

## UNA MIRADA A LOS MODELOS MULTIMODALES DE COMPREENSIÓN Y APRENDIZAJE A PARTIR DEL TEXTO\*

Jadranka Gladic Miralles\*\*  
Violeta Cautín-Epifani\*\*\*

### Resumen

El presente artículo tiene como objetivo realizar una revisión del enfoque cognitivo respecto de la comprensión de textos multimodales. Partimos de la premisa de que todos los textos son por naturaleza multimodales, ya que la construcción de significado se elabora por medio de la conjunción de diversos sistemas semióticos. Por esto, centramos nuestra primera mirada a la visión del texto desde la semiótica. Una vez establecido el concepto de texto como entidad multimodal, se realiza un recorrido por los diferentes modelos que abordan el fenómeno de la comprensión y aprendizaje a partir del texto desde esta perspectiva. Concluimos nuestra revisión con una reflexión en relación con los aportes y desafíos que presenta esta visión para futuras investigaciones en el área.

**Palabras Clave:** Multimodalidad, sistemas semióticos, texto multimodal, comprensión discursiva, Modelos De Comprensión Multimodal.

## AN OVERVIEW OF MULTIMODAL MODELS OF COMPREHENSION AND LEARNING FROM TEXT

### Abstract

The objective of this article is to provide an overview of the cognitive approach to text comprehension emphasizing multimodal models. We start from the premise texts are inherently multimodal due to the fact that in them the construction of meaning is performed through the combination of various semiotic systems. Consequently, we start our review looking at the concept of multimodal text from a semiotic perspective to then, revise the most important models that approach comprehension and learning from text from this perspective. We conclude our revision with a reflection regarding the contributions and challenges that this vision presents for future research in the area.

**Keywords:** Multimodality, semiotic systems, multimodal text, discourse comprehension, multimodal comprehension models.

Recibido: 05-10-2015

Aceptado: 16-06-2016

---

\* Este artículo deriva de la investigación realizada en el contexto de la tesis de Magíster *Niveles de comprensión y su relación con la predominancia de sistemas semióticos en textos del área de lingüística: Una aproximación al fenómeno multimodal desde el discurso académico universitario* (Gladic, 2012), enmarcada en el proyecto FONDECYT N°1090300 y posible gracias al apoyo del Programa de Formación de Capital Humano de CONICYT 22110119.

\*\* Chilena. Magíster en Lingüística Aplicada. Profesora de Castellano y Filosofía. Becaria PROGRAMA CONICYT 22110119. Docente del Programa de Lectura y Escritura Académica (PLEA) y Programa de Escritura Disciplinar (PED), Pontificia Universidad Católica de Chile, Santiago, Chile. jagladic@uc.cl

\*\*\* Chilena, Doctora en Lingüística. Profesora de Estado en Inglés, mención Traducción e Interpretariado Consecutivo. Becaria PROGRAMA CONICYT 22110119. Pontificia Universidad Católica de Valparaíso, Chile. vicautin@uc.cl

## 1. Introducción

Es indudable la influencia que han tenido los adelantos tecnológicos en los diferentes ámbitos de la sociedad, y la educación no es la excepción. Diversos investigadores coinciden en señalar que estas tecnologías han contribuido a la creciente incorporación de material educativo integrado por diversos *sistemas semióticos*<sup>1</sup>, lo que, por un lado, pone en tensión la anterior hegemonía del sistema verbal como preeminente para una alfabetización efectiva (Bezemer & Kress, 2010; Manghi, 2011; Kress & Seleander, 2012; Unsworth, 2010; Manghi, 2013; Manghi, González, Echeverría, Marín, Rodríguez & Guajardo, 2013; Unsworth & Thomas, 2014; Sabich, 2015; Soleas, 2015) y, por otro lado, destaca la importancia de la integración de estos sistemas en la interacción de aula y la comunicación en general (Manghi, 2010; Manghi & Córdova, 2011; Manghi, Arancibia, Illanes, Herrera, Zamora, 2014; Manghi, Badillo y Villacura, 2014).

Paralelamente, desde los estudios en comunicación mediada por computadora, cada vez se encuentran más investigaciones que describen nuevas formas de interacción, en especial entre jóvenes, las que surgen a raíz de la adopción de diversas plataformas de comunicación, también marcadas por una constante integración de distintos sistemas de construcción de significado (Jiménez & Vela Delfa, 2005; Noblía, 2007; Cantamutto, 2012; Cautín-Epifani, 2015a; 2015b; Giammatteo & Albano, 2013; Bourlai & Herring, 2014, Cantamutto, 2015; Herring, 2015; Vela Delfa & Cantamutto, 2015).

En este escenario, se hace insoslayable contar con constructos teóricos capaces de integrar estos aspectos en las explicaciones de los procesos cognitivos de comprensión y aprendizaje a partir de textos; aún más, teniendo en cuenta que estas configuraciones de construcción de significado requieren que los comprendedores dominen formas y estrategias particulares de lectura (Holmqvist, Holsanova, Barthelson & Lundquist, 2003; Holsanova, 2004; Hochpöchler, Schnotz, Rasch, Ullrich, Horz, McElvany & Baumert, 2013; Manghi et al, 2013; Schnotz, Ludwig, Ullrich, Horz, McElvany & Baumert, 2014). El estudio de la comprensión de textos ha sido una constante en diversas disciplinas

---

1 Se entenderá por sistema semiótico a los sistemas de significado que interactúan entre sí con el fin de elaborar y comunicar dentro de un contexto determinado Según Parodi (2010) en el texto escrito estos son el verbal, el gráfico, el matemático y el tipográfico.

del conocimiento; sin embargo, sólo en las últimas décadas, y debido a un cambio de paradigma hacia el estudio de las ciencias cognitivas, es que ésta comienza a centrarse en los procesos internos que realiza un individuo cuando se enfrenta a un texto. Es por eso que en el presente artículo se realiza una revisión de los diferentes modelos cognitivos de comprensión de textos que abordan explícitamente los componentes de integración de distintos sistemas semióticos en el procesamiento textual. En el primer apartado, damos cuenta de los fundamentos que sustentan la visión del texto como una entidad inherentemente multimodal. Luego, se realiza un recorrido por los diferentes modelos de comprensión del discurso escrito que abordan este fenómeno. Concluimos con una reflexión en torno a los aportes y desafíos que presenta esta visión para futuras investigaciones en el área.

