

## Procesamiento conjunto de lenguaje e imágenes en contextos didácticos: Una aproximación cognitiva

José César Perales López\* y Juan Francisco Romero Barriga

Universidad de Granada

**Resumen:** La imagen didáctica es una importante herramienta de complemento al lenguaje, tanto en la modalidad escrita como en la oral, para la transmisión de conocimientos en los contextos de enseñanza-aprendizaje. El presente trabajo revisa la literatura más significativa sobre procesamiento combinado de lenguaje e imagen, centrándose especialmente en aquellas situaciones en las que la imagen complementa a una narración (p.e. presentaciones multimedia, exposiciones orales...) o a un texto (p.e. libros de texto, enseñanza *on line*...).

Sobre la base de conocimientos firmemente establecidos sobre la estructura de la percepción, la atención y la memoria humanas, se propone un modelo cognitivo de integración texto-imagen –aplicable también con ciertas asunciones a presentaciones narración-imagen-. A diferencia de acercamientos anteriores, el modelo confiere una importancia fundamental a la memoria de trabajo y, sobre todo, a los procesos ejecutivos en el procesamiento activo del lenguaje y la imagen, y en el establecimiento de lazos referenciales entre ambos. El modelo es, además, específico sobre las funciones ejecutivas implicadas tanto en la construcción de representaciones unitarias que incluyan información visoespacial y proposicional (estabilización) como en la secuenciación de las acciones que el sujeto realiza en tiempo real durante la tarea de integración (secuenciación).

**Palabras clave:** Imagen didáctica; grafismo didáctico; codificación dual; comprensión lingüística; aprendizaje multimedia; procesos ejecutivos.

**Title:** Language-image integrative processing in instructional contexts: A cognitivist perspective.

**Abstract:** Didactical images are powerful tools to complement written and oral language comprehension in instructional situations. The present work summarizes the most relevant studies on conjoint processing of language and visual displays, focusing on those situations in which the image complements a narration (multimedia presentations, oral communications...) or a text (textbooks, online training materials...).

On the base of the well-established ground of human perception, attention and memory research, a cognitive model of text-image integrative processing is proposed –also applicable, taking into account certain considerations, to narration-image integrated presentations-. The model overcomes some of the limitations of previous approaches, and stresses the importance of executive processing in language and image interpretation, and in establishing referential links between their correspondent mental codes. The model specifies the executive functions involved in the elaboration of unitary-distributed representations that include visuospatial and propositional information (stabilization), and in sequencing the series of actions that the learner carries out in a real-time fashion during integration tasks (sequencing).

**Key words:** Didactical image; didactical graph design; dual coding; language comprehension; multimedia learning; executive processes.

Las imágenes, las composiciones gráficas, los mensajes visuales y los textos visuales<sup>1</sup> en general, han sido objeto de estudio y desarrollo con gran profusión en el ámbito socioeconómico, impulsados por los poderes fácticos de la sociedad de consumo, y esta circunstancia ha condicionado en gran medida la forma en que estos mensajes visuales han sido diseñados y producidos.

Como indica Costa (1998), dentro del conjunto de mensajes construidos con el lenguaje gráfico-visual, la imagen didáctica constituye un caso. Así, las imágenes publicitarias tienen como finalidad el incremento de las ventas de un determinado artículo de consumo y su estrategia se basa en manipular los elementos del lenguaje gráfico-visual para seducir al receptor. La pictografía señalética tiene como fin el diseño de información funcional y puntual para incorporarla al espacio de acción de los individuos y, para ello, se sirve de elementos gráficos normalizados. Las imágenes didácticas tienen como fin la transmisión de conocimiento y utilizan como estrategia composiciones gráficas

co-visuales que van desde imágenes puramente figurativas hasta esquematizaciones basadas en la abstracción que tratan de representar conceptos y fenómenos que no tienen una naturaleza visual.

En el ámbito de la educación formal las perspectivas de la didaxia y la autodidaxia se conjugan en un proceso sistemático que integra el conjunto de variables que intervienen en dicho proceso: unas provenientes de la acción de enseñar, otras de la acción de aprender, y otras de las condiciones físicas y materiales en las que tiene lugar el proceso de enseñanza-aprendizaje. Todo acto didáctico es un ejercicio de comunicación y consecuentemente, la Didáctica es responsable en última instancia de presentar métodos de intervención con capacidad de control sobre el conjunto de variables que han sido contempladas en modelos integrados de comunicación didáctica (Titone, 1981) con el fin de maximizar las ganancias del proceso. Pero para poder llevar a cabo ese cometido es muy importante que la Didáctica mantenga una relación interdisciplinar con áreas de conocimiento que pueden aportarle datos que han sido producidos científicamente.

La Psicología Cognitiva ha sido una fuente de saber de la que se ha surtido la Didáctica; de hecho los avances más significativos en metodología didáctica han aparecido tras incorporar los descubrimientos derivados de investigaciones realizadas desde el cognitivismo. Este trabajo es una reivindicación explícita de la necesidad de abordar el uso de la imagen en contextos didácticos bajo una supervisión conjunta de la Didáctica y de la Psicología Cognitiva, especialmente porque ésta última es la más capacitada para estudiar los mecanismos de los que disponen los

\* Dirección para correspondencia [Correspondence address]: José César Perales López. Deptº. Psicología Experimental (Secretaría). Facultad de Psicología. Universidad de Granada. Campus de Cartuja s/n. 18071 Granada (España). E-mail: [jcesar@ugr.es](mailto:jcesar@ugr.es)

<sup>1</sup> La expresión texto visual aparece aquí en el sentido dado por Eco (1979): “máquina semántico-pragmática que pide ser actualizada en un proceso interpretativo, cuyas reglas de generación coinciden con las propias reglas de interpretación”. Un significado equivalente alcanza en otros autores como por ejemplo (Vílchez, 2002, p. 39): “La imagen puede considerarse como un texto en el sentido de una unidad discursiva superior a una cadena de proposiciones visuales aisladas, que se manifiesta como un todo estructurado e indivisible de significación que puede ser actualizado por un destinatario”

agentes implicados en la comunicación didáctica para percibir, seleccionar y almacenar la información relevante.

Nuestro interés en este trabajo se centra en un modelo de situación en la que el docente utiliza la imagen como un recurso didáctico que se presenta junto con una narración<sup>2</sup> oral o texto escrito. Este tipo de situaciones es frecuente en el ámbito educativo: unas veces se produce cuando los docentes utilizan algún tipo de imagen como apoyo a sus explicaciones en el aula, en otras cuando los profesores proporcionan al alumnado algún tipo de texto escrito acompañado de un referente gráfico con la intención de reforzar la información contenida en el texto escrito. A estos dos formatos clásicos se ha añadido recientemente el de la enseñanza *online*. En el primer caso el emisor controla el momento y el tiempo de aparición de los códigos cifrados en el mensaje y en los otros dos, aunque el docente no puede ejercer el mismo dominio de la situación sí puede orientar al alumno sobre la mejor forma de integrar lenguaje e imagen para facilitar un procesamiento significativo de la información.

Pero la primera premisa para que pueda darse comunicación entre emisor y receptor en la comunicación didáctica viene dada por la necesidad de que ambos compartan el mismo lenguaje. En este caso, la utilización de una imagen didáctica, acompañada de un texto o narración implica ineludiblemente el uso de dos lenguajes autónomos entre sí, e independientes de otros tipos de expresiones. Todos los lenguajes poseen sistemas de significado y en las imágenes didácticas estos sistemas de significado vienen dados por una serie de códigos que son propios y exclusivos del lenguaje visual. Por tanto, la interpretación que el lector hace de la imagen no es un fenómeno exclusivamente perceptivo, sino que supone también una competencia lingüística distinta a la que tiene para la textual.

En una Educación Formal que ha recurrido tradicionalmente a la comunicación oral y escrita, esa capacidad lingüística es el primer requisito para seguir avanzando en el sistema educativo, sin embargo, no ocurre lo mismo con el lenguaje gráfico-visual al que apenas se le da importancia, posiblemente porque su introducción en el currículo escolar tiene una corta historia y, sobre todo, porque es un tema de estudio que se plantea de manera disgregada y, en el mejor de los casos, como una transversal (ver Ortega, 1998). El problema de la alfabetización visual es crucial y así es reconocido por muchos autores, como se verá más adelante. En el marco de la relación interdisciplinar Didáctica-Psicología algunas publicaciones (p.e. Ortega y Fernández, 1996) presentan programas de intervención válidos para alfabetizar visualmente en contextos educativos ya sean formales o no formales. En el terreno didáctico, algunos estudios se han encargado de indagar sobre la creatividad gráfica en la docencia universitaria y aportan un instrumento de evaluación del proceso seguido por los docentes para diseñar y elaborar sus propias gráficas didácticas, lo que puede sacar a la luz indicadores del grado de alfabetización visual entre los docentes que crean sus propios materiales didácticos (Romero, 2002a, b).

En este trabajo queremos resaltar la necesidad de abordar la comunicación didáctica desde una perspectiva que contemple el conocimiento de los mecanismos cognitivos mediante los cuales el sistema cognitivo humano procesa este tipo de códigos gráfico-visuales en conjunción con el lenguaje oral o escrito. Disponer de estos datos es fundamental para ensayar métodos de intervención didáctica que faciliten dichos procesos y eviten un buen número de interferencias en la comunicación.

## Una aproximación computacional al estudio de la imagen didáctica como ayuda al aprendizaje

El uso de la imagen en contextos didácticos requiere un abordaje conjunto desde las Ciencias de la Educación y las Ciencias Cognitivas. El objetivo primordial de este trabajo es dar una visión lo más coherente posible de la aportación de la Psicología Cognitiva a este panorama general. En última instancia el objetivo es contar con una serie de principios teóricos firmemente establecidos que permitan diseñar estrategias para optimizar el aprovechamiento de la imagen didáctica en relación a unos objetivos de aprendizaje.

De acuerdo con el marco de conocimiento propuesto por Marr (1982), el estudio de los procesos cognitivos requiere dos niveles de análisis mutuamente complementarios. En primer lugar, en el nivel llamado *computacional*, se considera la vertiente funcional del conjunto de procesos que vamos a estudiar. En nuestro caso, supone analizar la relación entre la información que el sujeto debe procesar (el *input*, en nuestro caso el texto o narración y la imagen) y el resultado de dichos procesos. Dicho de otro modo, en el nivel computacional el objeto de estudio son el 'por qué' y el 'para qué' de los procesos cognitivos. Sólo a partir del marco funcional que establece una teoría computacional puede realizarse un análisis más granular de los procesos concretos que se encargan de realizar esas funciones; esto es, sólo tomando como base los 'por qué' y los 'para qué' se puede abordar el estudio del 'cómo'. Este segundo nivel de análisis se denomina *nivel algorítmico*, y se ocupa de la identificación de las formas que toma la información en el interior del sistema cognitivo humano (*representaciones mentales*) y las diversas transformaciones y derivaciones que dicha información experimenta (*procesos cognitivos*).

En las últimas dos décadas se ha generado un interés creciente en relacionar los distintos tipos de imágenes (ver Romero, 2002a, b) con variables comportamentales y, específicamente, con los resultados del aprendizaje. Ello supone, si el acercamiento es lo suficientemente sistemático, un abordaje computacional del problema de la imagen como complemento didáctico.

Algunos de estos estudios se han centrado en una clase concreta de imágenes didácticas y han diseccionado las funciones que éstas cumplen. Así, por ejemplo, Fang (1996) ha analizado funciones que cumplen las ilustraciones en los libros de cuentos para niños y defiende su uso, bajo ciertas condiciones, en el aprendizaje de habilidades lectoras. En un contexto notablemente distinto, Shah y Hoeffner (2002) se han encargado de sistematizar las relaciones existentes entre distintas características de las gráficas matemático-estadísticas (formato, animación, color, tamaño, uso de leyendas) y la potencia de las mismas para

<sup>2</sup> A lo largo de este trabajo utilizaremos el término narración en sentido amplio, para referirnos a cualquier tipo de exposición oral que vaya acompañada de una imagen adjunta, ya sea dicha exposición en vivo, como, por ejemplo, en una clase estándar, o grabada, como en una presentación multimedia por ordenador.

ser utilizadas para determinados usos. Basándose en datos empíricos, estos autores llegan a la conclusión de que tanto los niños como los adultos poseen tendencias sistemáticas en la interpretación de estas gráficas, y que éstas pueden utilizarse intencionadamente para manipular, ocultar, priorizar o sesgar la información aparentemente objetiva que las gráficas aportan (ver también Shah y Carpenter, 1995, Gattis y Holyoak, 1996), pero también para maximizar la potencia didáctica de estas representaciones.

En esta misma línea, otros trabajos teóricos y experimentales, se han centrado en el uso de mapas topográficos (i.e. Griffin, 1995, Kulhavy, Lee y Caterino, 1985, Kulhavy, Pridemore y Stock, 1992, Kulhavy, Stock y Kealey, 1993, Rossano y Morrison, 1996, Schwartz y Kulhavy, 1981, Schwartz y Phillippe, 1991), mapas de conocimiento y mapas conceptuales (i.e. Chmielewski, Dansereau y Moreland, 1998, Hall y O'Donnell, 1996, Novak, 1998, O'Donnell, Dansereau y Hall, 2002), animaciones (Mayer y Anderson, 1992, Mayer y Moreno, 2002) y diagramas (Lowe, 1993a, 1993b). Es importante hacer notar que, aunque todos estos trabajos abordan de una forma u otra la cuestión de la funcionalidad de la imagen didáctica, esto es, el nivel de teorización computacional, muchos de ellos van más allá y se preocupan también de proporcionar un andamiaje procesual –algorítmico– de los efectos y las relaciones funcionales descritos.

