complejidad, se realizaron pruebas piloto con un grupo de estudiantes de características semejantes a los participantes de la prueba, quienes leyeron los textos y respondieron a la tarea de escritura. En el pilotaje se estableció la media de tiempo de lectura para cada texto. La tabla 1 presenta los datos de la media de tiempo.

**Tabla 1.** Media de tiempo (minutos) por texto

	Texto 1	Texto 2
Media de tiempo de lectura	6' 42"	7' 42"

Fuente: elaboración propia.

Como resultado del pilotaje se estableció el tiempo de lectura para la prueba de recuerdo libre en 15 minutos. Además, se confirmó que el texto de mayor complejidad (texto 2) tardaba más tiempo en ser leído. Adicionalmente, en el pilotaje se probó la

4 Se examinaron los índices de frecuencia de los sustantivos, verbos, adjetivos y adverbios.

5 Real Academia Española: Banco de datos (CREA) [en línea]. Corpus de Referencia del Español Actual <<http://corpus.rae.es/creanet.html>> [18 de octubre de 2018].

consigna de la tarea, con el fin de verificar su pertinencia y claridad. Los participantes manifestaron haber comprendido cabalmente la consigna por lo que se mantuvo igual para la prueba de recuerdo libre:

“Lea detenidamente el siguiente texto. Asegúrese de haber comprendido todo. Si es preciso relea.”

### Consigna de tarea de escritura

La tarea de escritura tenía como propósito hacer que el participante reprodujera el texto leído de la manera más fiel posible. Así que, una vez terminada la lectura, se le entregaba la hoja con la consigna de esta tarea:

“Reescriba lo más fielmente posible el texto que acaba de leer.”

El tiempo establecido para la tarea de escritura fue de 25 minutos.

### Matriz de análisis

Se ajustó y validó una matriz de registro y de categorización de las proposiciones que integran los textos de los estudiantes, con base en el trabajo previo de Mateus et al. (2012, p. 19-20). Este instrumento clasifica las proposiciones según los niveles de representación base del texto (BT) y modelo de situación (MDS) (Kintsch y van Dijk, 1978; Kintsch, 1998). Así, en BT, se registran las proposiciones provenientes de las ideas explícitas en el texto<sup>6</sup>; en el MDS, la información aportada por el lector se clasifica en términos del tipo de inferencia al que pertenece (asociación, explicación, predicción) y a las posibles evaluaciones que los participantes pueden registrar en la tarea. Adicionalmente, dentro de cada inferencia se analiza si la información guarda alguna relación con las ideas explícitas de la BT o si hace parte del conocimiento previo; también

6 Las proposiciones del nivel BT cobijan tanto las escritas de una forma muy cercana a la escritura original (literal), como las que presentan algunas modificaciones, las cuales, no obstante, implican que se conserva el significado explícito en el texto. Dichas modificaciones de la información corresponden a sustituciones léxicas (por ejemplo, mediante el uso de sinónimos), modificaciones sintácticas y elisión de palabras; no equivalen a las originadas por la operación de las macrorreglas propuestas por Van Dijk (1980) como pautas de proyección entre el nivel micro y macroproposicional.

se determina si la información integrada en forma de inferencia es o no consistente (contradictoria) con la contenida en el texto original (ver tabla 2).

### Procedimiento

La prueba se aplicó de forma colectiva, en una sesión de trabajo que duró alrededor de una hora. La sesión inició con una explicación detallada de la tarea; luego, se distribuyeron los textos a los estudiantes para su lectura y finalmente se dio paso a la tarea de escritura. Para la etapa de lectura, se pidió a los estudiantes que leyeran detenidamente el texto; se les señaló que se aseguraran de haber comprendido todo y se les dio la opción de releer. Esta fase tuvo una duración de 15 minutos. Para la etapa de escritura, se entregó una hoja en blanco y se les solicitó que reescribieran lo más fielmente posible el texto que acababan de leer. Para esta tarea de escritura se dieron 25 minutos.