## 2. El texto como construcción multimodal

Aun cuando la supremacía de la visión del texto puramente verbal ha definido la manera en que accedemos al conocimiento y la cultura, no es posible desconocer la relación que el sistema semiótico verbal ha mantenido con otros sistemas a lo largo de la historia de la humanidad<sup>2</sup>. Sin embargo, es sólo hasta la época actual en la que esta relación entre los diferentes sistemas semióticos parece haberse estrechado más que en los períodos anteriores. Debido a la expansión de los medios de comunicación de masas, la aparición de variadas tecnologías y el auge de la publicidad, es que los investigadores de distintas disciplinas del conocimiento como la psicología, la sociología, la lingüística y la semiótica ponen su atención e interés en la relación que se establece entre el sistema verbal y los demás sistemas semióticos (Hernández, 2001). De esta última disciplina surge la multimodalidad, área de investigación que se ocupa de la relación entre diferentes sistemas semióticos y las implicancias que esta relación produce en los individuos al momento de enfrentarse a los textos.

---

2 Para una revisión histórica de los estudios del discurso en relación con sus diferentes sistemas semióticos, ver Hernández (2001) y Kalternbacher (2007). Este último plantea que el surgimiento de los estudios sobre multimodalidad dataría de fines del siglo XVIII.

## 2.1. Concepto de multimodalidad

La multimodalidad, línea de investigación que desarrolla sus estudios a partir de las últimas décadas del siglo XX<sup>3</sup>, surge en un contexto de grandes cambios sociales y culturales que caracterizan este período. El desarrollo incipiente de las tecnologías y de los medios masivos de comunicación en el siglo pasado significó un avance en la manera en que las personas se comunican y una renovación en la difusión del conocimiento. La oralidad y la escritura parecen insuficientes para reportar la información y el conocimiento, por lo que se hace necesario un giro hacia lo visual. Esta nueva área de investigación fundamenta su línea de estudio en la Lingüística Sistémico-Funcional (Halliday, 1985)<sup>4</sup>. Las opciones que el lenguaje escrito y oral proporciona a los individuos corresponden a recursos semánticos, léxico-gramaticales y fonológicos-fonéticos, los que son seleccionados por los sujetos en relación a los factores relevantes de la situación comunicativa específica a la que se ven enfrentados.

Posteriormente, Hodge y Kress (1988) reformulan el concepto de semiótica social y postulan que la construcción de significado no sólo se limita al modo lingüístico. Estos autores analizan cómo la cultura -y en especial la ideología- se proyecta a través de distintos sistemas semióticos como las imágenes, los gestos, la mirada, la postura, entre otros que están a disposición de los sujetos y que les permiten construir significados y comunicar en contextos sociales y culturales específicos. Siguiendo los planteamientos de Kress y van Leeuwen (2001) se entiende por multimodalidad el uso de variados sistemas semióticos para la construcción de significado y el diseño de productos o eventos semióticos en contextos específicos.

Esta definición amplia de multimodalidad abarca una serie de dominios, procesos y productos. Parodi (2010) delimita el concepto

---

3 Desde la psicología ya se había estudiado la relación de la imagería con la construcción de conocimiento; sin embargo, no fue sino hasta que Paivio (1971; 1986) presentó sus postulados de la teoría de código dual que no se consideró esta área de estudio de manera seria en la corriente principal de esta disciplina (Paivio, 1991). Ahora bien, en este apartado consideramos los estudios de multimodalidad desde su vertiente semiótica y discursiva.

4 Corriente lingüística funcionalista que concibe el lenguaje como un sistema semiótico que los individuos utilizan para comunicar los significados sociales y culturales mediante múltiples opciones que se ajustan a diferentes situaciones comunicativas (ver Halliday, 1985).

de texto escrito multimodal<sup>5</sup> mediante la conceptualización de cuatro sistemas semióticos que se complementan e interactúan para construir significados. Estos sistemas semióticos son: el sistema verbal, el sistema gráfico, el sistema matemático y el sistema tipográfico. El sistema verbal está conformado por palabras, frases y oraciones que constituyen significado por medio de lo lingüístico; el sistema gráfico se compone de representaciones pictóricas de diversa índole como, por ejemplo, fotografías y espacios en blanco; el sistema matemático está conformado por grafías, signos o representaciones que permiten la codificación simbólica de la información; y el sistema tipográfico porta significado al texto por medio de su forma, dimensionalidad, color, tamaño de las letras, entre otros recursos. Cada uno de estos sistemas semióticos interactúa con los demás de diversas maneras con el fin de construir significado.

Igualmente, los sistemas semióticos pueden componer algunos recursos o unidades denominados *artefactos multisemióticos*<sup>6</sup>, cuya presencia permite la comunicación efectiva del conocimiento. Parodi (2010), en un estudio de corpus de textos disciplinares, identifica y define operacionalmente nueve de ellos entre los que se encuentran: esquema, figura geométrica, fórmula, gráfico estadístico, ícono, ilustración, mapa, red composicional y tabla<sup>7</sup>.

Ahora bien, el discurso cotidiano producido a través de medios virtuales, igualmente, emplea estos recursos y los ha integrado como propios de su actividad (Bourlari & Herring, 2014; Herring, 2015). Por un lado, la *recursividad*<sup>8</sup> o *multisimultaneidad*<sup>9</sup> de las interacciones en estos medios hace de estas una relación permanente de comunicación que se

5 Para este trabajo tomaremos la definición de texto que proporciona Gómez (1988) quien lo entiende como una creación humana convencional y simbólica. Este autor postula que en el texto todo está orientado a significar, siendo su estructura física destinada a ser utilizada como material significativo construido por quien otorga sentido al texto.