Junto con estos estudios, interesados específicamente en determinados tipos de representaciones, otras revisiones han procurado ofrecer un marco teórico funcional más general del procesamiento de la imagen didáctica como complemento al texto o la narración, independientemente del formato concreto de la misma. En ellas se hace un mayor hincapié en los aspectos que los distintos tipos de imágenes didácticas tienen en común, definidas éstas como representaciones visuales en sentido amplio, que sirven como instrumento para “presentar conjuntos de datos, ejemplificar conceptos abstractos, organizar conjuntos complejos de información, para integrar nuevo conocimiento en las estructuras preexistentes, para facilitar la retención de información y para servir de guía a los procesos de pensamiento y solución de problemas” (Schnotz, 1993).

Desde esta perspectiva que podemos llamar “independiente de formato”, Carney y Levin (2002) han enumerado las distintas funciones que cumple la imagen como complemento al texto. En primer lugar, la imagen puede servir a un objetivo puramente *representativo*, si únicamente representa o refleja parte del texto, como suele ocurrir con las ilustraciones de los cuentos. En segundo lugar, las imágenes *organizativas* proveen un marco estructural adecuado para el contenido del texto (e.g. una ilustración que muestra los distintos pasos necesarios para ejecutar una reanimación cardiopulmonar). En tercer lugar, las imágenes *interpretativas* ayudan a clarificar un texto difícil, y hacen uso de conocimiento previo a través del razonamiento por analogía (e.g. representando la presión arterial utilizando un sistema de bombeo como metáfora). Por último, las imágenes *transformacionales*, incluyen principios mnemónicos sistemáticos para mejorar el recuerdo por parte de los aprendices de la información contenida en el texto. Esta clasificación está realizada para categorizar distintos tipos de ilustraciones, principalmente de tipo figurativo (aquellas en las que existe un alto grado de iconicidad, como, por ejemplo, un mapa topográfico o un diagrama anatómico). Sin embargo, otras imágenes y representaciones visoespaciales

con un mayor nivel de abstracción (como los mapas conceptuales, los mapas de conocimiento o las gráficas matemático-estadísticas) también pueden incorporarse con facilidad a esta clasificación, sobre todo en la forma de imágenes organizativas e interpretativas.

Levin, Anglin y Carney (1987) realizaron un meta-análisis de los estudios experimentales realizados hasta esa fecha que demuestra la superioridad de estos cuatro tipos de imágenes (representativas, organizativas, interpretativas y transformacionales) sobre la ausencia de imágenes o las imágenes puramente decorativas, a la hora de comprender o memorizar un texto.

Aunque los autores hacen un acercamiento al ‘cómo’ de la facilitación de la comprensión y el recuerdo mediante la imagen (contribuyendo a los procesos de focalización de la atención, de reducción y concreción de la información, de coherencia y correspondencia con las estructuras de conocimiento preexistentes, y de codificación mnemónica), su descripción de dichos procesos y, sobre todo, la especificación de las relaciones existentes entre dichos procesos y la estructura de la percepción, la atención, y la memoria humanas está lejos aún de ser exhaustiva.

Una aproximación similar, aunque algo más completa, es la adoptada por Ainsworth (1999). El marco funcional propuesto por esta autora es aplicable a todos aquellos contextos en los que un individuo debe integrar varias representaciones a través de más de una modalidad sensorial, esto es, a cualquier situación multimedia o multi-representacional. Por tanto, es aplicable directamente a las situaciones de aprendizaje en la que el sujeto debe servirse de una imagen para comprender mejor el contenido de una narración. Indirectamente, es aplicable también a situaciones en las que la imagen se presenta como adjunto a un texto, puesto que, aunque la lectura y la percepción de la imagen compartan un mismo canal sensorial, utilizan, de hecho, sistemas de procesamiento y representación notablemente distintos.

Resumidamente, las funciones de la representación múltiple serían las siguientes. En primer lugar, pueden cumplir *funciones de complementación*, tanto de *procesos*, en aquellos casos en los que representaciones informativamente equivalentes facilitan o soportan distintos tipos de inferencias (Larkin y Simon, 1987, Meyer, Shinar y Leisser, 1997, ver también Johnson-Laird y Byrne, 1991); como de *información*, en el caso de que una de las representaciones aporte contenidos que la otra representación es incapaz de contener. Imaginemos, por ejemplo que utilizamos un modelo gráfico, junto con una descripción verbal completa, para describir la situación relativa de distintos objetos. Aunque es posible resolver problemas de localización espacial únicamente a través de la aplicación de reglas sintácticas a la información proposicional contenida en la descripción verbal (e.g. “si A está a la derecha de B, y B está a la derecha de C, entonces A está a la derecha de C”). La representación pictórica –el modelo gráfico– permite hacer las inferencias necesarias con mayor facilidad. En este caso, aunque las dos representaciones sean informativamente equivalentes, la representación pictórica cumple un papel de complementación de los procesos de inferencia necesarios para la respuesta. En otras ocasiones, la imagen conlleva información que no está presente en la descripción verbal. Por ejemplo, una descripción de una cara no es informativamente equivalente a una fotografía de la misma cara. El reconocimiento de dicho rostro será superior si se aporta una representación

pictórica (la fotografía) por que ésta complementa la información facilitada en el texto.

En segundo lugar, las representaciones múltiples permiten *constreñir* la información facilitada por cada una de las representaciones individuales. Esto puede llevarse a cabo, por ejemplo, utilizando una representación familiar para reducir la ambigüedad de la información facilitada por una representación menos familiar (haciendo uso del razonamiento analógico). En el caso que nos ocupa, la presentación simultánea de texto o narración e imagen, puede afirmarse que la imagen es menos ambigua y más económica que el texto para expresar información espacial y estructural (Stenning y Oberlander, 1995). Por tanto, uno de los usos más frecuentes de la imagen como complemento es facilitar y constreñir la interpretación de este tipo de información en el texto. En algunas ocasiones, aunque menos frecuentes en el contexto de enseñanza-aprendizaje se busca el efecto inverso: el texto que acompaña a una imagen didáctica puede disponerse para precisar el mensaje que transmite la imagen, subrayar los elementos denotativos e la imagen y suprimir los connotativos.

Por último, las representaciones múltiples pueden utilizarse para *facilitar un procesamiento más profundo* de la información. La combinación de dos representaciones, verbal y visual, no sólo porta la información contenida en cada una de ellas, sino que existen propiedades que emergen de la interacción entre ambas. Así, la imagen puede facilitar los procesos de *abstracción*, entendida ésta tanto como la eliminación de los detalles superficiales –reducción–, como la generación de una superestructura organizada y jerárquica de la información contenida en el texto –reontologización– (ver Kinstch y Van Dijk, 1978). Para cumplir esta función resultan en general menos útiles las ilustraciones puramente figurativas que aquellas que expresan relaciones funcionales y estructurales entre las distintas entidades y conceptos presentados en el texto (por ejemplo, diagramas, Lowe, 1993a, 1993b, y mapas de conocimiento, Chmielewski, Dansereau, y Moreland, 1998).

Además de la abstracción, las representaciones múltiples pueden facilitar la *generalización a nuevos contextos* y la *transferencia a la acción* (Mayer y Sims, 1994), sobre todo porque la representación múltiple puede facilitar la adquisición de los principios abstractos comunes a las diversas representaciones y extrapolables a otras situaciones (aunque no siempre la generalización y la transferencia requieren previamente de la abstracción, Anderson, 1983).

Por último, y entre las funciones facilitadoras de un procesamiento más profundo, la imagen didáctica puede facilitar la *realización de inferencias elaborativas* sobre el contenido verbal, esto es, puede servir de puente entre las estructuras de conocimiento del sujeto, como los esquemas previos, y la información contenida en el texto o narración.

El número de taxonomías similares a la de Carney y Levin (2002) y Ainsworth (1999) es, teóricamente, ilimitado, pero todas ellas serían parcialmente coincidentes. De hecho, como puede observarse, las dos taxonomías funcionales descritas son en gran medida traducibles entre sí. La razón de la aparente variabilidad, y de que no haya un acuerdo generalizado sobre cuáles son las funciones de la imagen como complemento didáctico, es la desconexión que existe entre la investigación en el nivel pragmático-funcional (que nosotros consideramos fundamentalmente de naturaleza computacional) y el nivel de teorización

algorítmica, esto es, sobre los mecanismos concretos encargados de realizar las operaciones cognitivas que están en la base de las funciones propuestas. Como ya hemos comentado anteriormente, ese esfuerzo de conexión si se ha hecho para formatos específicos de imágenes didácticas, tales como los mapas topográficos o las gráficas matemático-estadísticas, pero se ha realizado en mucha menor medida para la imagen didáctica en general.

Najjar (1998) ha elaborado un marco teórico tetradimensional que puede servir de puente para establecer una conexión entre los niveles computacional y algorítmico de estudio de los procesos involucrados en la integración lenguaje-imagen. Dicho marco intenta clasificar los factores que explican en que casos y por qué la imagen mejora el aprendizaje y la comprensión del texto o narración.

La primera dimensión del modelo (más bien un marco de organización teórica) esta compuesta por las variables propias del *aprendiz*. Dentro de estas variables se incluyen el conocimiento previo del individuo, sus aptitudes y su estilo cognitivo, que determinan las diferencias interindividuales en situaciones esencialmente equivalentes. En relación a esta primera dimensión diversos estudios han demostrado que las mismas condiciones de aprendizaje y las mismos procedimientos de evaluación dan lugar a resultados de aprendizaje notablemente distintos.

Mayer y Gallini (1990) encontraron que la presentación de imágenes junto a un texto técnico mejoraba la transferencia en un grupo de sujetos novatos, pero no en un grupo de sujetos expertos. Dicho de otro modo, la imagen tendía a compensar, al menos en parte, la falta de experiencia en la materia en cuestión. Ese dato parece corroborar los datos anteriores que indicaban que los sujetos expertos tienden a beneficiarse en menor medida de la presencia de imágenes accesorias al texto que los sujetos novatos (Dean y Enemoh, 1983, Kunz, Drewniak y Schott, 1983, citado por Najjar, 1998; ver también Kalyuga, Chandler y Sellar, 2000).

En relación a la influencia de las aptitudes cognitivas, Mayer y Sims (1994) comprobaron que la presentación simultánea de un diagrama y un fragmento de un texto era más eficaz que la presentación secuencial (efecto de contigüidad), pero que dicho efecto sólo aparecía en sujetos con una alta aptitud para el procesamiento visoespacial. Por su parte, Wardle (1977, citado por Levie y Lentz, 1982) demostraron que los sujetos con una baja aptitud lectora se beneficiaban más de la presencia de una imagen accesorias que los sujetos de alta aptitud lectora (teniendo en cuenta que se requieren unas habilidades lectoras mínimas en todos los casos en los que la información se presenta en forma de texto escrito, Carney y Levin, 2002). Tan importante como las habilidades lectoras ha demostrado ser la alfabetización visual, esto es, la posesión de esquemas gráficos que permitan leer e interpretar correctamente el significado de los elementos morfosintácticos de la imagen y las relaciones entre ellos (Lowe, 1993b, Schofield y Kirby, 1994, Kirby, 1985, citado por Moore, 1993).

Por último, en relación a la influencia de los estilos cognitivos Chmielewski, Dansereau y Moreland (1998) demostraron que los sujetos categorizados como dependientes de campo se beneficiaban más del uso de un mapa de conocimiento presentado por partes relevantes a cada porción del texto en cada momento

de la tarea, mientras que los independientes de campo mejoraban más su ejecución si el diagrama se mostraba íntegro en todo momento.

La segunda dimensión del modelo viene dada por los *materiales de aprendizaje*. Así, se ha demostrado que el recuerdo de series dibujos es mejor que el recuerdo de series de palabras. Este efecto se ha dado en conocer como *efecto de la superioridad del dibujo* (Clark y Paivio, 1991, Paivio, 1991), y tiene su explicación, como veremos más adelante, en la existencia de sistemas de representación distintos para el almacenamiento de información visual e información verbal. En relación a este dato, se ha observado que la presentación de imágenes facilita el recuerdo de información espacial, mientras que la presentación textual facilita el recuerdo de información secuencial y procedimental (Bartram, 1980, Garrison, 1978).

Cuando se utilizan texto e imagen simultáneamente, se han identificado también una serie de factores relativos a la presentación del material que pueden afectar a su comprensión, recuerdo y transferencia. Por ejemplo, como hemos comentado anteriormente, la presentación simultánea de narración e imagen es más beneficiosa que la presentación serial (efecto de contigüidad espacial, Mayer y Sims, 1994, Mayer, Moreno, Boire y Vagge, 1999). Además, cuando la presentación es serial, la presentación de la imagen previa a la presentación de la narración es más beneficiosa para la comprensión que la presentación de la imagen posteriormente a la narración (Verdi, Johnson, Stock, Kulhavy y Whitman-Ahern, 1997). En presentaciones computarizadas, por su parte, se ha mostrado también que la presentación de la imagen en formato “pop-up” mediante hipervínculos es más eficaz (probablemente por el componente de contigüidad espacial entre texto e imagen que este tipo de presentación conlleva) que el formato de pantalla dividida, en la que una imagen y un texto se presentan separados, cada uno en una parte de la pantalla, durante toda la tarea (Betrancourt y Bissret, 1998). Este efecto de contigüidad también se ha demostrado para el uso simultáneo de texto-en-pantalla y animaciones gráficas (Moreno y Mayer, 1999).