Para el examen de las pruebas se utilizó la técnica de análisis proposicional. Inicialmente se segmentaron los textos originales en las proposiciones que los integraban. De la misma manera, los textos producidos en las pruebas se segmentaron en proposiciones y estas se clasificaron según las categorías de análisis ya descritas. En los análisis de cada prueba se registró la frecuencia para cada una de las categorías en los niveles BT y MDS de la matriz y se consolidaron los resultados generales en términos de la frecuencia media de cada ítem de la matriz (ver tabla 2). En la última etapa del estudio, se realizaron varios análisis estadísticos con el fin de examinar la media de proposiciones producidas por la muestra, comparadas según las variables de complejidad y subtítulo. Se examinó la varianza a partir de pruebas ANOVA, en el programa SPSS.

### Resultados

La tabla 2 presenta los resultados de la prueba de recuerdo; se encuentran organizados según los dos niveles de la representación analizados y las subcategorías en que se desglosan.

Para dar cuenta de los resultados, en primer lugar se presentan los relativos a la generación de propo-



siones según las variables *complejidad* (mayor y menor) y *subtítulo* (sin subtítulos y con subtítulos); enseguida se desarrollan los correspondientes al contraste entre las proposiciones según los niveles base del texto (BT) y modelo de situación (MDS); finalmente, se exponen los específicos para cada nivel de acuerdo con las categorías específicas (BT: literal/ transformada; MDS: tipo de inferencia [asociación/ explicación/predicción] y origen de la información

[BT / conocimientos previos]). Los resultados de las categorías con valor cero (0,0) no se tienen en cuenta en los resultados por no tener ocurrencias.

### 1. Generación de proposiciones

La media de proposiciones de los estudiantes que leyeron el texto más complejo fue de 32,4 (figura 2), en tanto que la de los que leyeron el menos complejo fue de (31,2), sin que el factor *número de proposicio-*

**Tabla 2.** Resultados por categoría. (Excepto los datos de la última fila de tabla, las cifras presentadas corresponden a la media de frecuencia de cada categoría).

		Complejidad		Mayor		Menor		
		subtítulo		No	Si	No	Si	
Base del texto	Literal			8,3	3,1	2,6	1,4	
	Modificada			13,5	17,0	9,1	10,8	
	Total base texto			21,8	20,1	11,8	12,2	
	Total subtítulo colapsado			<b>21</b>		<b>12,0</b>		
Asociaciones	Base del texto (BT)	Consistente		4,3	5,9	6,3	7,9	
		No consistente		2,4	2,6	4,1	4,0	
	Conocimiento previo (CP)	Consistente		1,4	2,6	4,5	3,3	
		No consistente		2,0	1,6	2,4	3,1	
	Total fuente BT			6,6	8,4	10,4	11,9	
	Total fuente CP			3,4	4,1	6,9	6,4	
	Total Asociaciones			10,0	12,6	17,3	18,3	
<b>Total Asociaciones subtítulo Colapsado</b>				<b>11,3</b>		<b>17,8</b>		
Modelo de la situación	Explicaciones	Base del texto	Consistente		0,0	0,0	0,4	1,3
			No consistente		0,0	0,0	0,0	0,1
		Conocimiento Previo	Consistente		0,0	0,0	0,0	0,6
			No consistente		0,0	0,0	0,0	0,3
	Total Explicaciones			0,0	0,0	0,4	2,3	
	<b>Total Explicaciones subtítulo colapsado</b>				<b>0,0</b>		<b>1,4</b>	
	Predicciones	Consistente			0,0	0,0	0,0	0,0
		No consistente			0,0	0,0	0,0	0,0
	Evaluaciones	Contenido			0,0	0,0	0,0	0,0
		Forma			0,0	0,0	0,0	0,0
Tarea				0,0	0,0	0,0	0,0	
Total Modelo de Situación				10,0	12,6	17,6	20,7	
<b>Total Modelo de Situación subtítulo colapsado</b>				<b>11,2</b>		<b>19,2</b>		
<b>Número de proposiciones producidas</b>				<b>31,8</b>	<b>32,7</b>	<b>29,4</b>	<b>32,9</b>	

Fuente: elaboración propia.

nes se presente como estadísticamente significativo ( $F=0,66$   $p < 0,05$ ). En tanto la variable *subtítulo* no presenta efecto significativo, en los siguientes apartados se da cuenta de los resultados según la variable complejidad.

En consecuencia, el número de proposiciones generadas para cada texto (más complejo vs menos complejo) no presentó diferencias, ni tampoco la presencia o ausencia de subtítulos en los mismos durante su lectura.