6 Son artefactos multisemióticos tablas, gráficos, mapas, esquema, entre otros (Parodi, 2010).

7 Las definiciones presentadas por Parodi (2010) fueron elaboradas mediante tres criterios orientadores: criterio de modalidad, que corresponde a los sistemas que componen el artefacto; criterio de función, que determina el uso que se da a cada uno de los artefactos; y el criterio de composición, que indica la constitución del artefacto.

8 Gobato (2014) y, especialmente, Cantamutto (2013) describen el fenómeno de la *recursividad* como la capacidad que se presenta en las interacciones en medios virtuales de mantener una conversación a través de distintas plataformas (teléfono, chat, redes sociales, etc.). Vela Delfa & Cantamutto (2015, p. 496) denominan este fenómeno como *multisimultaneidad* y lo definen de la siguiente manera: "Multisimultaneity refers to the maintenance of co-occurring interactions, i.e., the fact that speakers are involved simultaneously in several parallel exchanges. This is an intrinsic condition of the digital medium, enabled by the confluence of several factors: 1) textual persistence, 2) the deferred nature of the exchanges, and 3) the lack of physical co-presence".

9 También llamada escritura Ideofonemática (Torres i Villartasana, 2003).

desarrolla en distintas plataformas (Alcántara-Pla, 2014) y con distintos grados de integración multimodal. Adicionalmente, podemos encontrar en la escritura en medios virtuales el empleo de variados recursos que los usuarios adaptan para conseguir sus objetivos comunicacionales, por ejemplo, uso creativo de la grafía que puede combinar el sistema semiótico verbal y matemático de modo de crear artefactos similares a fórmulas algebraicas (4ever, 2U, re100), uso de íconos, conocidos como emoticones, que confieren expresividad al discurso en este medio, abundante uso de fotografías y videos que se entremezclan con las interacciones escritas, entre otros.

### 3. Modelos multimodales de comprensión del texto

Considerando la complejidad que ha adquirido la comprensión del discurso escrito en estos nuevos escenarios, se hace necesario describir la forma en que los sujetos comprenden y aprenden en modelos que recogen esta problemática. Este cuestionamiento no ha sido ajeno para la psicología discursiva, por lo que, a continuación, realizaremos una revisión de algunos de los modelos cognitivos más importantes que tratan de dar un sustento explicativo a la comprensión de textos multimodales.

Cabe señalar que, tal como fundamentamos en la primera parte de este artículo, concebimos la definición de texto como una entidad inherentemente multisemiótica, por lo que desde esta perspectiva no cabría realizar una distinción entre texto e imagen, sino más bien entre sistema verbal y otros sistemas semióticos. Sin embargo, desde la tradición cognitiva de la investigación multimodal, sí se trabaja con los conceptos de *texto* e *imagen*. En lo que sigue recogeremos esta tradición terminológica entendiendo que esta distinción es una construcción teórica que permite operacionalizar los conceptos que desde esta visión se trabajan. Ello, debido a que los modelos distinguen entre los procesos de comprensión de elementos verbales (sistema semiótico verbal), lo que denominan *texto*, y los procesos de comprensión no verbal (sistema semiótico gráfico), lo que generalmente conceptualizan como *imágenes*.

#### 3.1. Teoría del Código Dual

La Teoría del Código Dual (TCD) propuesta por Paivio (1971; 1986) descansa en el supuesto de que las representaciones mentales son de modalidad específica debido a que ellas “retain some of the concrete

qualities of the external experiences from which they derive” (Sadoski & Paivio, 2004, p. 1331). En este sentido, se oponen a la conceptualización de representaciones mentales basadas en constructos abstractos como proposiciones o esquemas. Debido a esta concepción, algunos expertos postulan esta teoría como la precursora del giro hacia la cognición corpórea (Dove, 2010)<sup>10</sup>.

Para estos investigadores, la lectura y la escritura implican la actividad de dos subsistemas cognitivos de codificación independientes. Un subsistema de codificación verbal que captura y procesa la información lingüística y un subsistema de codificación no verbal que procesa y analiza la información no verbal o referente a eventos. Esta información es representada y almacenada de manera interna por el sujeto, manteniendo las características de la modalidad en la que se percibió.

Esta teoría postula que cada uno de estos códigos detentaría sus propias unidades básicas, las que corresponden a estructuras y vías neuronales que poseen los sujetos. Estas estructuras no serían homogéneas, ya que además de ser de modalidades distintas, sus tamaños también diferirían. Por un lado, las unidades básicas del subsistema verbal, denominadas *logogens*, pueden organizarse en *logogens visuales* más pequeñas para letras y mayores para palabras o frases; también pueden organizarse en *logogens auditivas* pequeñas para fonemas y mayores para pronunciaciones de palabras o frases, y así sucesivamente con las diferentes modalidades. En este sentido, las *logogens* poseen una organización jerárquica secuencial que deriva de la naturaleza serial del lenguaje verbal y la forma en que los seres humanos las percibimos.

Por su parte, las *imagens* asumen una jerarquía sincrónica, ya que la información se encuentra disponible de manera simultánea. Sin embargo, esto no implica que los sujetos puedan acceder a ella y puedan procesarla de una sola vez. La diferencia entre ambas unidades cognitivas no solo se debe a los estímulos que las originan, sino que también a la forma en que estas se procesan (Sadoski & Paivio, 2001; Sadoski & Paivio, 2004; Sadoski, 2015).

10 Se entiende por cognición corpórea: “las teorías corpóreas proponen, como hemos visto, que el significado se basa en simulaciones modales, corporeizadas en el cerebro humano, donde la experiencia juega un rol central en la representación del conocimiento.” (Urrutia & De Vega, 2012)

Esta teoría propone una estructura de conexiones en y entre los subsistemas de codificación, y tres niveles de procesamiento de la información. En una primera instancia, se presentan los estímulos verbales y visuales a los que están expuestos los sistemas sensoriales de los lectores. Las conexiones entre estas representaciones sensoriales y la activación de las logogens y las imágenes corresponden al primer nivel de procesamiento de la información: el procesamiento representacional.