Mayer, Heiser y Lonn (2001) demostraron que cuando se utiliza material auditivo (narraciones) junto con el material textual y visual, la redundancia audio-texto reduce la comprensión de la narración, así como el aprovechamiento de la imagen (efecto de redundancia, ver también Kalyuga, Chandler y Sweller, 1999, 2000). En ese mismo estudio, la utilización de imágenes o animaciones interesantes pero poco relevantes para la comprensión de las ideas centrales del texto, provocaba que la retención y la transferencia, no sólo no se incrementaran, sino que se redujeran significativamente (efecto de coherencia, ver también, Moreno y Mayer, 2000).

Por último, en esta dimensión también se sitúan los factores referidos a la dificultad del texto. En general, cuando más complejo es el contenido del texto o la narración, mayor es el efecto facilitador de la imagen (Levin y Mayer, 1993). Este factor hay que considerarlo, lógicamente, en relación al efecto del conocimiento previo.

De especial importancia de cara al desarrollo de aplicaciones educativas es la dimensión referida a la *tarea de aprendizaje*. En general, para alcanzar los máximos beneficios es recomendable que los aprendices procesen la imagen de forma activa (Dean y Kulhavy, 1981, Rieber, 1989, 1990, Moreno, Mayer y Lester,

2000), realizando operaciones sobre la misma, como etiquetar sus componentes más importantes (Peeck, 1993), elaborando ellos mismos la imagen (Rubman y Waters, 2000), o contestando a preguntas elaborativas (Adler, 1993). En este punto es esencial la figura del didacta, que en los contextos de enseñanza-aprendizaje es el responsable de dirigir dicho procesamiento activo.

Obviamente, el tipo de procesamiento que se realiza no es indiferente. En general, las tareas que conducen a un procesamiento más elaborativo y al establecimiento de conexiones referenciales cruzadas entre lenguaje e imagen (Mayer y Sims, 1994, Schnotz, 2002) producen una mejor retención, comprensión y transferencia de lo aprendido.

La última dimensión a tener en cuenta son las *características del test* que permite evaluar lo aprendido. Dicho test puede ser una prueba de retención, de comprensión o de transferencia que, en contextos didácticos, estará inseparablemente ligada a los objetivos de aprendizaje. La validez de la medida va a venir por tanto dada por su relación con las habilidades que se pretende que el aprendiz adquiera. El principio de interacción entre codificación y recuperación señala, además, que el nivel de ejecución va depender en gran medida del ajuste existente entre la representación o representaciones generadas durante el aprendizaje (codificación) y las demandas de la tarea a través de la que se evalúa dicho aprendizaje (ejecución). En general, puede afirmarse que “globalmente, el grado en que un objetivo educativo se ve beneficiado por la presencia de imágenes e ilustraciones depende del énfasis que se dé al conocimiento visoespacial en el test de aprendizaje” (Levie y Lentz, 1982). Si entendemos el término visoespacial en sentido amplio –tanto referido a características perceptuales individuales como estructurales– esta afirmación viene refrendada por numerosos estudios experimentales (Garrison, 1978, Beagles-Roos y Gat, 1983, Poon, Szabo y Ally, 1997, citado por Najjar, 1998).

El modelo teórico propuesto en este trabajo pretende sintetizar los argumentos expuestos hasta este momento. Todas las funciones propuestas para realizar taxonomías basadas en las funciones cognitivas de la imagen como complemento al texto o narración pueden resumirse en la siguiente: el objetivo del procesamiento conjunto de texto e imagen es elaborar una representación compleja (y, probablemente, distribuida entre varios subsistemas cognitivos) con elementos provenientes de los resultados del procesamiento visoespacial, los procesos de comprensión lingüística, y los elementos, tanto conceptuales y proposicionales como estructurales y espaciales, preexistentes en la memoria semántica. Dicha representación debe cumplir dos requisitos para resultar útil: primero, que dicha representación sea *estable*, esto es, que no contenga elementos de información contradictoria o ambigua, y, segundo, que la información contenida en ella sea *funcional*, esto es, que se ajuste a las demandas de la tarea de recuerdo, comprensión o transferencia que el aprendiz debe realizar. El resto de funciones –las citadas en los párrafos anteriores– quedan subordinadas a ésta, y su efecto queda explicado por el éxito o el fracaso para alcanzar este objetivo. Los factores que *facilitan el procesamiento unitario de la imagen* (factores perceptivos y atencionales como, por ejemplo, los principios de organización de la Gestalt); los factores que *facilitan la comprensión del texto o narración*, simplificándolo y facilitando su interpretación (por ejemplo, posibilitando la aplicación de macrorreglas

de reducción, construcción y generalización, y la aplicación de inferencias elaborativas; Kintsch y Van Dijk, 1978); los factores que *reducen la carga cognitiva* provocada por el procesamiento conjunto de dos fuentes de información; los factores que *reducen la interferencia* entre los procesos realizados y las representaciones intermedias generadas (especialmente en memoria de trabajo); los factores que *facilitan el establecimiento de conexiones* entre los distintos elementos de información activos en un momento dado, tanto provenientes del exterior (texto e imagen) como del conocimiento previo del sujeto (memoria semántica); y los factores que *posibilitan la recuperación unitaria de la representación generada*, tendrán un efecto positivo sobre la ejecución final. En el siguiente apartado describiremos una estructura cognitiva con capacidad para llevar a buen término esta tarea.

## Un modelo algorítmico de la integración lenguaje-imagen

### Mecanismos de procesamiento de lenguaje e imagen

Por separado, la imagen y el lenguaje siguen rutas de procesamiento muy diferentes. Por una parte, la representación de la imagen sigue una serie de etapas que van construyendo progresivamente sus características computacionales más complejas. Así, a partir de rasgos visuales simples, como los contrastes de intensidad y color, la textura o el movimiento, se componen bordes, manchas, barras o terminaciones, que son los elementos constitutivos de la forma. Por otra parte, los productos del procesamiento visual deben poder situarse en el espacio, primero en función de las coordenadas definidas por el observador, y después en un sistema de coordenadas independiente del observador (Marr, 1982). Tanto la identificación de las características de la imagen, como la disposición relativa de las mismas en el espacio (estructura) son necesarios para el reconocimiento de los patrones visuales para los cuales tenemos significados almacenados (ver, por ejemplo, Rumelhart y Zipser, 1986; Biedermann, 2000). En este punto, es necesario señalar que la capacidad para retener y recuperar una imagen va a depender en gran medida de que las distintas partes que la componen tengan una serie de características específicas y una distribución espacial que la hagan reconocible.

El procesamiento de la forma (a través del llamado sistema 'qué') y el procesamiento de la localización espacial (a través del sistema 'dónde') de un objeto siguen a su vez rutas diferentes, aunque interrelacionadas (ver Van Essen, Anderson y Felleman, 1992). Tanto los estudios de neuroimagen (e.g. McCarthy, Blamire, Puce, Nobre, Bloch, Hyder, Goldman-Rakic y Shulman, 1994, Jonides, Smith, Koeppe, Awh, Minoshima y Mintum, 1993, Courtney, Ungerleider, Keil y Haxby, 1996), como los estudios de registro celular en primates (e.g. Funahashi, Bruce y Goldman-Rakic, 1989) han demostrado que la franja dorsolateral de los lóbulos parietales es responsable del procesamiento y la representación de la información espacial, mientras que la franja ventral de los lóbulos occipitales y temporales es la principal responsable del procesamiento y representación de la forma. También se han identificado circuitos específicos para el procesamiento del movimiento.

La mayor parte de los subproductos del procesamiento visual son transitorios y no accesibles a la conciencia. Sin embargo, otros pueden mantenerse activos si las demandas de la tarea o los objetivos marcados por el sujeto lo requieren. Esta capacidad para mantener activa y manipular transitoriamente la información visoespacial, ya provenga ésta del exterior o de la memoria a largo plazo, es lo que llamamos *imaginaria mental* (Kosslyn y Rabin, 1999, Paivio, 1988) o, más coloquialmente, imaginación. El conjunto de estructuras y procesos que mantienen activa la información visoespacial a corto plazo reciben el nombre de *agenda visoespacial* (Baddeley, 2002, Baddeley y Hitch, 1974, 1994). Recientemente se ha demostrado que la agenda visoespacial contiene al menos dos subsistemas, un subsistema encargado de la retención de la información referida al espacio, la localización y el movimiento (el componente espacial), y otro referido a aspectos visuales no espaciales (el componente pictórico, ver Goldman-Rakic, 1999). En el caso que nos ocupa, estos dos subsistemas almacenarían temporalmente y manejarían los dos aspectos fundamentales de la imagen, a saber, los referidos a la identificación de los elementos morfosintácticos de la misma (puntos, iconos, formas, etc.), y los referidos a localización espacial de cada uno de dichos elementos (estructura compositiva). La disociación entre estos dos subsistemas se ha podido establecer a través de estudios conductuales, neurofisiológicos y neuropsicológicos (ver Courtney, Ungerleider, Keil y Haxby, 1996).

Por su parte, el texto o narración escrito sigue un procesamiento perceptual prelingüístico dependiente de la modalidad sensorial que es similar en sus primeras etapas al procesamiento de otros patrones sensoriales. En lo referido a este trabajo, sin embargo, interesa mucho más el procesamiento de tipo lingüístico.

Independientemente del modelo teórico que se adopte (ver, por ejemplo, Just y Carpenter, 1980, Kintsch y Van Dijk, 1978), el objetivo del procesamiento lingüístico a partir de un texto o una narración es llegar a una representación organizada, jerárquica y económica del contenido del mismo. En la elaboración de esta representación, la memoria de trabajo cumple un papel fundamental. Mantiene activos los elementos fonológicos, léxicos y gramaticales de un fragmento del discurso hasta que su contenido puede engarzarse en una o más proposiciones; después los elementos menos relevantes se dejan decaer y, con toda probabilidad, se olvidan.

La lista de proposiciones contenida en un texto compone el texto-base. Sin embargo, la memoria de trabajo tampoco mantiene activo todo el texto base simultáneamente, sino que va generando una estructura más reducida, estructurada y elaborada del texto, conforme va avanzando la lectura. Finalmente es esta estructura, que incorpora el contenido básico del texto junto con elementos provenientes del conocimiento previo del sujeto, la que se almacena a largo plazo (Schmalhofer y Glavanov, 1986, Kintsch, Welsch, Schmalhofer y Zimny, 1990); los elementos menos centrales, y que han recibido una menor profundidad de procesamiento tienden también a olvidarse.

La memoria de trabajo, por tanto, necesita también de un componente auditivo-verbal (Gathercole y Baddeley, 1997), que mantenga activos el segmento lingüístico que está siendo procesado en un momento determinado y, al menos en parte, la macrorepresentación del discurso, progresivamente más abstracta

y elaborada, que se va generando durante la comprensión. Este componente puede disociarse de los componentes visoespaciales descritos anteriormente, como demuestran múltiples estudios conductuales (e.g. Brooks, 1968), neuropsicológicos (e.g. McCarthy y Warrington, 1990) y de neuroimagen (e.g. Smith, Jonides y Koeppel, 1996).

En resumen, el procesamiento del patrón visual da como resultado una *representación eidética* que incorpora la identificación de los elementos básicos de la escena, así como su localización y las relaciones estructurales que mantienen. El procesamiento del texto o narración da lugar, por su parte, a una *representación proposicional*, que incorpora, fundamentalmente, conceptos y relaciones entre ellos, a distintos niveles jerárquicos. Ambas representaciones –o al menos parte de ellas– pueden mantenerse activas en distintos subsistemas de la memoria de trabajo, que también puede operar con ellas. Y lo que es más importante, puesto que en la mayor parte de los contextos de aprendizaje el texto-narración y la imagen están presentes de forma simultánea, ambas representaciones van generándose en paralelo y los dos flujos de procesamiento tienen la posibilidad de influirse mutuamente. Según el modelo propuesto, las operaciones responsables de dicha interacción involucran fundamentalmente a *elementos de los sistemas de representación visoespacial y verbal de la memoria de trabajo* provenientes de los resultados de procesamiento de la imagen y el texto o narración, y a *elementos tanto espaciales-estructurales como conceptuales-proposicionales de la memoria a largo plazo* transitoriamente activos.

La coordinación de esta tríada de subsistemas –agenda visoespacial, almacén verbal-fonológico, y memoria a largo plazo– corre a cargo del ejecutivo central (Baddeley, 2002, Baddeley y Hitch, 1974, 1994; para una revisión ver Georgopoulos, Taira, Lukashin, Wurtz, Tach, Mink, Goodkin, Keatin, Norman, Shallice y Goldman-Rakic, 2000), un conjunto de sistemas de procesamiento encargado de la coordinación y secuenciación de las acciones asignadas al resto de subsistemas activos.

De los distintos factores y variables que en el apartado anterior hemos citado como relevantes a la hora de maximizar o minimizar el aprovechamiento por parte del aprendiz de la imagen como complemento a un texto o a una narración, algunos de ellos ejercen su acción en la facilitación del procesamiento visual o verbal por separado, previamente al establecimiento de relaciones referenciales entre los subsistemas verbal y visoespacial (la construcción de relaciones referenciales y la implicación de los procesos ejecutivos en la misma serán analizados con detalle en el apartado siguiente).

Por una parte, el procesamiento de la imagen y su posterior recuperación, pueden potenciarse confiando a la misma una estructura familiar o reconocible para el aprendiz, o disponiendo sus elementos de acuerdo a las leyes de la Gestalt (Hoffman, 1998, Wallace, West, Ware y Dansereau, 1998, O'Donnell, 1994). En tanto que la imagen sea codificada de forma más unitaria y significativa, también podrá ser utilizada como una clave de recuerdo más potente en la ejecución posterior.