## 2. Base del texto (BT) vs Modelo de Situación (MDS)

No obstante los resultados presentados en el anterior apartado, sí se presenta una interacción entre los factores que involucran el *Nivel de la representación* (Base del texto o Modelo de Situación)  $\times$  *complejidad* ( $F=7,63$   $p < 0,05$ ). La figura 3 da cuenta de las diferencias en la media de proposiciones en los niveles BT y MDS en los dos textos.

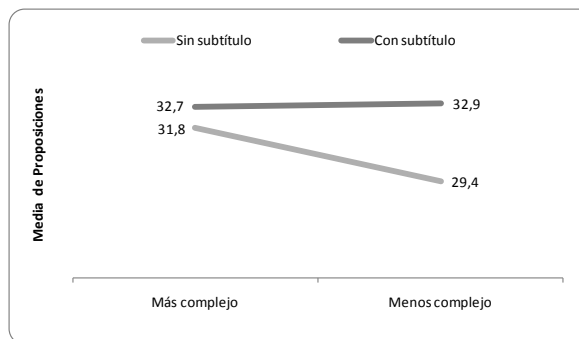
En tanto en el texto más complejo la media de proposiciones en BT es mayor que en MDS; en el menos complejo, las diferencias se presentan a favor de las proposiciones del nivel MDS sobre las correspondientes a BT.

La comparación entre la media de proposiciones del nivel bt en los dos textos revela que se produjeron significativamente más en el texto de mayor complejidad ( $F=7,36$   $p < 0,05$ ).

La diferencia entre las medias de proposiciones correspondientes al nivel mds, superior en el texto menos complejo, es estadísticamente significativa ( $F=7,34$   $p < 0,05$ ).

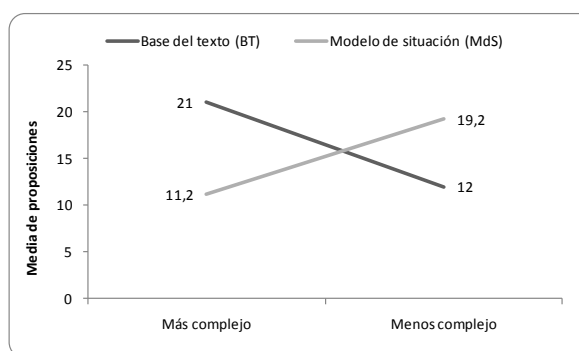
## 3. Base del texto

El análisis de las proposiciones del nivel BT evidencia que en los dos textos se produjeron mayor número de ellas *modificadas* (figura 4). Las diferencias entre proposiciones modificadas y literales es estadísticamente significativa tanto *inter* como *intra* textos (BT,  $F=4,63$   $P < 0,05$ ; MDS,  $F=4,9$   $P < 0,05$ ).



**Figura 2.** Media de proposiciones generadas  $\times$  complejidad  $\times$  subtítulo

Fuente: elaboración propia.



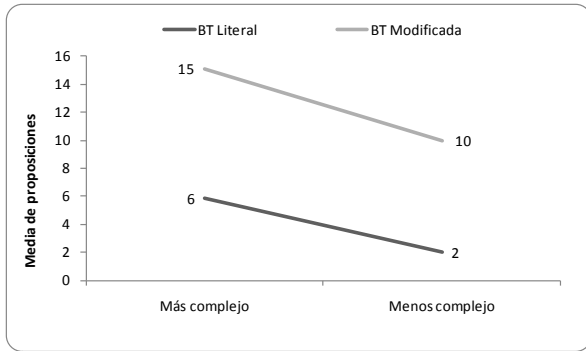
**Figura 3.** Media de proposiciones Nivel de la representación  $\times$  complejidad

Fuente: elaboración propia.

## 4. Modelo de situación

De las categorías establecidas para la información del nivel MDS, no se presentaron ocurrencias para las *inferencias predictivas*, ni para las *evaluaciones* (comentarios, forma y tema). Por tanto el análisis se concentra en las ocurrencias de las categorías *inferencias asociativas* e *inferencias explicativas*.

La figura 5 presenta la media de *inferencias asociativas* y *explicativas* producidas en los dos textos. En los dos casos es mayor la media de *asociativas* que de *explicativas*. Estas diferencias son estadísticamente significativas (*asociativas*,  $F=5,3$   $p < 0,05$ ; *explicativas*,  $F=8,1$   $p < 0,05$ ). Del mismo modo, en el texto *menos complejo* las *inferencias asociativas* como las *explicativas* fueron más numerosas que en el texto *más complejo*.