Cada uno de los subsistemas organiza la información jerárquicamente. El subsistema verbal organiza la información de manera secuencial, mientras que el subsistema visual la organiza de manera sincrónica. La conexión que se genera entre las unidades de los subsistemas, es decir, la relación entre *logogens* e *imágenes* y viceversa, se lleva a cabo a través del nivel de procesamiento referencial y la actividad que se produce entre las unidades de un mismo subsistema se realiza por medio del procesamiento asociativo (Sadoski & Paivio, 2001).

Los subsistemas son capaces de funcionar de tres maneras: de manera independiente cuando uno de ellos se activa y el otro se mantiene pasivo; en paralelo cuando ambos subsistemas se activan separadamente al mismo tiempo; o de manera conectada cuando existen conexiones entre las unidades del subsistema verbal y visual. Las relaciones que se establecen en y entre los subsistemas no sólo dependen de los estímulos externos a los que los sujetos han sido expuestos, sino que también de la activación de conocimientos previos y del número de conexiones que se logran establecer entre los subsistemas.

### 3.2. Modelo de aprendizaje multimedia

El modelo propuesto por Mayer (1997) y reformulado en publicaciones posteriores (Mayer & Moreno 2002; Mayer 2005; Mayer, 2009; Mayer, 2014) se origina a partir de las investigaciones que estos expertos realizan en torno al aprendizaje de lo que ellos denominan mensajes instruccionales multimedia. Estos corresponden a la presentación simultánea de información destinada a fomentar el aprendizaje, que combinan diferentes modos semióticos, a través de distintos canales (auditivo, visual) y que, por lo tanto, requiere de distintas modalidades sensoriales de percepción. Tal como comentan Farias, Obilinovic & Orrego (2011), Mayer y colaboradores utilizan el concepto de multimedia en tres sentidos distintos:



as a delivery media (combining two or more delivery devices, as a PowerPoint presentation and the lecturer's voice), presentation modes (representations that include words and pictures, as on-screen text and animation) and sensory modalities (visual and auditory senses, as used to process slides and narration, for example). (p. 137).

El objetivo que persiguen Mayer y colaboradores es investigar en profundidad la extensión y limitaciones del *principio de multimedia*, que postula que los estudiantes aprenden mejor cuando son enfrentados a textos que combinan los sistemas semióticos verbal y gráfico. La consideración de estos sistemas permitiría la creación de un modelo que da cuenta de los procesos cognitivos implicados en el aprendizaje a partir de *mensajes multimediales* (Mayer, 2005). Además, estos investigadores pretenden desarrollar principios que orienten el diseño instruccional de los materiales de aprendizaje.

El modelo que Mayer y colaboradores propone integrar postulados de la teoría de la carga cognitiva (Chandler & Sweller, 1991; Sweller, 1999) que afirma que la capacidad de procesamiento de la memoria de trabajo de los sujetos es limitada, por lo que la comprensión y aprendizaje serían mayores si se evita la sobrecarga de ésta. Además, integra principios de la teoría constructivista del aprendizaje, pues postula que el aprendizaje significativo se produce cuando los sujetos seleccionan activamente la información relevante de los textos, organizándola en representaciones coherentes e integrándola con sus conocimientos previos.

Este modelo, que se basa en la TCD, propone dos sistemas cognitivos de codificación de la información: un sistema verbal, que se encargaría del procesamiento de la información que ingresa al sistema en forma de audio, y un sistema visual que se encargaría del procesamiento de la información no verbal y de la información verbal que ingresa al sistema a través de la visión. Cabe destacar que estos investigadores postulan que "in working memory printed words can be converted into sounds and images can be converted into spoken words" (Mayer, 2014).

Asimismo, proponen la existencia de tres etapas o momentos en los que se realiza el proceso de comprensión. En la primera etapa, el aprendiz selecciona en la memoria a corto plazo la información relevante tanto del sistema verbal como del visual. En el sistema verbal el sujeto selecciona

las palabras más importantes del texto, construyendo un modelo de representación proposicional, y en el sistema visual los aprendices seleccionan imágenes que permiten la construcción de un modelo de representación visual. Una vez que el aprendiz ha construido ambos modelos mentales, la siguiente etapa corresponde a la organización del material seleccionado de una manera coherente. En esta etapa el sujeto utiliza mecanismos de asociación entre los elementos de cada uno de los sistemas. Una vez que el aprendiz ha construido un modelo basado en la información verbal y uno en la información visual, debe ser capaz de realizar conexiones entre ambas representaciones. Este proceso, denominado por Mayer (1997) integración, describe la conexión entre la información organizada en las etapas anteriores y los conocimientos previos que se encuentran ya almacenados en la memoria del aprendiz. Mayer y Moreno (2002) reformulan este modelo y proponen que estos procesos pueden generarse de manera alternada e interactivamente. En última instancia, estos autores postulan que la comprensión y el aprendizaje ocurrirían cuando los sujetos integran la información de ambos modelos representacionales en un modelo representacional coherente de la situación descrita en el texto. En este sentido, los autores se alinean con la tradición de la psicología discursiva que plantea la preponderancia de la construcción del modelo de situación para alcanzar una comprensión efectiva de lo leído (Cautín-Epifani, 2014; Martínez, 2015).