Igualmente, la codificación y el recuerdo del texto o narración también serán más efectivos cuanto más claro, menos ambiguo y mejor estructurado sea éste. Aunque no se ha realizado una investigación exhaustiva de la influencia de estos factores en contextos de aprendizaje de textos o narraciones con imágenes, existen datos muy claros que apuntan en esta dirección. Por

ejemplo, Mayer, Heiser y Lonon (2001) han demostrado que la utilización de un lenguaje con un mayor número de referencias personales, y un menor número de tecnicismos mejora el recuerdo posterior y la integración lenguaje-imagen. Igualmente, como ya se comentado anteriormente, en presentaciones multimedia narración-animación la adición de detalles interesantes pero poco relevantes a la animación (Mayer, Heiser y Lonon, 2001, Moreno y Mayer, 2000) no sólo no mejora, sino que dificulta la comprensión de la narración, su integración con la imagen, y la transferencia del conocimiento adquirido a la ejecución. Estos efectos pueden explicarse tanto por la dificultad para generar una macroestructura proposicional simple y coherente, como por la sobrecarga cognitiva producida por el exceso de información, que deja pocos recursos ejecutivos libres para la construcción de lazos referenciales.

En el apartado anterior, considerábamos las aptitudes y la experiencia previa del sujeto como variables importantes a tener en cuenta al analizar el procesamiento de textos-narraciones e imágenes. Como veremos más adelante, el efecto de la experiencia y las aptitudes se localiza en parte en el establecimiento de conexiones lenguaje-imagen, pero también se debe en gran medida al efecto que tiene sobre el procesamiento de texto-narración e imagen por separado. Por ejemplo, los sujetos clasificados como ‘visualizadores’ (aquellos que tienen una mayor aptitud para el procesamiento visoespacial complejo) hacen un mejor aprovechamiento de la imagen como complemento a un texto (Schofield y Kirby, 1994, Gyselinck, Ehrlich, Cornoldi, De Beni y Dubois, 2000, Pazzaglia y Cornoldi, 1999, ver también Mumaw y Pellegrino, 1984). Es de imaginar que dicho efecto se debe a que estos sujetos elaboran representaciones visuales más detalladas y estructuradas, o bien tienen una mayor capacidad para recuperar la información visual almacenada.

En este nivel es necesario tener en cuenta que, al igual que la aptitudes lectoras son un determinante importante de la comprensión del texto (de tal manera que las personas con baja aptitud lectora se benefician más de la inclusión de imágenes que los buenos lectores, Wardle, 1977, citado por Levie y Lentz, 1982), la *alfabetización visual* es igualmente crucial en la comprensión y la retención de imágenes. Se tiende a pensar equivocadamente que la interpretación de las imágenes, incluso de aquellas que contienen un alto nivel de abstracción, dependen de habilidades intuitivas, insertas de forma natural en los sistemas de procesamiento visual. Este es uno de los errores más frecuentes que los docentes y los autores de libros de texto tienden a cometer en su labor. Numerosos estudios han demostrado empíricamente la importancia de una alfabetización visual paralela a la alfabetización verbal, y acorde a la edad del aprendiz. Por ejemplo, se ha observado que los niños preescolares tienen habilidades de comprensión visual que van poco más allá de la representación figurativa realista (Schnotz, 2002, Shah y Hoeffner, 2002). En estudios relacionados se ha demostrado que los individuos requieren esquemas cognitivos específicos –llamados *esquemas gráficos*– para ser capaces de leer información de configuraciones visoespaciales (Lowe, 1993b), y muchos sujetos, incluso adultos, carecen de dichos esquemas (Kirby, 1985, citado por Moore, 1993, Schofield y Kirby, 1994).

Aunque este hecho no está suficientemente recalado en la literatura, el estudio del efecto de la alfabetización visual entronca directamente con los estudios clásicos sobre el papel del co-

nocimiento experto en el recuerdo de configuraciones visoespaciales, tales como posiciones del juego del ajedrez (Chase y Simon, 1973, Simon y Chase, 1973), o diagramas de circuitos electrónicos (Egan y Schwartz, 1979). Esta línea de investigación ha demostrado suficientemente que en estos casos el mejor recuerdo no se debe a la existencia de cambios de las estructuras de memoria propiamente dichas, sino en un intercambio más eficaz de información entre la memoria a largo plazo y la memoria de trabajo.

Por último, la formación de representaciones de alto orden del contenido verbal y de la imagen será tanto más eficaz cuanto menor sea el grado de interferencia que se produzca en los distintos niveles de ambas rutas de procesamiento (para una revisión, ver Kirby, 1993). Así, en presentaciones multimedia en las que la presentación de la información es rápida y el sujeto no tiene control sobre la misma (por ejemplo, la presentación de una animación mientras una voz en *off* narra lo que está ocurriendo), es preferible que la información visual (la animación) y la información verbal (la narración) utilicen canales sensoriales distintos. Si las dos fuentes de información utilizan el mismo canal, por ejemplo, presentando la narración en formato texto o duplicándola (texto + narración oral), se produce una situación de competencia entre las dos fuentes de información por los mecanismos perceptivos visuales que dificulta la comprensión, retención y transferencia (Mayer y Moreno, 1998, Moreno y Mayer, 1999, Mayer, Heiser y Lonn, 2001). En un estudio reciente, Romero (2002b) demostró que esta reduplicación texto-narración es frecuente en los profesores universitarios.

### **El papel de la memoria de trabajo en el procesamiento coordinado de texto e imagen**

Para dar el alcance más global posible a esta revisión, hasta el momento hemos considerado el procesamiento lingüístico de narraciones y textos de forma indiferenciada. La mayor parte de lo dicho hasta ahora sirve, por tanto, igualmente a presentaciones multimedia imagen-narración, como para presentaciones texto-imagen. Ciertamente, la lectura y la comprensión del lenguaje oral implican procesos muy distintos, aunque partimos de la suposición de que la naturaleza de la representación final que se alcanza sobre la información recibida es de similar naturaleza en ambos casos. En este apartado, sin embargo, nos centraremos en el papel de la memoria de trabajo en la integración de textos e imagen. Puesto que texto e imagen utilizan el mismo canal sensorial, ello impone constricciones extra sobre los procesos implicados en la integración, derivados de la alternancia necesaria entre texto e imagen (en el caso de narraciones, los dos canales pueden recibir información simultáneamente). La presentación en forma de texto incrementa además el control del sujeto sobre la tarea, puesto que, en condiciones normales, éste puede avanzar y retroceder en la lectura del texto con libertad (lo que se ha dado en llamar lectura recursiva). Dado que uno de nuestros objetivos es resaltar la importancia de las estrategias activas de procesamiento, la integración texto-imagen sirve como una situación más adecuada para nuestra reflexión.

La visión actual de la memoria de trabajo es muy diferente al concepto de memoria a corto plazo propio de las teorías estructuralistas propuestas en los albores de la Revolución Cognitiva (Atkinson y Shiffrin, 1974). Por una parte, la evidencia empírica

demuestra abrumadoramente que la memoria de trabajo no es una estructura unitaria, sino múltiple, y compuesta por una serie de subsistemas disociables. Por otra, se hace cada vez más hincapié, no sólo en las labores de almacenamiento de la memoria de trabajo como *buffer* de retención temporal de los resultados del procesamiento, sino también en sus funciones de procesamiento. En la última década, los procesos de alto nivel que se encargan de la secuenciación y la organización de las distintas acciones llevadas a cabo por los distintos subsistemas de la memoria de trabajo, así como de la detección de conflictos entre dichas acciones –los procesos ejecutivos– se ha convertido en un tema central (Baddeley, Della-Sala, Papagno y Spinnler, 1997, Georgopoulos *et al.*, 2000, Meyer y Kieras, 1997, Posner y DiGirolamo, 1998).

De la visión actual de la memoria de trabajo se derivan las dos características fundamentales de las teorías que pretenden explicar como los individuos interpretan conjuntamente texto e imagen en contextos didácticos. La primera de dichas características es la *codificación dual*, esto es, el reconocimiento de que el contenido del texto y el contenido de la imagen tienen formatos internos de representación esencialmente distintos. La segunda característica es la relevancia dada al establecimiento de *lazos referenciales* entre estas dos representaciones. Sorprendentemente, aunque el estudio de la segunda de estas características resulta más retador, lo cierto es que ha inspirado un volumen de investigación notablemente menor que la primera.

El primer punto de contacto entre la investigación básica en Psicología Cognitiva sobre el procesamiento conjunto de texto e imagen y la investigación aplicada en contextos didácticos es la *hipótesis de la codificación dual* (Paivio, 1986, 1991, Paivio y Clark, 1986, Clark y Paivio, 1991). De acuerdo a esta hipótesis, la información verbal y la información pictórica son procesadas y representadas en subsistemas cognitivos distintos. Esta afirmación se apoya en el efecto de la superioridad del dibujo, que consiste en la demostración de que las personas recordamos mejor una lista de dibujos que una lista de palabras equivalentes. Ello se debería, según la propuesta de Paivio (1986), a que los dibujos son inmediatamente traducidos a un formato fonológico-verbal en la memoria de trabajo, por lo que quedarían doblemente representados a largo plazo, verbal y eidéticamente; por el contrario las palabras no se traducen automáticamente a un formato eidético, por lo que su codificación sería sencilla, exclusivamente en formato verbal. El efecto de superioridad del dibujo se debería meramente al efecto cuantitativo producido por esta superposición de representaciones.

Los dos sistemas de codificación propuestos por Paivio coinciden a grandes rasgos con los dos subsistemas verbal-auditivo y visual propuestos por Baddeley y Hitch (1974, 1994); el *lazo fonológico* y la *agenda visoespacial*. Además, la hipótesis de la codificación dual mantiene que las características de las representaciones en la memoria de trabajo, a corto plazo, se conservarían en los trazos de memoria producidos por las mismas a largo plazo. Esta idea ha sido fuertemente criticada en el pasado (Pyshlyln, 1974), pero cuenta con un importante apoyo empírico procedente de estudios comportamentales y neuropsicológicos (ver Vaidya, Zhao, Desmond y Gabrieli, 2002). Una de las características importantes del modelo presentado en el presente trabajo, y congruente con esta idea, es que los trazos almacenados en la memoria a largo plazo son la huella generada por las



operaciones de codificación y procesamiento. Esta idea es la base de la concepción procesual (opuesta a la estructural) de la memoria humana (Craick y Lockhart, 1972). En nuestro modelo, esto significa que la representación generada durante el procesamiento conjunto de texto e imagen estará distribuida a lo largo de varios subsistemas y que, durante su codificación, incluirá trazos activos de la memoria a largo plazo, tanto de naturaleza semántico-proposicional como visoespacial, así como elementos provenientes del medio y representados en los dos subsistemas hipotetizados en la memoria de trabajo, entre los que se establecen conexiones referenciales. Existen varios modelos de recuperación en memoria, principalmente de tipo matricial o conexionista (i.e. Pike, 1984) que permiten explicar cómo esta representación distribuida puede reactivarse de forma unitaria cuando es requerida.

Existe una gran cantidad de evidencia que apoya la implicación concurrente de la agenda visoespacial y del subsistema verbal-fonológico de la memoria de trabajo en las tareas de integración texto-imagen en contextos didácticos. Gyselinck, Ehrlich, Cornoldi, De Beni y Dubois (2000) demostraron que los aprendices con una mayor capacidad de almacenamiento visual, medida mediante el test de bloques de Corsi (Kessels, van Zandvoort, Postma, Kapelle y de Haan, 2000), también se benefician más de la inclusión de imágenes como ayuda a la comprensión de un texto. En un estudio posterior, este mismo grupo de investigadores (Gyselinck, Cornoldi, Dubois, De Beni y Ehrlich 2002) demostró que la inclusión de una tarea concurrente de *tapping* (una tarea que ocupa el componente espacial dentro de la agenda visoespacial de la memoria de trabajo) hacía desaparecer el efecto facilitador de la imagen en la comprensión de un texto, pero no afectaba a la comprensión del texto mismo, en la condición sin imagen adjunta. Por el contrario, una tarea de supresión articulatoria, que interfiere con el lazo fonológico de la memoria de trabajo, produce un efecto deletéreo de igual magnitud en el grupo que recibía sólo el texto y el grupo que recibía el texto con una imagen adjunta. Estos datos se interpretan como una prueba de que la comprensión del texto requiere del correcto funcionamiento del lazo fonológico<sup>3</sup>, tanto si se dispone de una imagen adjunta como si no, mientras que el efecto facilitador de la imagen depende críticamente del correcto funcionamiento de la agenda visoespacial (ver también Kruley, Sciana y Glenberg, 1994, Sims y Hegarty, 1997, Pearson, Logie y Gilhooly, 1999). En este estudio también se replicaron las diferencias entre buenos y malos visualizadores obtenidas en el trabajo anterior. La comparación entre los resultados de ambos estudios permite concluir, además, que el efecto facilitador de la imagen

se debe, sobre todo, a la participación del componente espacial de la memoria de trabajo (que codifica la estructura), y no tanto al componente pictórico (que representa la identidad y la forma).

Los principales modelos de integración texto-imagen hacen uso de esta evidencia, y por tanto, postulan un doble sistema de representación para la información verbal y la información visoespacial. Sin embargo, estos modelos no sólo incorporan el efecto cuantitativo producido por la doble representación, sino que también incorporan el efecto cualitativo debido al establecimiento de lazos referenciales entre los elementos de ambas representaciones.