**Figura 4.** Media de proposiciones del nivel Base del texto

Fuente: elaboración propia.

#### 4.1. Origen de las asociaciones

En tanto las inferencias explicativas fueron las más frecuentes, se analiza su procedencia, es decir, si se fundamentan en el texto mismo y su base del texto, o si surgen del conocimiento previo de los participantes. La figura 6 ilustra cómo predominan las inferencias asociativas cuyo origen es la BT en los dos textos. Las diferencias entre valores de las inferencias provenientes de la BT son estadísticamente significativas ( $F= 4,2 p < 0,05$ ), mientras que las que corresponden al conocimiento previo no presentan diferencias significativas ( $F= 2,08 p > 0,05$ ).

#### 4.2. Consistencia vs inconsistencia de las inferencias asociativas

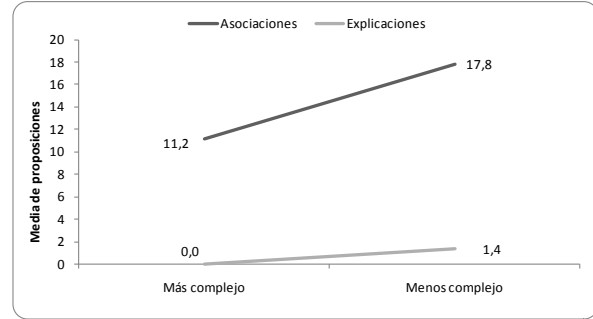
Si bien las diferencias en cuanto a la consistencia en los significados de las inferencias no son estadísticamente significativas, se presentan los resultados específicos en cada texto (figura 7).

*Más complejo.* En este texto las inferencias consistentes son mayores que las inconsistentes, especialmente en las que provienen de la BT.

*Menos complejo.* De igual modo, en este texto, las inferencias consistentes son más frecuentes, siendo las originadas en la BT las más numerosas.

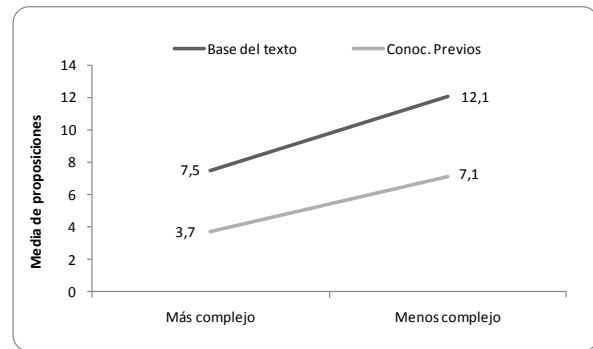
### Discusión

Los resultados del porcentaje de proposiciones escritas por los participantes demuestran que, sin importar la complejidad de los textos, ni la presencia de subtítulos, la cantidad de información recordada



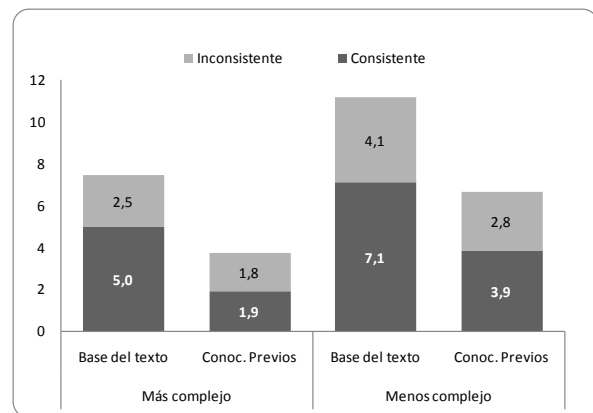
**Figura 5.** Tipos de inferencias x proposiciones

Fuente: elaboración propia.



**Figura 6.** Origen de las inferencias asociativas

Fuente: elaboración propia.