A partir de estos postulados teóricos, Mayer y sus colaboradores han realizado investigaciones empíricas sobre la efectividad de los *mensajes multimedia* en el aprendizaje. Por medio de una serie de experimentos, los autores proponen algunas dimensiones que permiten el mejor diseño instruccional de textos de aprendizaje multimedial. Reseñamos, a continuación, las cuatro dimensiones más importantes<sup>11</sup>: contigüidad, coherencia, modalidad y redundancia (Mayer & Moreno, 2002; Mayer, 2005). La dimensión de contigüidad se cumple cuando lo verbal y lo visual se coordinan en el tiempo, lo que permite que los aprendices tengan más probabilidades de mantener estas representaciones en la memoria de trabajo y, por lo tanto, de realizar las conexiones correspondientes entre

---

11 Por razones de espacio, sólo nos remitiremos a las dimensiones más importantes. Para una revisión acabada de las diferentes dimensiones postuladas por Mayer y colaboradores se sugiere ver Mayer (2005; 2014) y Sorden (2013).

ambas. Mayer y Moreno (2002) postulan que la dimensión de coherencia se cumple cuando los textos multimediales exponen información fundamental y excluyen datos superfluos. Esto se debe a que los textos, que poseen un exceso de información, sobrecargan la memoria de trabajo de los aprendices e impiden que puedan realizar las conexiones necesarias entre las representaciones verbales y visuales. Las dimensiones de modalidad y redundancia, según Mayer y Moreno (2002) se cumplen cuando la información verbal y visual es presentada a los sujetos a través de distintos modos sensoriales. Mediante los experimentos llevados a cabo por los autores, se comprueba que los aprendices que observan animaciones y escuchan narraciones de manera simultánea presentan mejores resultados en pruebas de transferencia que los sujetos que, paralelamente, observan animaciones, escuchan texto y leen subtítulos en el mismo idioma que la narración. Esto se debería a que, cuando la información se presenta a partir del mismo modo sensorial, por ejemplo, a través de la vista, la memoria de trabajo se sobrecarga, la atención se divide, la energía cognitiva disminuye y las conexiones que se pueden establecer entre las representaciones verbales y visuales se reducen.

Otro de los postulados que surgen de la investigación empírica de estos investigadores dice relación con las condiciones limitantes en las que las dimensiones recién mencionadas serían más efectivas. En este sentido los investigadores postulan que los principios del modelo de aprendizaje multimedia serían más efectivos en sujetos con menor conocimiento del tema, cuando el tema es complejo y el material se presenta de manera más rápida (Mayer, 2009). Es necesario destacar que Mayer (2005; 2009; 2014) enfatiza que este modelo está en evolución, por lo que se espera que algunos de sus postulados se vayan modificando a medida que avanza la investigación en el área.

### 3.3. Modelo integrador de la comprensión de textos e imágenes

Schnotz y Bannert (2003) y Schnotz (2005) consideran las implicancias que los textos multimodales poseen sobre el aprendizaje de los estudiantes, las condiciones en las que el aprendizaje multimodal es efectivo y las razones de su efectividad en situaciones específicas de aprendizaje. Para ello, Schnotz y Bannert (2003) proponen un modelo de integración de la comprensión de textos e imágenes, que Schnotz

(2005) reformula considerando algunos aspectos no abordados en el primer modelo.

En la primera versión del modelo se propone que el texto y las imágenes corresponden a representaciones externas que son comprendidas por los individuos una vez que han construido representaciones mentales internas de éstas. Los investigadores proponen que el procesamiento de la información verbal, que ellos denominan texto, debido a su carácter simbólico, son representados descriptivamente; las imágenes, en cambio, debido a su naturaleza icónica, son representadas depictivamente. Así, el modelo de Schnotz y Bannert (2003) posee una rama *descriptiva* y una rama *depictiva* de procesamiento de la información.

En la rama *descriptiva*, el lector construye una representación mental de la estructura de la superficie textual, genera una representación proposicional del contenido semántico del texto, es decir, una base textual y, finalmente, construye, a partir de la base del texto, un modelo mental de la situación descrita (van Dijk & Kintsch, 1983). Para ello, estos investigadores postulan que los sujetos emplean procesos ascendentes y descendentes de activación de esquemas cognitivos, los que poseen una función selectiva y organizadora de la información pertinente a la tarea de lectura.

Por su parte, la rama *depictiva* comprende la representación externa de una imagen construida por medio de un procesamiento perceptual, una representación interna de la imagen y, posteriormente, se construye un modelo mental, así como una representación proposicional del contenido de la imagen por medio del procesamiento semántico. La construcción del modelo mental de una imagen se realiza a través de un proceso de *mapeo*, guiado por los esquemas mentales que posee el sujeto. Este *mapeo* es considerado como un proceso analógico entre un sistema viso-espacial y un sistema de relaciones semánticas, el que puede ser llevado a cabo de manera ascendente o descendente según sea el tipo de imagen a la que se enfrenta el sujeto (Schnotz & Bannert, 2003).

En esta misma línea, Schnotz (2005) reelabora el modelo y distingue un nivel perceptuales y uno cognitivo. El nivel perceptual se caracteriza por el funcionamiento de diferentes canales sensoriales, el auditivo y el visual, que permiten la transferencia de la información externa verbal y visual a la memoria de trabajo (Baddeley, 1999, 2000). La comprensión

de textos, por tanto, comienza cuando la información verbal o visual ingresa a través de los ojos o los oídos y se reenvía a la memoria de trabajo auditivo y visual. Estos canales sensoriales tienen una capacidad limitada de procesamiento de la información, la que es almacenada por un período breve de tiempo.

El nivel cognitivo se refiere al procesamiento de la información dentro de la memoria de trabajo y el intercambio de información que se establece entre la memoria de trabajo y la memoria a largo plazo, a partir de dos canales representacionales: canal verbal y canal visual. En este nivel, el lector construye una representación proposicional del contenido del texto y un modelo mental del contenido de la imagen a partir de la relación que establece entre la información procesada y seleccionada de la memoria de trabajo y los conocimientos previos que posee en la memoria a largo plazo.

En base a estos modelos y a investigaciones empíricas realizadas por distintos autores (Ayres & Sweller, 2005; Low & Sweller, 2005; Sweller, 2005) es que Schnotz (2005) sostiene que los beneficios o costes implicados en el aprendizaje multimodal dependen de distintos factores. El autor señala que para demostrar la validez del modelo, este debe ser capaz de predecir las condiciones bajo las que la integración de texto e imagen es beneficiosa para el aprendizaje. Sin embargo, el modelo también debe ser capaz de predecir las condiciones en que esta integración provoca efectos perjudiciales en el aprendizaje de los sujetos.