La *teoría del procesamiento conjunto* de Kulhavy y colaboradores (Kulhavy, Stock y Kealy, 1993) enfatiza la importancia de la disponibilidad simultánea de información verbal y pictórica en la memoria de trabajo, para que puedan establecerse conexiones cruzadas entre las dos representaciones. En tanto que esta teoría está basada en los datos obtenidos con mapas pictóricos como ayuda al texto, estas conexiones cruzadas se refieren, sobre todo, a la identificación de los términos en el texto y los iconos en la imagen que se refieren a un mismo elemento en la realidad. Una propuesta similar es la *teoría del aprendizaje multimedia* (Mayer, 1997, 2002), según la cual el procesamiento del discurso base da lugar a un modelo mental del mismo, mientras que el procesamiento de la imagen da lugar a un segundo modelo mental de la imagen. Por último, se construirían las conexiones entre los elementos de ambos modelos mentales. Esta teoría también resalta la importancia del procesamiento activo, esto es, la selección del material relevante, su organización en una representación coherente y su integración con el conocimiento previo. Sin embargo, este procesamiento activo no se detalla más allá de la simple mención de su existencia.

Schnotz (2002) ha propuesto un *modelo integrador* que no se diferencia en gran medida de los anteriores, aunque sí es más detallado en algunos aspectos. Por una parte, el procesamiento semántico del texto de superficie (que incorpora el resultado de todos los procesos perceptivos prelingüísticos) lleva a una representación proposicional del mismo, que incorpora los conceptos y proposiciones contenidos en el texto. Por su parte, la imagen visual, fruto de los procesos sensoriales y perceptivos de bajo nivel, lleva a un modelo mental que surge de la interacción de esos procesos perceptivos de bajo nivel con el conocimiento del sujeto. El modelo mental se diferencia de la representación visual en que no es dependiente de ninguna modalidad sensorial. Aunque el modelo mental guarda una relación de analogía con aquello que representa (Johnson-Laird y Byrne, 1991) no depende necesariamente de la modalidad visual, sino que puede incorporar información de varios modalidades sensoriales (visual, háptica, auditiva).

En la teoría de Schnotz las proposiciones del modelo proposicional pueden traducirse a un formato estructural, esto es, su contenido puede incorporarse al modelo mental, a través de un proceso de *construcción del modelo*. A la inversa, la información estructural del modelo mental puede traducirse a conceptos y proposiciones, mediante un proceso de *inspección del modelo*. Dicho de otro modo, el modelo mental influye en la construcción de la representación proposicional y viceversa. Además, hay que considerar el flujo de procesamiento de arriba-abajo, que va a) de las representaciones conceptuales de la memoria a largo pla-

<sup>3</sup> El papel del lazo fonológico en la lectura y la comprensión del lenguaje oral es controvertida, dado que se han relatado casos de personas con una capacidad fonológica a corto plazo extraordinariamente baja (Warrington y Shallice, 1970) que, sin embargo, no presentaban déficits de comprensión catastróficos. Al parecer, en estos sujetos, la capacidad fonológica a corto plazo residual es suficiente para ir generando una representación proposicional-semántica del contenido del discurso, a costa de una disminución moderada de la velocidad de comprensión y de la presencia de dificultades de comprensión de oraciones muy complejas. Es de suponer, por tanto, que la representación de los contenidos verbales requiere al menos de dos subsistemas, uno superficial, exclusivamente fonológico-auditivo y otro más abstracto, proposicional-semántico (algunos autores han postulado el término *memoria de trabajo a largo plazo* para referirse a este segundo sistema, Butcher y Kintsch, 2003).

zo tanto al modelo mental como a la representación proposicional, y que determina la influencia del conocimiento previo, y b) de la representación proposicional y el modelo mental a los procesos perceptivos, lo cual permite seleccionar y deshacer la ambigüedad de la información entrante.

Propuestas similares a las descritas pueden encontrarse en Kirby (1993), en Mayer, Heiser y Lonon (2001), y en Gyselinck, Ehrlich, Cornoldi, De Beni y Dubois (2000). En todas ellas se repiten los presupuestos básicos de codificación dual y establecimiento de lazos referenciales, pero, también en todos los casos, la descripción de cómo se establecen esos lazos referenciales es bastante pobre. Por otra parte, a pesar de la importancia teórica que se concede a la memoria de trabajo, se olvida de forma sistemática el papel central de los procesos ejecutivos en el establecimiento de dichos lazos, y en la planificación general de la tarea.

Nuestro acercamiento teórico procura solventar estas limitaciones, partiendo de la evidencia existente. En primer lugar, compartimos el presupuesto de la codificación dual. El texto se codifica primero de forma superficial y después en un formato más abstracto y elaborado conforme va avanzando la lectura. La construcción de dicha representación, que vamos a llamar proposicional o macroproposicional, según su nivel de abstracción, requiere de un conjunto muy complejo de procesos: acceso léxico, aplicación de reglas gramaticales para la extracción de proposiciones, aplicación de reglas locales para la construcción de macroproposiciones, y aplicación de arriba-abajo de esquemas de conocimiento previo en los tres niveles.

En el modelo de Baddeley y Hitch (1974, 1994) de memoria de trabajo, sólo se contempla la existencia de un almacén fonológico-auditivo para el material verbal entrante. La información semántica y proposicional referida a esa información entrante se considera información a largo plazo. Hasta este momento, en este trabajo hemos utilizado el término componente verbal o verbal-fonológico, en lugar del más tradicional lazo articulatorio, precisamente para hacer ver que dicho componente (que posiblemente incorpora varios subcomponentes) no puede ser exclusivamente fonológico, sino que debe incorporar también información conceptual y proposicional transitoriamente activa<sup>4</sup>. Es más, consideramos que la memoria a corto plazo no es más que la información transitoriamente activa en las estructuras encargadas de su procesamiento. Esta afirmación es congruente con los estudios de imagen cerebral que demuestran que parte de las estructuras encargadas del 'habla interna' son las mismas que se encargan de la percepción del habla (Posner y Raichle, 1994). La dificultad tradicional para encontrar la localización de

la memoria a largo plazo se debería, por tanto, a que las huellas de memoria a largo plazo se encuentran distribuidas por una multiplicidad de localizaciones, relacionadas estrechamente con las mismas estructuras encargadas de procesar los formatos de información correspondientes (Roesler, Heil y Hennighausen., 1995, 1996). El término memoria de trabajo es sin embargo aún útil, porque es imprescindible postular mecanismos responsables de mantener la activación en los lugares correspondientes mientras la información relevante es necesaria (Cohen, Forman, Braver, Casey, Servant-Schreiber, y Noll, 1994, Goldmann-Rakic, 1999, Kandel, 2003), y de secuenciar las acciones que determinan la consecución de un objetivo (Meyer y Kieras, 1997).

Por motivos de claridad expositiva vamos a mantener en nuestro esquema la diferenciación entre información proposicional en memoria de trabajo e información proposicional activa en memoria a largo plazo, con el objetivo de indicar que una buena parte del procesamiento del texto depende de la aplicación de procesos de abajo-arriba y de reglas de coherencia local que no involucran decisivamente al conocimiento previo sobre la materia (Kintsch, 1988).

Lo mismo puede decirse de la representación generada en el subsistema visoespacial de la memoria de trabajo. Este subsistema incorpora información pictórica y espacial transitoriamente activa en las estructuras encargadas de su procesamiento (Kosslyn, Flynn, Amsterdam y Wang, 1990), que, a su vez, produce la activación de información almacenada a largo plazo, tanto de naturaleza conceptual-proposicional (por ejemplo, dos imágenes similares unidas por una flecha puede interpretarse como una transición entre dos estados de un mismo objeto), como de carácter espacial-estructural (por ejemplo, esquemas previos almacenados para determinadas configuraciones espaciales).

Por lo tanto, en cualquier momento durante la realización de una tarea de integración texto-imagen lo que podríamos observar (si dispusiésemos de los instrumentos necesarios para ello) es una colección de elementos activos de forma simultánea: información visoespacial en los componentes espacial y pictórico de la memoria de trabajo, proveniente de la imagen; información verbal y proposicional provenientes del texto; e información relativa a la información entrante, tanto proposicional como espacial-estructural, que se ha activado en la memoria a largo plazo.

El primer proceso importante en este modelo es establecer las relaciones posibles entre los elementos transitoriamente activos. Aunque este proceso es computacionalmente complejo, entendemos que se produce en parte de forma automática. El sistema por sí mismo tiende a reducir su ambigüedad, restringiendo la interpretación de un elemento por la influencia paralela de otros elementos. El aprendiz, en cualquier caso, dispone de mecanismos de selección atencional que pueden permitirle buscar activamente coincidencias entre elementos concretos. Esas coincidencias se refieren a dos tipos de relaciones. Por una parte, elementos discretos de la imagen como un nodo en un mapa conceptual, puntos de referencia o iconos en un mapa topográfico, o incluso elementos dinámicos como una flecha o una transición de color, pueden corresponderse uno a uno con conceptos o acciones en el texto (un sujeto, un predicado o un complemento). Así, una expresión tal como "A se transforma en B" puede referirse en la imagen como dos iconos represen-

<sup>4</sup> Existen varias aproximaciones sobre estructura de la memoria que más se ajusta a los datos disponibles en comprensión del lenguaje. Una cuestión fundamental es dónde se almacena la representación proposicional que se va generando durante la lectura o la comprensión del texto. Dicha representación podría almacenarse en memoria de trabajo a corto plazo, pero a costa de una fuerte sobrecarga de la misma. Por ello, Kintsch y sus colaboradores (ver, por ejemplo, Butcher y Kintsch, 2003) han propuesto que las habilidades fuertemente practicadas generan unas estructuras de memoria a largo plazo pero que son muy rápidamente accesibles. Estas estructuras, que no están realmente activas durante la tarea, pero son accesibles, y la información que contienen recuperable, rápidamente y sin ningún esfuerzo consciente (y que por tanto poseen características típicas tanto de la memoria de trabajo como de la memoria a largo plazo) se han dado en llamar *memoria de trabajo a largo plazo* o *LTWM* (*Long-Term Working Memory*).

tando a A y B unidos por una flecha, que representa la transición. Por otra parte, hay relaciones menos explícitas, como las que se dan entre proposiciones que definen relaciones espaciales y las relaciones espaciales mismas en la imagen. Por ejemplo, “el pueblo A está 250 kms al noroeste del pueblo B” puede estar representado en un mapa a escala por dos iconos que representan a A y B. La distancia entre esos dos puntos y su posición relativa expresada verbalmente en el texto no está marcada por ningún elemento o icono concreto en la imagen, sino por la distancia relativa entre A y B y su posición relativa en el mapa en relación a otras distancias o a una escala estandarizada, de acuerdo con ciertas convenciones. Relaciones parecidas pueden darse, por ejemplo, en el caso de las gráficas matemático-estadísticas.

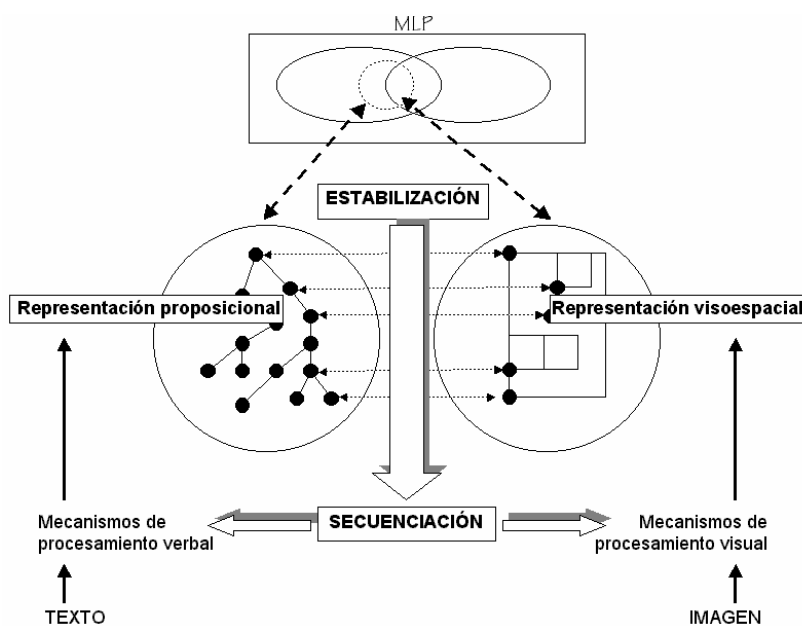
Este primer proceso, al que llamaremos *estabilización*, lleva el tiempo necesario para conseguir una configuración única y estable de relaciones entre los elementos activos en un momento determinado. El *componente automático* del proceso de estabilización puede ser llevado a cabo por una red conexionista que incluya mecanismos excitatorios de transmisión de activación junto con mecanismos inhibitorios encargados de reducir la activación residual. Una vez que una serie de elementos se han activado, las mutuas influencias inhibitorias y excitatorias darán lugar a un patrón que será estable hasta que se de paso a nueva información al sistema. Por su parte, el *componente controlado* del proceso de estabilización debe depender de un sistema de supervisión estrechamente relacionado con los procesos de selección atencional (Shallice, 1994, Posner y Badgaiyan, 1998). Esta combinación de procesos controlados y automáticos puede resultar llamativa, puesto que las aproximaciones conexionistas de la atención suelen adoptar la idea de que la atención no requiere de procesos activos, sino que es fruto de la dinámica misma del sis-

tema (atención como resultado). Sin embargo, la posibilidad de combinar mecanismos asociativos y mecanismos de control está siendo recientemente explorada en áreas como la atención (Gilbert y Shallice, 2002) y el aprendizaje (Perales, Catena y Maldonado, 2002a, b).