**Figura 7.** Análisis de la consistencia de las inferencias asociativas

Fuente: elaboración propia.

corresponde a un porcentaje cercano al 30%. Estos hallazgos concuerdan con los antecedentes, en los que se encontró que en el caso de estudiantes de secundaria, su recuerdo corresponde a un tercio del

100% de las proposiciones que conforman un texto (Mateus et al., 2012). Sin embargo, en un estudio previo realizado por Kintsch y Van Dijk (1978) con textos narrativos se encontró que el porcentaje de recuerdo inmediato fue de 72%. A este respecto es necesario anotar que las investigaciones sobre comprensión textual han demostrado que los textos narrativos evidencian mejores índices de recuerdo que los expositivos (Langer, 1990; Narváez et al., 1999; Zabrchy y Ratner, 1992).

Por otra parte, en este estudio se observa que la variable *subtítulo* no tiene efecto alguno sobre la recuperación de información y la elaboración de inferencias en la lectura de textos de complejidad diferente. Esto puede explicarse, por un lado, en el hecho de que la tarea de recuerdo libre exige que se genere una representación mental general del texto, a diferencia de los estudios previos donde el efecto del título sobre el recuerdo es examinado a partir de preguntas de información local o con fines de resolver problemas de ambigüedad (Sellán et al., 1995; Migueles y García, 1991). Al respecto, se requiere el diseño y aplicación de una prueba en la que se deba resolver ambas tareas; por un lado, elaborar una representación general del texto y, por otro lado, activar información local.

Además, los estudios previos que mostraban diferencias en función de la presencia o no de títulos son más acentuados en textos experimentales, es decir, en textos que necesitaban la presencia de este paratexto para completar la coherencia. Una posible explicación de que la variable *subtítulo* no haya tenido efecto en este estudio, es decir, que no presentara diferencias estadísticamente significativas en la cantidad de información recuperada de los textos fuente, radica en que las pruebas se realizaron con textos auténticos y no experimentales.

De otra parte, en cuanto al origen de las proposiciones, se encontró que los estudiantes de pregrado activan información de origen distinto (bien sea BT o MDS), según la complejidad del texto leído: en un texto con menor complejidad predomina la representación del modelo de situación (i.e. proposiciones de la base del texto y mayor producción de inferencias),

mientras que en el texto más complejo predomina la base del texto, lo cual parece lógico, pues frente a un texto que integra léxico más frecuente se activa más información proveniente del conocimiento previo, mientras que cuando el texto implica el procesamiento de léxico especializado el recuerdo se concentra en la información proporcionada por el texto mismo. De acuerdo con Mateus et al. (2012), las dificultades que presentan los estudiantes de secundaria a la hora de recuperar información relevante del texto y de diferenciar estas ideas de aquello que es prescindible se asocia a factores como las características de texto (en esta investigación, el léxico), los conocimientos previos del lector, el tipo de tarea e incluso los intereses del individuo.

Hasta aquí, se puede establecer que el principal efecto de la variable de complejidad del texto es: un texto con léxico menos frecuente genera la activación de un mayor número de proposiciones provenientes de BT; además, estas tienden a ser más literales que modificadas, en comparación con los textos de menor complejidad.

En lo que respecta al modelo de situación, el primer resultado que se observa es que sin importar la complejidad del texto, el tipo de inferencia que predomina es la asociación; enseguida se encuentran las explicaciones. No se generaron predicciones ni evaluaciones. Estos hallazgos son consistentes con los datos aportados por Narváez et al. (1999) y Mateus et al. (2012). Adicionalmente, el origen de tales asociaciones parece ser mayormente el conocimiento previo. La causa del mayor número de inferencias asociativas podría residir en la naturaleza de la tarea, pues al haberse solicitado escribir el texto de la manera más fiel al original, los estudiantes pudieron añadir información relacionada con los referentes de texto en forma de contexto de lo leído sin “atreverse” a generar lazos explicativos entre las informaciones leídas y activadas.

Por otro lado, desde la perspectiva del lector, es decir, basados en las características de la población, se encontró que los estudiantes de pregrado elaboran proposiciones cuya representación corresponde a información consistente (cerca del 66%), tanto con

la base del texto como con el conocimiento previo, en textos de distinta complejidad. Lo cual apunta a un proceso de comprensión de lectura más cualificado, diferente al reportado por Mateus et al. (2012), en el cual se encontró que los estudiantes de secundaria presentan dificultades en el proceso de lectura, pues integran un porcentaje considerable de información inconsistente (cerca al 50%) con la base de texto y el conocimiento previo. Es de notar que la población del presente estudio corresponde a un grupo de estudiantes universitarios en ciencias humanas, en donde parte de su formación específica se da a través de lecturas especializadas.