Es por esto que el autor explicita que los estudiantes aprenden mejor cuando imágenes y textos se relacionan en condiciones de coherencia y contigüidad. Estas condiciones se cumplen cuando las imágenes y el texto están relacionados semánticamente y se presentan en el mismo espacio y tiempo. Asimismo, se estima que el aprendizaje de los estudiantes mejora si se les enfrenta a textos orales con imágenes que a textos escritos con imágenes, ya que el acceso a la memoria de trabajo se puede realizar a través de distintos canales sensoriales y, por ende, procesar la información de manera diferenciada.

En cuanto a la influencia del conocimiento previo en el aprendizaje multimodal, Schnotz (2005) postula que la presentación de imágenes puede resultar más beneficiosa para estudiantes con bajo nivel de conocimientos previos. Las imágenes proporcionan al estudiante información adicional

que les permiten construir un modelo mental supliendo la falta de conocimientos internos. En cambio, para los estudiantes con alto nivel de conocimientos, la utilización de varias representaciones externas puede interferir en la construcción de modelos mentales satisfactorios, ya que el procesamiento cognitivo se ve obstaculizado por información adicional que en realidad estos no necesitan.

Asimismo, Schnotz (2005) propone que uno de los efectos perjudiciales que posee la combinación de textos e imágenes es la redundancia de información que puede presentarse tanto en el texto como en las imágenes. El autor postula que el efecto de redundancia tiene un impacto negativo en el aprendizaje, debido a que la información adicional ocupa capacidad de la memoria de trabajo y genera un costo cognitivo innecesario que no conduce finalmente a una mejor comprensión de los textos.

En estudios posteriores, Schnotz y colaboradores (Hochpöchler *et al.*, 2013) investigaron a través de la medición de movimientos oculares las estrategias que los sujetos utilizan al procesar información verbal y gráfica al mismo tiempo y las funciones de la información que extraen. Estos investigadores concluyen que el procesamiento de la información verbal se utilizaría en la creación del modelo de situación inicial, siendo utilizada solo para la actualización de dicho modelo en contextos en que la tarea de lectura lo requiera. La información gráfica se usaría más como un apoyo o andamiaje (*scaffolding*) para la creación inicial del modelo de situación y sólo se utilizaría dependiendo de los requerimientos de los sujetos en relación con sus características y la tarea que deben realizar. Específicamente, estos investigadores encontraron que los lectores de escuelas de rendimiento más alto y de mayor edad tienden a utilizar la información gráfica con mayor frecuencia como apoyo a medida que las exigencias de la prueba de lectura se tornan más complejas. En este sentido, los autores postulan que el procesamiento de la información derivada de los gráficos estaría guiada por estrategias de procesamiento de coherencia global, en una primera instancia y, luego, por estrategias de procesamiento de selección de información específica para la tarea, lo que podría indicar que estos estudiantes poseen mejores estrategias metacognitivas que les permiten reconocer la necesidad de contar con información adicional para completar tareas más complejas. Por otro lado, los estudiantes más jóvenes y de escuelas de rendimiento más bajo

tienden a utilizar la información que derivan del texto en un primer momento y, luego, escasamente complementan esta información con la que se deriva de la información gráfica. En este sentido, no adaptarían su procesamiento a los requerimientos de la tarea.

#### 4. Reflexiones finales

Los modelos de comprensión multimodal presentados en el apartado anterior permiten contemplar el desarrollo de los estudios relacionados con el procesamiento de la información considerando distintos sistemas semióticos. La TCD de Paivio (1971; 1986) descansa en el supuesto de que el procesamiento de información implica la actividad de dos subsistemas cognitivos de codificación independientes. Además, plantea que las representaciones que emanan del procesamiento de la información codificada mantendrían ciertas características de la modalidad sensorial con la que fueron percibidas. En este sentido, se aparta de los modelos siguientes que descansan en la existencia de representaciones amodales, como los esquemas o proposiciones. El modelo propuesto por Mayer y colaboradores, aun cuando asume la existencia de estos dos subsistemas cognitivos, postula la integración de ambos. Además, incorpora postulados de la teoría de la carga cognitiva y la teoría constructivista del aprendizaje. Schnotz y Bannert (2003) y Schnotz (2005) intentan superar lo propuesto en los modelos anteriores, tomando en cuenta la interacción del texto y la imagen en distintos momentos de su procesamiento y estableciendo niveles perceptuales y cognitivos que relacionan los estímulos externos con la memoria a corto y largo plazo. Proponen también la creación de dos modelos mentales, con funciones distintas, que obedecerían a diferentes estrategias de procesamiento según la tarea de lectura y la habilidad de los sujetos. En este sentido, se acercan más a un estudio integrado de comprensión de textos, considerando diversos factores como la tarea y los objetivos del lector.

Una reflexión importante que surge de esta revisión dice relación con la concepción del texto como inherentemente multimodal. Esto pone en tensión las prácticas de la enseñanza con respecto a la comprensión y producción de textos. Pese a que los esfuerzos por mejorar los niveles de comprensión de los estudiantes han sido cuantiosos, en general sólo se basan en la comprensión de textos en los que predomina eminentemente el sistema verbal. En este sentido, los modelos que hemos revisado

avanzan hacia un entendimiento de los procesos de comprensión de textos desde una perspectiva multimodal. Estos avances teóricos deberían orientar los futuros trabajos de investigación en esta área y la aplicación de estos postulados en el diseño curricular de los planes de formación docente.

Aun cuando estos modelos se encuentran en desarrollo, los postulados que proponen no pueden ser desatendidos, debido al amplio soporte empírico que los sustenta. Ahora bien, no podemos dejar de considerar que la evidencia en la que se basan proviene, en su mayoría, de investigaciones experimentales que no se ajustan a los contextos educativos reales. Se aprecia, entonces, una brecha que futuras investigaciones deberán encargarse de saldar.