El proceso de estabilización tiene dos resultados posibles, uno positivo -la consecución de una representación distribuida estable-, y uno negativo -la imposibilidad de alcanzar tal representación-, por la existencia de inconsistencias o conflictos en la información activa, o bien porque existe una ambigüedad irreductible en la interpretación de dicha información. Este segundo resultado es una de las condiciones que, se sabe, involucran más directamente a los procesos ejecutivos; conlleva la toma de conciencia de un conflicto, y activa zonas concretas de la corteza cerebral (concretamente, el tercio anterior del giro del cíngulo, en la corteza cerebral medial, Posner y DiGirolamo, 1998).

A continuación entrará en juego un segundo proceso, que hemos llamado de *secuenciación*. Una vez que se alcanza una representación distribuida estable de la información procesada hasta el momento se da entrada a una nueva pieza de información, que dará lugar a que se ponga en marcha de nuevo el proceso de estabilización. En el caso negativo, el proceso de secuenciación dará entrada a una pieza de información ya procesada, bien en el texto, bien en la imagen, para buscar interpretaciones alternativas que reduzcan la ambigüedad o el conflicto.

En resumen, el modelo propone una alternancia estabilización-secuenciación que va dando lugar a una representación progresivamente más elaborada, más abstracta y menos ambigua. Si se analiza el patrón conductual del sujeto, lo que se observará es que atiende alternativamente a la imagen y al texto durante toda la tarea, con retrocesos puntuales a partes de la información que ya habían sido procesados anteriormente.



**Figura 1:** Diagrama del modelo de estabilización-secuenciación. La representación visoespacial y la representación proposicional, simultáneamente activas en la memoria de trabajo entran a formar parte, junto con la información proposicional-semántica y la información espacial-estructural de la memoria a largo plazo de una representación distribuida. Una vez que dicha representación distribuida se hace estable (estabilización), se da paso a una nueva pieza de información entrante (secuenciación).

Este modelo es a grandes rasgos aplicable también a situaciones multimedia narración-imagen, con algunas matizaciones. Puesto que la narración y la imagen no comparten los mismos canales sensorio-perceptivos, no es necesario que el sujeto deba alternar entre texto e imagen para dar entrada simultánea a las dos fuentes de información. Además, en estas situaciones las posibilidades de interacción del sujeto con la narración son pequeñas. En una comunicación estándar, en la que el didacta habla directamente a sus estudiantes ayudándose de imágenes, el oyente tiene la posibilidad de interrumpir y preguntar pero, aún así, no puede controlar el avance y retroceso sobre la información como puede hacer en el caso de la lectura. Por tanto, en esta situación el proceso de secuenciación no entraría en juego, y el sujeto debe aplicar el proceso de estabilización, en sus dos vertientes, controlada y automática, en tiempo real, estableciendo conexiones entre los elementos conceptuales y proposicionales de la narración y los elementos de la imagen conforme dichos componentes van apareciendo. Entendemos que aunque la pérdida de control dificulta en cierta medida la tarea, en esta situación se dispone de la gran ventaja para el didacta de poder potenciar el procesamiento activo a través de la atención dirigida, manipulando los elementos de la imagen y de su propio discurso de acuerdo a la estructura cognitiva del sujeto descrita en este apartado.

### Ventajas del modelo de estabilización-secuenciación

En primer lugar, el modelo propuesto, a diferencia de los anteriores, es específico sobre como actúan los procesos encargados de establecer conexiones entre texto e imagen en situaciones didácticas. Además, lo hace de forma verosímil, de acuerdo con los descubrimientos recientes de la Psicología Básica, la Neuropsicología y los estudios de neuroimagen, que animan a sustituir la 'metáfora del ordenador' por la 'metáfora del cerebro'. La interpretación de los diagramas de flujo, tan frecuentes en la disciplina, se da en un sentido funcional, más que estructural, y se proponen de hecho estructuras más congruentes con la configuración real del cerebro humano.

En segundo lugar, el modelo describe qué hay realmente de activo en el llamado procesamiento activo, cuya importancia se había señalado en los modelos anteriores, y se relaciona aquí directamente con los procesos ejecutivos. Es más, se supera la crítica habitual que se hace al concepto de ejecutivo central —que reproduce el problema del homúnculo—, diseccionando hasta donde es posible las funciones que éste cumple en estas tareas. La separación de los dos procesos, secuenciación y estabilización, y la descripción de sus condiciones de aplicación, reducen en gran medida la ambigüedad de los términos comúnmente utilizados para referirse a los procesos ejecutivos.

Sin embargo, desde el punto de vista teórico, la mayor ventaja del modelo es que permite incorporar con facilidad los resultados experimentales encontrados hasta la fecha, y permite explicar otros que resultaban difíciles de acomodar en los modelos anteriores. Además, permite realizar predicciones concretas que pueden ponerse a prueba en futuras investigaciones.

Por ejemplo, los efectos de orden en la presentación de texto e imagen tiene ahora una explicación unívoca. Tal y como

demuestran los estudios antes citados (ver Mayer y Moreno, 2002, para una revisión), no es igualmente eficaz presentar el texto (o una narración) de forma simultánea a la imagen que secuencialmente, y no es idéntico presentar la imagen antes o después. Si la imagen se presenta antes del texto, se está obligando al sistema a codificarla y retenerla como una unidad con todas las especificaciones posibles, y a ocupar toda la capacidad de la agenda visoespacial y una parte de los recursos ejecutivos en esa tarea. En cualquier caso, mientras la lectura avanza, los distintos conceptos y acciones referidas en el texto pueden ir relacionándose con la representación almacenada de la imagen. Puesto que normalmente es la imagen la que sirve para reducir la ambigüedad de la información verbal (sobre todo si el texto contiene proposiciones relativas a relaciones espaciales), y no al contrario, la imagen almacenada sirve de guía para interpretar de forma unívoca el contenido del texto. Esto no ocurre si la imagen se presenta al final. En ese caso, el sujeto debe mantener simultáneamente durante la lectura varias interpretaciones posibles del texto (o una representación inestable) con la sobrecarga del sistema que eso conlleva; o decidirse por una, lo cual conlleva el riesgo de que dicha interpretación entre en conflicto con la imagen, sin posibilidad de volver atrás en el texto, y sin la posibilidad, por tanto, de alcanzar una representación distribuida estable.

La presentación serial imagen-texto es, por tanto, en la mayor parte de los casos (cuando la imagen se utiliza para reducir la ambigüedad del texto) más eficaz que la presentación texto-imagen. En cualquier caso, la presentación que permite alternar entre texto e imagen, o mantener la representación verbal y visoespacial simultáneamente en la memoria de trabajo sin que se produzca interferencia entre ambas es preferible a la secuencial. Es más, si el sujeto no tiene control sobre el ritmo de acceso a la información verbal (por ejemplo, en una narración acompañada de una animación), es preferible que la información verbal y visual no compartan código, para no entrar en competencia por los mismos canales. El más eficaz de los casos es aquel en el que el sujeto puede determinar el ritmo de presentación de la información, alternando libremente entre texto e imagen, y avanzando adelante y atrás en el texto. Ello supone una máxima economía de los subsistemas de la memoria trabajo, una interferencia mínima y una máxima eficacia de los procesos ejecutivos (Verdi, Johnson, Stock, Kulhavy y Whitman-Ahern, 1997, Mayer y Sims, 1994, Mayer, Heiser y Lonn, 2001).

Entre los datos aparentemente contraintuitivos que el modelo consigue explicar están la correlación que existe entre la dificultad del texto y el número de consultas que el aprendiz hace de la imagen. Se ha demostrado que los textos difíciles provocan un mayor número de referencias a la imagen que los textos fáciles, y que los malos aprendices tienen una distribución de referencias a la imagen distinta a la de los buenos aprendices. Esto se ha demostrado en estudios que recogen protocolos verbales o hacen el seguimiento de algún correlato conductual del funcionamiento cognitivo del aprendiz durante la tarea (Schnotz, Picard y Hron, 1993). Datos preliminares procedentes de nuestro laboratorio (Perales y Romero, 2003) indican que, en una presentación informatizada texto-imagen, en la que el sujeto puede visualizar un diagrama en formato "pop-up" en cualquier

punto a lo largo de la lectura de un texto, indican que los buenos aprendices no hacen un mayor número de visualizaciones de la imagen, sino que las distribuyen, en el primer quinto de la tarea, mientras los malos aprendices, por el contrario, no consultan la imagen en la primera parte de la tarea, sino más adelante, cuando la lectura del texto ya se encuentra avanzada. Estos datos son congruentes con la noción de estabilidad propuesta por nuestro modelo: el establecimiento de relaciones tempranas entre los componentes del texto y de la imagen constriñe la interpretación del texto, y permite generar una representación más económica.

Por último, el modelo resalta la importancia del procesamiento activo como búsqueda dirigida de relaciones entre los elementos del texto, la imagen, y el conocimiento previo, y de la búsqueda de referentes en la imagen de las relaciones expresadas en el texto. Lógicamente, el sujeto puede recibir entrenamiento en la materia sobre la que versa el texto, u orientado a conseguir una alfabetización visual suficiente (Peeck, 1993), pero también puede ser entrenado expresamente en las estrategias metacognitivas que dan lugar a un mejor aprovechamiento de la imagen. Dichas estrategias van encaminadas, sobre todo, a establecer relaciones de forma controlada entre los elementos del texto y la imagen y a comprobar periódicamente su verosimilitud (Moore, 1993, Moore y Schofield, 1994).

Por tanto, en resumen, el modelo señala la importancia del procesamiento activo como el establecimiento controlado de las relaciones entre los distintos elementos activos en un momento determinado. También indica que el estudio exclusivo de los resultados del aprendizaje (con contadas excepciones, como el estudio de Schnotz, Picard y Hron, 1994), sin atender al comportamiento del aprendiz durante la tarea provoca la pérdida de información valiosa sobre el desarrollo de los procesos de comprensión en tiempo real. Por tanto, el análisis de los resultados del aprendizaje debería combinarse con el seguimiento de la conducta del sujeto en tiempo real. Una fuente de información posible son los protocolos verbales, aunque estos pueden provocar una desnaturalización de los procesos que se pretende estudiar. En el futuro será necesario considerar medidas objetivas tales como las respuestas de seguimiento ocular, las alternancias texto-imagen, los retrocesos-avances en la lectura del texto, o el análisis cualitativo de la distribución de los puntos del texto en los que el sujeto espontáneamente consulta la imagen adjunta. El resultado de esta línea de investigación debería llevar a la identificación de patrones que diferencien a los buenos y malos aprendices, y, por tanto de los aspectos que pueden corregirse (Perales y Romero, 2003).

### Algunas consideraciones evolutivas

El presente trabajo está centrado principalmente en las capacidades del sistema cognitivo adulto. Sin embargo, es conveniente tener en cuenta que en una buena parte de los casos, el contexto de enseñanza-aprendizaje involucra a individuos en distintas etapas de la niñez y la adolescencia. Aunque describir exactamente los cambios reseñables que experimentan los sistemas perceptivos, de memoria y atencionales que pueden afectar al procesamiento conjunto lenguaje-imagen está mucho más allá de los objetivos de esta revisión, enumeraremos en esta sección algunas consideraciones que conviene tener en cuenta.

Las habilidades perceptivas sufren un cambio importante en los primeros años de vida. Ya antes de la edad escolar, sin embargo, podemos afirmar que el sistema perceptivo es estructuralmente idéntico al del adulto. Incluso los componentes pictórico y espacial de la memoria de trabajo se encuentran diferenciados entre sí, y de otros componentes de la memoria de trabajo (Isaacs y Vargha-Khadem, 1989, Pickering *et al.*, 1997, citado por Gathercole, 1998). En la edad infantil y los primeros años escolares será necesario aún que evolucionen las capacidades de categorización, que adquirirán un carácter progresivamente más conceptual y funcional, y menos apegado a las características sensoriales superficiales (Cohen y Cashon, 2003); y también están sometidas a cambio las estructuras de almacenamiento semántico (fundamentalmente esquemas de conocimiento) que son aplicables a la interpretación de patrones visuales. Como ya comentamos anteriormente, los niños en los primeros años de primaria tienen habilidades de interpretación de imágenes que van poco más allá de las representaciones figurativas (con un alto contenido icónico) y algunos tipos de modelos esquematizados (Schnotz, 2002, Shah y Hoeffner, 2002, Lowe, 1993b, Schofield y Kirby, 1994, Kirby, 1985, citado por Moore, 1993).

Una buena parte de este proceso de desarrollo —de carácter más funcional que estructural—, tiene que ver con la alfabetización visual. Aunque en parte ésta se adquiere de forma incidental, sin esfuerzo consciente (por ejemplo, la mayoría de los adultos sabemos interpretar intuitivamente un mapa meteorológico sin haber recibido ninguna educación formal al respecto), también es cierto que dicho proceso de aprendizaje debería recibir más atención en los currícula escolares.

Algo parecido puede decirse respecto de la estructura de la memoria. Los estudios realizados con niños verbales, de tres años en adelante, muestran una organización de la memoria similar a la del adulto (para una revisión, ver Gathercole, 1998). A partir de esta edad ya puede demostrarse la existencia de un componente fonológico-auditivo y otro visual de la memoria a corto plazo, como demuestra el hecho de que los niños sean sensibles tanto a los efectos de longitud de la palabra y similitud fonológica cuando el material de recuerdo a corto plazo es verbal, como a la interferencia con tareas visoespaciales cuando el material de recuerdo son figuras o dibujos (Ford y Silber, 1994, Gathercole y Adams, 1994, Hitch, Halliday, Dodd y Littler, 1989, Hitch, Halliday, Schaafstal y Schraagen, 1988). Sin embargo, el funcionamiento de estos sistemas y, sobre todo, su coordinación está aún lejos de mostrar las características adultas.