En lo que respecta al efecto de la frecuencia léxica, gran parte de las investigaciones han sido realizadas en experimentos con tareas en las que se requiere el procesamiento de palabras aisladas, de emparejamiento o tareas de decisión léxica en oraciones simples (Arya, Hiebert y Pearson, 2011; Fajardo, Hernández y González, 2012). En este caso, se ha evidenciado que la frecuencia léxica influye en niveles de procesamiento lingüístico superior, como es el procesamiento de textos auténticos compuestos por distintos párrafos de diversa extensión. Estos hallazgos tendrían implicaciones en la didáctica de la comprensión lectora, específicamente en lo relacionado con la selección, incluso con la edición de textos, por ejemplo, explicitar las relaciones internas entre las ideas que integran los textos (Vidal-Abarca y Gilabert, 1990); sería preciso añadir un análisis concienzudo del léxico y las estrategias adecuadas para abordarlo según se trate de léxico frecuente o de otro más especializado. Sumado a la variable frecuencia léxica, podrían llevarse a cabo otras investigaciones que tuvieran en cuenta las redes semánticas de las cuales hacen parte las unidades léxicas seleccionadas en los textos (Henríquez, Mahecha y Mateus, 2016; Mahecha y Mateus, 2017; Mateus, Castiblanco y Álvarez, 2018) y así, poder examinar otra vía de incidencia del léxico en la comprensión.

## Referencias

- Adam, J. y Revaz, F. (1996). (Proto)tipos: La estructura de la composición en los textos. *Textos de Didáctica de la Lengua y la Literatura*, 10, 9-22.
- Adam, J. M. (1997). Los textos: heterogeneidad y complejidad. En F. J. Cantero, A. Mendoza y C. Romea (eds.). *Didáctica de la lengua y la literatura para una sociedad plurilingüe del siglo xxi* (pp. 3-12). Barcelona: Sociedad Española de Didáctica de la Lengua y la Literatura, Universitat de Barcelona.
- Arya, D., Hiebert, E. y Pearson, P. (2011). The Effects of Syntactic and Lexical Complexity on the Comprehension of Elementary Science Texts. *International Electronic Journal of Elementary Education*, 4, 107-125.
- Bernardez, E. (1995). *Teoría y epistemología del texto*. Madrid: Cátedra.
- Bransford, J. y Johnson, M. (1972). Contextual Prerequisites for Understanding: Some Investigations for Comprehension and Recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717-726.
- Cutler, A. (1983). Lexical Complexity and Sentence Processing. In G. B. Flores d'Arcais y R. J. Jarvella (eds.), *The Processes of Language Understanding* (pp. 43-79). New York: Wiley.
- Dooling, D. J., y Lachman, R. (1971). Effects of Comprehension on Retention of Prose. *Journal of Experimental Psychology*, 88(2), 216-222.
- Fajardo, A., Hernández, J. y González, A. (2012). Acceso léxico y comprensión lectora: un estudio con jóvenes universitarios. *Revista Electrónica de Investigación Educativa*, 14(2), 25-33. Recuperado de <http://redie.uabc.mx/vol14no2/contenido fajardoetal.htm>
- Fincher-Kiefer, R. (1994). The Effect of Inferential Processes on Perceptual Identification. *Discourse Processes*, 18, 1-17.
- González, B., Salazar, A. y Peña, L. (2015). *Formación inicial en lectura y escritura en la universidad*. Bogotá: Pontificia Universidad Javeriana.
- Henríquez, M., Mahecha S. V. y Mateus, G. E. (2016). Análisis de los mecanismos cognitivos del léxico disponible de cuerpo humano a través de grafos. *Lingüística y Literatura*, 69, p. 229-251.
- ICFES (2017). Informe Nacional de resultados Examen Saber 11 2016. Recuperado de [www.icfes.gov.co](http://www.icfes.gov.co)
- ICFES (2017). Informe Nacional de resultados Examen Saber Pro 2016. Recuperado de [www.icfes.gov.co](http://www.icfes.gov.co)