Finalmente, debemos recalcar que no podemos seguir orientando las prácticas educativas en base a un concepto de texto eminentemente verbal, sino más bien es necesario que el proceso educativo se conciba desde un concepto multimodal del texto. La educación y las prácticas de enseñanza deben adaptarse, por tanto, a los requerimientos del contexto en el que se insertan los estudiantes y las características de sus procesos, es decir, a una realidad de comunicación multimodal.



## Referencias bibliográficas

- Alcántara-Pla, M. (2014). Las unidades discursivas en los mensajes instantáneos de wasap. *Estudios de Lingüística del Español*, 35, 223-242.
- Ayres, P., & Sweller, J. (2005). The Split-Attention Principle in Multimedia Learning. En R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp.135-146). New York: CUP.
- Baddeley, A. (1999). *Memoria humana: teoría y práctica*. Madrid: MacGrawhill.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Buckingham, D. (2007). *Beyond Technology: Children's Learning in the Age of Digital Media*. Cambridge, UK: Polity Press.
- Bezemer, J., & Kress, G. (2010). Changing Text: A Social Semiotic Analysis of Textbooks. *Designs for Learning*, 3(1-2), 10-29.
- Boulari, E., & Herring, S. (2014). Multimodal communication on Tumblr: "I have so many feels!". *Proceedings of the 2014 ACM conference on Web science*, 171-175.
- Cantamutto, L. (2015). Enseñanza de la lengua, comunicación por SMS y estilo electrónico: Un modelo posible. *El Toldo de Astier*, 6(10), 2-18.
- Cantamutto, L. (2012). *El uso de los mensajes de texto en el habla adolescente del español bonaerense*. Tesis de Licenciatura. Universidad Nacional del Sur, Bahía Blanca, Argentina.
- Cantamutto, L. (2013). La recursividad de las interacciones contemporáneas. Límites teórico-metodológicos del estudio de los SMS como conversación. *Revista de Ciencias Sociales de la Universidad Nacional de Quilmes*, 23, 83-104.
- Cautín-Epifani, V. (2014). Enfoque cognitivo para la comprensión de narraciones: una mirada desde la psicología discursiva y el modelo de indexación de eventos. *Literatura y Lingüística*, 29, 271-292.

- Cautín-Epifani, V. (2015a). *Formas de tratamiento en interacciones verbales escritas en la Biografía Facebook de hablantes de la provincia de Iquique* (tesis doctoral no publicada). Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile.
- Cautín-Epifani, V. (2015b). Poder virtual y formas de tratamiento en el discurso mediado por computadora: exploración en una red comunicativa virtual. *Forma y Función*, 28(01), 55-78.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.
- Dove, G. (2010). On the need for embodied and disembodied cognition. *Frontiers in Psychology*, 1, 129-141.
- Fariás, M., Obilinovic, K., & Orrego, R. (2011). Engaging multimodal learning ans second/foreign language education in dialogue. *Trabalhos de Lingüística Aplicada*, 50(1), 133-151.
- Giammatteo, M., & Albano, H. (2013). El léxico: de la vida cotidiana a la comunicación cibernética. *Signo y Señal: Revista del Instituto de Lingüística*, 24, 189-195.
- Gobato, F. (2014). *La escritura secundaria: oralidad, grafía y digitalización en la interacción contemporánea*. Buenos Aires: Biblos.
- Gómez, L. (1988). Elementos para un marco teórico de la comprensión textual. *Revista Signos. Estudios de Lingüística*, 21(26), 121-126.
- Hernández, C. (2001). Narración breve medieval e imagen (de la ilustración al cine). *Revista Estudios románicos*, 2(13-14), 109-125.
- Halliday, M. (1985). *An introduction to Functional Grammar*. Londres: Arnold.
- Herring, S. (2015). New frontiers in interactive multimodal communication. En A. Georgopoulou & T. Spilloti (Eds.), *The Routledge handbook of language and digital communication* (pp. 398-402). London: Routledge

- Hochpöchler, U., Schnotz, W., Rasch, T., Ullrich, M., Horz, H. McElvany, N. & Baumert, J. (2013). Dynamics of mental model construction from text and graphics. *European Journal of Psychology of Education*, 28(4), 1105-1126.
- Hodge, B. & Kress, G. (1988). *Social Semiotics*. Cambridge: Polity.
- Holmqvist, K., Holsanova, J., Barthelson, M. & Lundqvist, D. (2003): Reading or scanning? A study of newspaper and net paper reading. En J. Hyönä, & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 657-670). Amsterdam: Elsevier.
- Holsanova, J. (2004). Tracking multimodal interaction with new media [en línea]. Disponible en: [https://www.researchgate.net/publication/244137303\\_Tracking\\_multimodal\\_interaction\\_with\\_new\\_media](https://www.researchgate.net/publication/244137303_Tracking_multimodal_interaction_with_new_media)
- Jiménez, J. & Vela Delfa, C. (2005). La escritura electrónica: ¿progreso o decadencia? *Interlingüística*, 16(2), 621-630.
- Kaltenbacher, M. (2007). Perspectivas en el análisis de la multimodalidad: desde los inicios al estado del arte. *Revista Latinoamericana de Estudios del Discurso*, 7(1), 31-57.
- Kress, G. & Selander, S. (2012). Multimodal design, learning and cultures of recognition. *The Internet and Higher Education*, 15(4), 265-268.
- Kress G, & van Leeuwen T. (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Routledge.
- Low, R. & Sweller, J. (2005). The Modality Principle in Multimedia Learning. En R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp.147-158). New York: CUP.
- Manghi, D. (2010). Recursos semióticos del profesor de matemática: funciones complementarias del habla y los gestos para la alfabetización escolar. *Estudios Pedagógicos*, 36(2), 99-115.