Por ejemplo, los niños preescolares desconocen que el contenido de su memoria a corto plazo va a decaer y desaparecer en pocos segundos si no hace poco por retenerlo. La forma más sencilla de evitar este decaimiento, el repaso por repetición, no aparece hasta poco antes de los siete años (algo antes si se trata de material significativo y el contexto es lo suficientemente motivante, Johnston, Johnson y Gray, 1997). Por otra parte, numerosos estudios demuestran que la capacidad real del lazo fonológico alcanza un desarrollo temprano; sin embargo el incremento progresivo de la eficiencia de los intercambios de información entre el lazo fonológico y la memoria semántica continúa de forma ininterrumpida hasta la adolescencia (Roodenrys, Hulme y Brown, 1996). Este hecho explica por qué la ejecución en tareas de amplitud de palabras no deja de crecer durante todo ese periodo (Hulme, Muir, Thomson y Lawrence, 1984).

Los cambios más notables en la memoria humana en desarrollo ocurren en la estructuración de la información en la memoria a largo plazo y en los procesos ejecutivos. La entrada en la edad escolar supone, por una parte, el comienzo de una serie de procesos de cambio conceptual que deben modificar notablemente las teorías intuitivas que el niño tiene sobre el mundo (Vosniadou, 1996), y por otra, la adquisición gradual de habilidades procedimentales motoras y cognitivas (entre ellas, fundamentalmente, la lectoescritura), que producen un importantísimo cambio, no sólo cuantitativo, sino también cualitativo, en sus capacidades de comprensión y memorización.

Los procesos ejecutivos, por su parte, se desarrollan lentamente. Los indicadores objetivos de procesamiento ejecutivo indican una evolución constante hasta la adolescencia tardía (Siegel, 1994), paralelamente a los procesos de mielinización de los lóbulos prefrontales. En realidad, todos los índices de funcionamiento de memoria a corto plazo (por ejemplo amplitud de dígitos, presencia de repaso, amplitud visoespacial, tareas de bloques) están contaminados de una u otra forma por el funcionamiento ejecutivo. Dado el papel central que estos procesos tienen en la integración lenguaje-imagen, el estudio su desarrollo debería ser un punto clave en esta área. Sin embargo, prácticamente ningún trabajo empírico se ha centrado en cómo el desarrollo de los procesos ejecutivos afecta a las capacidades de integración entre formatos de información o modalidades diversas (ver, sin embargo, Moore, 1993).

Del modelo propuesto se deriva que los procesos ejecutivos son imprescindibles en los componentes activos, autoiniciados, de la tarea de integración. Si el nivel de desarrollo de los mismos es aún bajo, el establecimiento de lazos referenciales debe ser un proceso guiado por el didacta. Por tanto, las interrelaciones entre texto e imagen deben ser lo más expresas posible. A esto se une que en los dos primeros años de la edad escolar (6-7 años) muchos niños aún carecen de estrategias que les permitan mantener de forma prolongada y manipular la información de los almacenes fonológico y visoespacial, y, por tanto, se hace evidente que el principio de contigüidad adquiere aún más importancia que en los sujetos adultos.

## Conclusión: Principios para el desarrollo de aplicaciones didácticas

La revisión realizada en este trabajo tenía como objetivo resumir la literatura sobre procesamiento conjunto de lenguaje e imagen en contextos didácticos, y ofrecer un marco teórico coherente que dotara de sentido a dicha información. Dicho marco teórico permite superar una estrategia de intervención basada en manipulaciones aisladas y más o menos azarosas por una manipulación sistemática que permita aislar variables relevantes en la cadena de procesamiento.

La enumeración de dichas variables debe primero tener en cuenta el procesamiento del discurso y la imagen por separado. En términos generales, hay que partir del hecho de que existen importantes variantes interindividuales en las aptitudes de procesamiento. Variables como la capacidad eidética, o aquellas que dependen de la madurez evolutiva del individuo, deben ser tenidas en cuenta, pero no son susceptibles de manipulación. Otras, como la alfabetización visual o la aptitud lectora si son suscepti-

bles de intervención. En el caso de la alfabetización visual, el sujeto de intervención no es sólo el aprendiz, sino también el didacta, para que éste adquiera las habilidades de generación de imágenes utilizando un código o lenguaje común con el aprendiz, y teniendo en cuenta sus habilidades de interpretación intuitivas.

En caso de encontrarnos ante un individuo o un grupo con baja aptitud lectora, debe procurarse utilizar textos simples, y dar un mayor peso a la imagen, teniendo en cuenta que el sujeto va a invertir una mayor carga de procesamiento en ella. Ello implica la necesidad de que exista un fuerte solapamiento entre texto e imagen. Por otra parte, ante individuos con una baja alfabetización visual, o con esquemas visuales pobres, debe prestarse especial atención a aquellos elementos de la imagen que tienen un valor más arbitrario que icónico. Debe dedicarse esfuerzo a explicar el sentido de los componentes de la imagen, esto es, a hacer explícito qué representan los distintos elementos, tanto conceptuales como funcionales, y cómo su disposición puede tener una carga de significado. También debe cuidarse especialmente en este caso la coherencia de la imagen, utilizando las leyes de la Gestalt para agrupar a los elementos que tienen un rasgo común, especialmente si dicho rasgo tiene un cierto peso en la organización del discurso.

Estos factores no sólo van a facilitar la interpretación de la imagen, sino que también van a decrementar el uso de recursos ejecutivos y a facilitar el recuerdo. En tanto que la imagen se codifique de forma unitaria y significativa, los elementos eidéticos de dicha representación van a tener una mayor potencia como clave de recuerdo. En otras palabras, puesto que la representación final es una representación distribuida, su reactivación será tanto mejor cuanto más elaborada y distintiva sea su codificación.

Al igual que es necesario maximizar la influencia de los factores que facilitan un procesamiento eficiente de lenguaje e imagen, también es importante minimizar la de los que lo dificultan. Frecuentemente, el didacta reduplica la información, utilizando texto y narración oral simultáneamente (la lectura en voz alta por parte del profesor del libro de texto, mientras el alumno también lee, es un recurso muy utilizado en el aula), añade detalles pretendidamente interesantes, aunque irrelevantes, con el objetivo de mantener el interés, o añade un gran número de detalles a sus imágenes didácticas. Como ya hemos visto, se ha demostrado suficientemente que estas estrategias tienen un coste cognitivo importante, puesto que producen interferencia en los canales perceptivos, y, por tanto, deben utilizarse de forma muy cuidadosa.

En cualquier caso, finalmente vamos a encontrarnos con una colección de elementos activos en memoria de trabajo. Parte de estos elementos provienen del texto, fundamentalmente conceptos y proposiciones; parte de ellos pertenecen a la imagen, sobre todo elementos icónicos y estructurales; y otra parte incluye elementos tanto proposicionales-semánticos como estructurales-espaciales previamente almacenados en la memoria a largo plazo. El objetivo último es conseguir la integración efectiva de todos esos elementos.

Algunas variables que tienen efecto a este nivel son variables del material. Utilizar presentaciones "pop-up", posibilitar la disponibilidad simultánea de texto e imagen, o maximizar la contigüidad temporal y espacial entre los elementos de uno y de otra

son estrategias que se han demostrado eficaces para posibilitar la presencia conjunta elementos del texto y la imagen en la memoria de trabajo. Otros factores tienen su efecto en la reducción del consumo general de recursos atencionales. Como vimos anteriormente, el uso de claves perceptivas (como, por ejemplo, colores) para establecer conexiones directamente del texto a la imagen es más eficaz que el uso de leyendas, que requieren un esfuerzo de interpretación extra. De forma similar, la integración de lenguaje e imagen se dificulta si el etiquetado de los elementos de la imagen se hace a través de llamadas, y no sobre la imagen misma.

Otras variables son fundamentalmente variables de tarea, y estas son las que el didacta puede utilizar con mayor flexibilidad. Las variables que tienen un efecto más potente son las que posibilitan que el sujeto adopte un papel activo (involucre procesos atencionales controlados) en el establecimiento de lazos referenciales entre texto e imagen.

Los lazos referenciales constriñen la interpretación semántica del léxico y las proposiciones contenidas en el discurso, dando lugar a elementos representacionales más unívocos y, por tanto, más distintivos, y a representaciones distribuidas más estables.

Tanto una parte del proceso de estabilización como el proceso de secuenciación son controlados, y, por tanto, son sensibles al uso de estrategias. Estas, a su vez, son susceptibles de aprendizaje. Las comparaciones entre 'buenos' y 'malos' aprendices están empezando a señalar una serie de diferencias importantes en el uso de estrategias. Los datos indican que el uso de la imagen desde el comienzo de la tarea (entre los 200-300 segundos después del comienzo de la lectura según los datos de nuestro laboratorio), es uno de los indicadores más precisos de un buen aprendizaje. Otro predictor importante es el índice de lectura retrospectiva inducida por la imagen, esto es, el regreso a fragmentos del texto ya leídos cuando se detecta una incongruencia entre lo que se cree saber y el contenido de la imagen. La primera variable puede utilizarse como un índice del uso estratégico del componente controlado del proceso de estabilización contemplado en nuestro modelo. La segunda como un índice de la capacidad para detectar las condiciones que requieren la puesta en marcha del proceso de secuenciación.

Esta línea de investigación, centrada en índices objetivos de procesamiento estratégico, resulta extraordinariamente prometedora por sus implicaciones prácticas, ya que puede ofrecer las claves para una intervención centrada en los aspectos centrales de las tareas de integración lenguaje-imagen.

Por último, un aspecto de la integración lenguaje-imagen que no ha sido suficientemente resaltado en la literatura es la estrecha interacción que se da entre la tarea de aprendizaje (codificación) y la aplicación de lo aprendido (recuperación). Si se produce una verdadera integración entre los elementos representacionales activos en los distintos sistemas implicados en el procesamiento, se producen una mejor organización, elaboración y abstracción del contenido del texto, a la vez que los elementos estructurales-espaciales almacenados sirven como clave de recuerdo. Por tanto, las tareas de recuperación de naturaleza puramente verbal (exámenes orales o escritos, pruebas objetivas, etc.), la transferencia a la solución de problemas y las pruebas de carácter más visoespacial (por ejemplo, rellenar los espacios en blanco de un mapa topográfico) pueden beneficiarse de la pre-

sentación de una imagen durante el aprendizaje. En términos generales, hay tareas que requieren una codificación puramente declarativa o proposicional; otras se ven máximamente beneficiadas por una codificación visual-espacial; por último, las tareas que requieren la puesta en práctica, una comprensión profunda del material, y una verdadera reestructuración conceptual, requieren la integración de ambos formatos. Se ha demostrado que la presentación de una imagen puede incluso dañar el recuerdo de la información verbal recogida en el texto que no tiene un correlato en la imagen (Schnotz, 2002). En la misma línea, si se pretende enseñar a un aprendiz a adquirir una serie de habilidades, la imagen puede facilitar estrategias de aprendizaje alternativas que le impidan aprender las habilidades objetivo. Por ejemplo, en el caso de las habilidades lectoescritoras, utilizar ilustraciones figurativas puede retrasar el aprendizaje, ya que el niño puede utilizar estrategias de comprensión alternativas a las necesarias para la lectoescritura (Fang, 1996, Carney y Levin, 2002).

La especificidad en la codificación es un serio problema a la hora de producir un verdadero cambio conceptual. Puesto que en los métodos didácticos más tradicionales tanto en la codificación de la información como en las pruebas de recuerdo se hace mayor hincapié en el aprendizaje verbal, a menudo pasa desapercibido que la información que se almacena a largo plazo con formatos representacionales no proposicionales (estructurales-espaciales) no se ve afectada por el proceso de aprendizaje. Por ejemplo Vosniadou y sus colaboradores (ver, por ejemplo, Samarapungavan, Vosniadou y Brewer, 1996, Vosniadou, 1993) han demostrado repetidamente que los niños integran la información puramente verbal que reciben sobre contenidos astronómicos (p.e. "la tierra es redonda", "la tierra gira alrededor del sol"... ) con la información perceptual a la que están expuestos en la vida cotidiana. A partir de la integración de estos dos formatos de información, el niño construye modelos mentales internamente congruentes pero que no se corresponden con los modelos científicamente correctos (por ejemplo, pueden construir un modelo mental en el que la tierra es redonda y achatada, y la que las personas viven sólo en la parte de arriba). Si la prueba de recuperación no evalúa los aspectos estructurales y espaciales del conocimiento del niño, es posible que sus respuestas 'parezcan' correctas. Si a un niño se le informa de que la tierra es redonda, y en la prueba de recuerdo posterior se le pregunta de qué forma es la tierra, y éste contesta 'redonda', dicha respuesta se dará por válida, aunque el modelo mental que hay en su base sea incorrecto. El verdadero cambio conceptual se dará sólo si el aprendiz reorganiza los aspectos estructurales y espaciales de su conocimiento en concordancia con la información verbal facilitada. De ahí, que las ilustraciones y los modelos tengan un papel fundamental como complemento en el aprendizaje de contenidos científicos (Chi, Slotta y de Leeuw, 1994).

En resumen, como señalamos en la sección dedicada al marco computacional de nuestro modelo, el objetivo último del uso de imágenes como complemento al discurso es generar una representación distribuida, estable y funcional. El término estable implica que dicha representación ha de ser unitaria y congruente en sí misma, y suficientemente distintiva y elaborada como para recuperarse de forma lo más completa posible cuando se requiera. Por su parte, el término funcional indica que dicha representación debe ajustarse a las demandas de la tarea que se ha de

realizar, de acuerdo con el principio de especificidad de la codificación. El modelo propuesto señala un camino para diseñar estrategias que permitan cumplir estos dos objetivos.