- Just, M. y Carpenter, P. (1987). *The Psychology of Reading and Language Comprehension*. Newton, MA: Allyn and Bacon.
- Keefe, D. y McDaniel, M. (1993). The Time Course and Durability of Predictive Inferences. *Journal of Memory and Language*, 32, 446-463.
- Kintsch, W. y van Dijk, T. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review*, 85(5), 363-394.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Kintsch, W., Welsch, D., Schmalhofer, F., y Zimny, S. (1990). Sentence Memory: A Theoretical Analysis. *Journal of Memory and Language*, 29(2), 133-159.
- Klare, G. R. (1968). The Role of Word Frequency in Readability. In J. R. Bormuth (ed.), *Readability* (pp. 7-17), National Conference on Research in English, New York.
- Langer, J. (1990). The Processes of Understanding: Reading for Literary and Informative Purposes. *Research in the Teaching of English*, 24, 229-260.
- León, J. A. (coord.) (2003). *Conocimiento y discurso. Claves para inferir y comprender*. Madrid: Ediciones Pirámide.
- León, J., Escudero, I., y van den Broek, P. (2003). La influencia del género del texto en el establecimiento de inferencias elaborativas. En José A. León (coord.), *Conocimiento y discurso. Claves para inferir y comprender* (pp. 153-170). Madrid: Ediciones Pirámide.
- Mahecha, S. V. y Mateus, G. E. (2017). El léxico disponible y sus mecanismos de asociación: un análisis con grafos. En F. de la Rosa (ed), *Palabras Vocabulario Léxico. La lexicología aplicada a la didáctica y a la sincronía* (pp. 123-142). Italia: Edizioni Ca' Foscari.
- Mateus, G. E., Castiblanco, L. y Álvarez, P. (2018). Mecanismos lógicos y analógicos en la producción del léxico disponible. *Folios*, 47, p.133-152.
- Mateus, G., Castillo, M., Santiago A. W. y Rodríguez, L. (2012). Lectura y representación mental de textos expositivos en estudiantes de educación media. *Folios*, (36), 115-130.
- McKoon, G. y Ratcliff, R. (1995). The Minimalist Hypothesis: Directions for Research. In C. A. Weaver, S. Mannes y C. Fletcher (eds.) *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 97-116). Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Miguelés, M. y García, E. (1991). Influencia del recuerdo inmediato y del contexto en el procesamiento de un texto ambiguo. *Revista de Psicología General y Aplicada*, (44), 215-219.
- Narváez, D., Van den Broek, P. y Ruiz, A. (1999). The Influence of Reading Purpose on Inference Generation and Comprehension in Reading. *Journal of Educational Psychology*, 91(3), 488-496.
- Rayner, K. y Duffy, S. (1986). Lexical Complexity and Fixation Times in Reading: Effects of Word Frequency, Verb Complexity, and Lexical Ambiguity. *Memory y Cognition*. 14(3), 191-201.
- Real Academia Española. Banco de datos (CREA) (octubre de 2018). *Corpus de referencia del español actual*. Recuperado de <http://www.rae.es>
- Santiago A., Castillo, M. y Mateus, G. (2014). *Didáctica de la lectura: una propuesta sustentada en metacognición*. Bogotá: Alejandría libros.
- Sellán, C., Sanz, M. y Valle, F. (1995). Influencia de un título pospuesto en la comprensión de textos. *Revista de Psicología General y Aplicada*, (48), 283-299.
- Trabasso, T. y Magliano, J. P. (1996). Conscious Understanding During Comprehension. *Discourse Processes*, 21, 255-287.
- Van den Broek, P. (1994). Comprehension and Memory of Narrative Texts: Inferences and Coherence. In M. A. Gernsbacher (ed.), *Handbook of psycholinguistics* (pp. 539-588). San Diego, US: Academic Press.
- Van den Broek, P., Fletcher, C. y Risden, K. (1993). Investigations of Inferential Processes in Reading: A Theoretical and Methodological Integration. *Discourse Processes*, 16, 169-180.
- Van Dijk, T. (1980). *Macrostructures*. New Jersey: Lawrence Erlbaum Associates.
- Van Dijk, T. y Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.
- Vidal-Abarca, E. y Gilabert, R. (1990). Comprender para aprender, un programa para mejorar la comprensión y el aprendizaje de textos informativos. *Comunicación, lenguaje y educación*, 6, 113-124.
- Zabrocky, K. y Ratner, H. (1992). Effects of Passage Type on Comprehension, Monitoring and Recall in Good and Poor Readers. *Journal of Reading Behavior*, 24, 373-391.