- Manghi, D. (2011). La perspectiva multimodal sobre la comunicación. Desafíos y aportes para la enseñanza en el aula. *Revista electrónica Diálogos Educativos*, 22, 3-14.
- Manghi, D. (2013). Representación y comunicación del conocimiento en Educación Media: análisis multimodal del discurso de materiales utilizados para la enseñanza escolar de la historia y de la biología. *Onomázein*, 27(1), 35-52.
- Manghi, D. & Córdova, J. (2011). Definiciones y explicaciones multimodales: Potencial semiótico en la enseñanza de la biología en Educación Media. *Logos: Revista de Lingüística, Filosofía y Literatura*, 21(2), 17-39.
- Manghi, D., González, D., Echeverría, E., Marín, C., Rodríguez, P. & Guajardo, V. (2013). Leer para aprender a partir de textos multimodales: los materiales escolares como mediadores semióticos. REXE. *Revista de Estudios y Experiencias en Educación*, 12(24), 77-91.
- Manghi, D., Arancibia, M., Illanes, O., Herrera, P. & Zamora, M. (2014). Medios semióticos y definiciones multimodales en las clases de Ciencias Naturales y Ciencias Sociales en una escuela para jóvenes sordos. *Onomázein*, Número Especial IX ALSFAL, 37-56.
- Manghi, D., Badillo, C. y Villacura, P. (2014). Alfabetización semiótica en clases de historia, estrategias de mediación desde un enfoque multimodal. *Revista Perfiles Educativos*, 36(146), 63- 79.
- Martínez, J. (2015). Tres aproximaciones a los modelos de situación: Una propuesta integradora para la construcción de los modelos de situación del texto argumentativo. *Revista Nexus Comunicación*, 16, 6-27.
- Mayer, R. (1997). Multimedia Learning: Are We Asking the Right Questions. *Educational Psychologist*, 32(1), 1-19.
- Mayer, R. (2005). Cognitive Theory of Multimedia Learning. En R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31-48). New York: CUP.

- Mayer, R. (2009). *Multimedia Learning*. New York: Cambridge University Press.
- Mayer, R. (2014). Research-Based Principles for Designing Multimedia Instruction. En V. Benassi, C. Overson, & C. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* [en línea]. Disponible en: <http://teachpsych.org/ebooks/asle2014/index.php>
- Mayer, M. & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12, 107-119.
- Noblía, M. (2007). De amores y odios: la construcción multimodal de la identidad en los fotolog. En A. Spiegel (Ed.), *Nuevas tecnologías, saberes, amores y violencias. Construcción de identidades dentro y fuera de la escuela* (pp. 111-156). Buenos Aires: Noveduc.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- Paivio, A. (1986). *Mental Representations*. New York: Oxford University Press.
- Paivio, A. (1991). Dual Coding Theory: Retrospect and Current Status. *Canadian Journal of Psychology*, 45(3), 255-287.
- Parodi, G. (2010). Multisemiosis y lingüística de corpus: Artefactos (multi) semióticos en los textos de seis disciplinas en el corpus PUCV-2010. *Revista de Lingüística Teórica y Aplicada*, 48(2), 33-70.
- Sabich, A. (2015). El discurso pedagógico en tiempos de internet: aproximaciones sociosemióticas al estudio de los portales educativos. *Tonos Digitales*, 29. Recuperado de: <http://www.tonosdigital.es/ojs/index.php/tonos/article/view/1302/778>
- Sadoski, M. (2015). Reading comprehension, embodied cognition and Dual Coding Theory. En S. Parris & K. Headly (Eds.), *Comprehension Instruction: Research-based Best Practices* (pp. 45-55). New York: Guilford.

- Sadoski, M., & Paivio, A. (2004). A dual coding theoretical model of reading. En R. Ruddell & N. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 1329-1362). Newark, DE: International Reading Association.
- Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schnotz, W. (2002). Aprendizaje multimedia desde una perspectiva cognitiva. *Boletín Red Estatal de Docencia Universitaria*, 2(2), 31-40.
- Schnotz, W. (2005). An Integrated Model of Text and Picture Comprehension. En R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp.49-69). New York: CUP.
- Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141-156.
- Schnotz, W., Ludewig, U., Ullrich, M., Horz, H., McElvany, N. & Baumert, J. (2014). Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology*, 106(4), 974-989.
- Soleas, E. (2015). *New Teacher Perceptions of Inclusive Pedagogies: Designing New Futures for the Changing Classroom* (Tesis de maestría no publicada). Brock University, Ontario, Canadá.
- Sorden, S. (2013). Cognitive theory of multimedia learning. En B. Irby, G. Brown & R. Lara-Alecio (Eds.), *Handbook of educational theories*. Charlotte, NC: Information Age Publishing.
- Sweller, J. (1999). *Instructional Design in Technical Areas*. Camberwell, Australia: ACER Press.
- Sweller, J. (2005). The Redundancy Principle in Multimedia Learning. En R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp.159-167). New York: CUP.

- Torres i Vilartasana, M. (2003). *Aproximació a l'escriptura ideofonemàtica o netspeak català*. Ponencia presentada en 2n Congrés "Cap a quina ciutat del coneixement?", congreso en línea, Observatorio para la Cibernsiedad.
- Unsworth, L. (2010). Resourcing multimodal literacy pedagogy: Towards a description of the meaning-making resources of language-image interaction. En T. Locke (Ed.), *Beyond the grammar wars: A resource for teachers and students on developing language knowledge in English/literacy classroom* (pp. 276-293). London, UK: Routledge/Taylor & Francis.
- Unsworth, L. & Thomas, A. (2014). *English teaching and New Literacies Pedagogy: Interpreting and Authoring Digital Multimedia Narratives*. Nueva York: Peter Lang.
- Urrutia, M., & De Vega, M. (2012). Lenguaje y acción: Una revisión actual a las teorías corpóreas. *Revista de lingüística teórica y aplicada*, 50(1), 39-67.
- Van Dijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vela Delfa, C. & Cantamutto, L. (2015). Methodological Approach to the Design of Digital Discourse Corpora in Spanish. Proposal of the CODICE Project. *Procedia – Social and Behavioral Sciences*, 198, 494-499.