## Referencias

- Adler, C. (1993). Directed picture processing: The effects for learners on recall of related text. *Dissertation Abstracts International. Section A: The Sciences and the Engineering*, 54 (3-A), 863.
- Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, 33, 131-152.
- Anderson, J.R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Atkinson, R.C. y Shiffrin, R.M. (1974). Human memory: a proposed system and its control processes. In Kenneth W. Spence y Janet T. Spence (eds.), *The Psychology of Learning and Motivation: II*. Oxford: Academic Press.
- Baddeley, A.D. (2002). Is working memory still working?. *European Psychologist*, 7, 85-97.
- Baddeley, A.D., Della-Sala, S., Papagno, C. y Spinnler, H. (1997). Dual task performance in dysexecutive and nondysexecutive patients with a frontal lesion. *Neuropsychology*, 11, 187-194.
- Baddeley, A.D. y Hitch, G.J. (1974). Working Memory. In G.H. Bower (ed.), *The Psychology of Learning and Motivation: VIII*. New York: Academic Press.
- Baddeley, A.D. y Hitch, G.J. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485-493.
- Bartram, D.J. (1980). Comprehending spatial information: The relative efficiency of different methods of presenting information about bus routes. *Journal of Applied Psychology*, 65, 103-110.
- Beagles-Roos, J. y Gat, L. (1983). Specific impact of radio and television on children's story comprehension. *Journal of Educational Psychology*, 75, 128-137.
- Betrancourt, M. y Bisseret, A. (1998). Integrating textual and pictorial information via pop-up windows: An experimental study. *Behavior and Information Technology*, 17, 263-273.
- Biederman, I. (2000). Recognition by components: A theory of human image understanding. In Steven Yantis (ed.), *Visual Perception: Essential readings. Key readings in cognition*. Philadelphia: Psychology Press.
- Brooks, L.R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 22, 349-368.
- Butcher, K.R. y Kintsch, W. (2003). Text comprehension and discourse processing. In A.F. Healy y R.W. Proctor (eds.), *Handbook of Psychology: Experimental Psychology. Vol. IV*. New York: Wiley and Sons.
- Carney, R.N., y Levin, J.R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5-26.
- Chase, W.G. y Simon, H.A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M.T.H., Slotta, J.D. y de Leeuw, N. (1994). From things to processes: A theory for conceptual change for learning science concepts. *Learning and instruction*, 4, 27-43.
- Chmielewski, T.L., Dansereau, D.F. y Moreland, J.L. (1998). Using common region in node-link displays: The role of field-dependence/independence. *Journal of Experimental Education*, 66, 197-207.
- Clark, J.M. y Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3, 149-210.
- Cohen, J.D., Forman, S.D., Braver, T.S., Casey, B.J., Servant-Schreiber, D. y Noll, D.C. (1994). Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Human Brain Mapping*, 1, 293-304.
- Cohen, L.B. y Cashon, C.H. (2003). Infant perception and cognition. In R.M. Lerner (ed.), *Handbook of Psychology: Developmental Psychology. Vol. VI*. New York: Wiley and Sons.
- Costa, J. (1998). *La esquemática. Visualizar la información*. Barcelona: Paidós Ibérica.
- Courtney, S.M., Ungerleider, L.G., Keil, K. y Haxby, J.V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex*, 6, 39-49.
- Craik, F.I. y Lockhart, R.S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Dean, R.S., y Enemoh, P.A.C. (1983). Pictorial organization in prose learning. *Contemporary Educational Psychology*, 8, 20-27.
- Dean, R.S., y Kulhavy, R.W. (1981). Influence of spatial organization on prose learning. *Journal of Educational Psychology*, 73, 57-64.
- Eco, U. (1982). *Lector in fábula*. Barcelona: Lumen.
- Egan, D.E. y Schwartz, B.J. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition*, 7, 149-158.
- Fang, Z. (1996). Illustrations, texts, and the child reader. What are pictures in children story-books for?. *Reading Horizons*, 37, 130-142.
- Ford, S. y Silber, K.P. (1994). Working memory in children: A developmental approach to the phonological coding of pictorial material. *British Journal of Developmental Psychology*, 12, 165-175.
- Funahashi, S., Bruce, C.J. y Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61, 331-349.
- Garrison, W.T. (1978). The context bound effects of picture-text amalgam: Two studies. *Dissertation Abstracts International. Section A: The Sciences and the Engineering*, 39, 4137.
- Gathercole, S.E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, 39, 3-27.
- Gathercole, S.E. y Adams, A. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language*, 33, 672-688.
- Gathercole, S.E. y Baddeley, A.D. (1997). Sense and sensitivity in phonological memory and vocabulary development: A reply to Bowery. *Journal of Experimental Child Psychology*, 67, 290-294.
- Gattis, M. y Holyoak, K.J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 231-239.
- Georgopoulos, A.P., Taira, M., Lukashin, A., Wurtz, R.H., Tach, W.T., Mink, J.W., Goodkin, H.P., Keating, J.G., Norman, D.A., Shallice, T., Goldman-Rakic, P. (2000). Action and Executive Function. In Michael S. Gazzaniga, *Cognitive Neuroscience: A reader*. Malden, USA: Blackwell.
- Gilbert, S.J. y Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology*, 44, 297-337.
- Goldman-Rakic, P.S. (1999). Neural basis of working memory. In R.A. Wilson y F.C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*. London: MIT Press.
- Griffin, M.M. (1995). You can't get there from here: Situated learning transfer and map skills. *Contemporary Educational Psychology*, 20, 65-87.
- Gyselinck, V., Ehrlich, M.F., Cornoldi, C., de Beni, R. y Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer-Assisted Learning*, 16, 166-176.
- Gyselinck, V., Cornoldi, C., Dubois, V., De Beni, R. y Ehrlich, M.F. (2002). Visuospatial memory and phonological loop in learning from multimedia. *Applied Cognitive Psychology*, 16, 665-685.
- Hall, R.H., y O'Donnell, A.M. (1996). Cognitive and affective outcomes of learning from knowledge maps. *Contemporary Educational Psychology*, 21, 94-101.
- Hitch, G.J., Halliday, M.S., Dodd, A. y Littler, J.E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, 7, 347-362.
- Hitch, G.J., Halliday, M.S., Schaafstal, A.M. y Schraagen, J.M.C. (1988). Visual working memory in young children. *Memory and Cognition*, 16, 120-132.
- Hoffman, D.D. (1998). *Visual Intelligence: How we create what we see*. New York: Norton.
- Hulme, C., Muir, C., Thomson, N. y Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology*, 38, 241-253.
- Isaacs, E.B. y Vargha-Khadem, F. (1989). Differential course of spatial and verbal memory span: A normative study. *British Journal of Developmental Psychology*, 7, 377-380.
- Johnson-Laird, P.N. y Byrne, R.M.J. (1991). *Deduction*. Hillsdale: Erlbaum.
- Johnston, R.S., Johnson, C. y Gray, C. (1987). The emergence of the word length effect in young children: The effects of overt and covert rehearsal. *British Journal of Developmental Psychology*, 5, 243-248.

- Jonides, J., Smith, E.E., Koeppe, R.A., Awh, E., Minoshima, S. y Mintum, M.A. (1993). Spatial working memory in humans as revealed by PET. *Nature*, 363, 623-625.
- Just, M.A. y Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 4, 329-354.
- Kalyuga, S., Chandler, P. y Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-371.
- Kalyuga, S., Chandler, P. y Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, 92, 126-136.
- Kandel, E. (2003). Pastillas contra el olvido. Entrevista concedida a Der Spiegel, traducida y reproducida en *El País*, 11-05-2003.
- Kessels, R.P.C., van Zandvoort, M.J.E., Postma, A., Kapelle, L.J. y de Haan, E.H.F. (2000). The Corsi block tapping task: Standardization and normative data. *Applied Neuropsychology*, 7, 252-258.
- Kintsch, W. (1988). The role of knowledge in discourse: A construction-integration. *Psychological Review*, 95, 163-182.
- Kintsch, W. y Van Dijk, T.A. (1978). Towards a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Kintsch, W., Welsch, D., Schmalhofer, F. y Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, 29, 133-159.
- Kirby, J.R. (1993). Collaborative and competitive effects of verbal and spatial processes. *Learning and Instruction*, 3, 201-214.
- Kosslyn, S.M. y Rabin, C. (1999). The representation of left-right orientation: A dissociation between imagery and perceptual recognition. *Visual Cognition*, 6, 497-508.
- Kosslyn, S.M., Flynn, R.A., Amsterdam, J.B. y Wang, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition*, 34, 203-277.
- Kruey, P., Sciana, S.C. y Glenberg, A.M. (1994). On-line processing of textual illustration in the visual sketchpad: Evidence from dual-task studies. *Memory and Cognition*, 22, 261-272.
- Kulhavy, R.W., Lee, J.B. y Caterino, L.C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology*, 10, 28-37.
- Kulhavy, R.W., Pridemore, D.R. y Stock, W.A. (1992). Cartographic experience and think aloud about thematic maps. *Cartographica*, 291, 1-9.
- Kulhavy, R.E., Stock, W.A. y Kealy, W.A. (1993). How geographic maps increase recall of instructional texts. *Educational Technology Research and Development*, 41, 47-62.
- Kulhavy, R.W., Stock, W.A., Verdi, M.P., Rittschotf, K.A. y Saveyne, W.A. (1993). Why maps improve memory for texts: The influence of structural information on working memory operations. *European Journal of Cognitive Psychology*, 5, 375-392.
- Larkin, J.H. y Simon, H.A. (1987). Why a diagram is (sometimes) worth a thousand words. *Cognitive Science*, 11, 65-99.
- Levie, W.H., y Lentz, R. (1982). Effects of texts illustrations: A review of research. *Educational Communication and Technology Journal*, 30, 195-232.
- Levin, J.R., Anglin, G.J. y Carney, R.N. (1987). On empirically validating functions of pictures in prose. En, D.M. Willows y H.A. Houghton (eds.), *The Psychology of Illustration: I. Basic Research*. New York: Springer.
- Levin, J.R. y Mayer, R.R. (1993). Understanding illustrations in text. En, B.K. Britton, A. Woodward y M. Brinkley (eds.), *Learning from textbooks*. Hillsdale: Erlbaum.
- Lowe, R.K. (1993a). Diagrammatic information: techniques for exploring its mental representation and processing. *Information Design Journal*, 7, 3-17.
- Lowe, R.K. (1993b). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction*, 3, 157-179.
- McCarthy, G., Blamire, A.M., Puce, A., Nobre, A.C., Bloch, G., Hyder, F., Goldman-Rakic, P.S. y Shulman, R.G. (1994). Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. In *Proceedings of the National Academy of Sciences*, 91, 8690-8694.
- McCarthy, R.A. y Warrington, E.K. (1990). *Cognitive Neuropsychology: A clinical introduction*. San Diego: Academic Press.
- Marr, D.R. (1982). *Vision*. New York: Freeman.
- Mayer, R.E. (1997). Multimedia learning: Are we asking the right questions?. *Educational Psychologist*, 32, 1-19.
- Mayer, R.E. (2002). Using illustrations to promote constructivist learning from science text. In José Otero y José León (eds.), *The Psychology of Science Text Comprehension*. Mahwah: Erlbaum.
- Mayer, R.E. y Anderson, R.B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R.E. y Gallini, J.K. (1990). When is an illustration worth ten thousand words?. *Journal of Educational Psychology*, 82, 715-726.
- Mayer, R.E., Heiser, J., y Lonn, S. (2001). Cognitive constraints on multimedia learning: when presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187-198.
- Mayer, R.E. y Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90, 312-320.
- Mayer, R.E. y Moreno, R. (2002). Animation as an aid to multimedia learning. *Journal of Educational Psychology*, 14, 87-99.
- Mayer, R.E., Moreno, R., Boire, M. y Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, 91, 638-643.
- Mayer, R.E. y Sims, V.K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86, 389-401.
- Meyer, D. y Kieras, D.E. (1997). A computational theory of executive cognitive processes and multiple-task performance: I. Basic Mechanisms. *Psychological Review*, 104, 3-65.
- McCarthy, G., Blamire, A.M., Puce, A., Nobre, A.C., Bloch, G., Hyder, F., Goldman-Rakic, P.S. y Shulman, R.G. (1994). Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. In *Proceedings of the National Academy of Sciences*, 91, 8690-8694.
- McCarthy, R.A. y Warrington, E.K. (1990). *Cognitive Neuropsychology: A clinical introduction*. San Diego: Academic Press.
- Meyer, J., Shinar, D. y Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human factors*, 39, 268-286.
- Moore, P.J. (1993). Metacognitive processing of diagrams, maps and graphs. *Learning and Instruction*, 3, 215-226.
- Moore, P.J. y Schofield, N.J. (1994). Metacognitive instruction in map reading. *Educational Psychology*, 14, 259-268.
- Moreno, R. y Mayer, R.E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.
- Moreno, R. y Mayer, R.E. (2000). A coherence effect in multimedia learning: The case for irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, 92, 117-125.
- Moreno, R., Mayer, R.E. y Lester, J.C. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interactions. In *Proceedings of ED-MEDIA 2000*. Charlottesville: AACE Press.
- Mumaw, R.J., y Pellegrino, J.W. (1984). Individual differences in complex spatial processing. *Journal of Educational Psychology*, 76, 920-939.
- Najjar, L.J. (1998). A framework for learning from media: The effects of materials, tasks and tests on performance. *Dissertation Abstracts International. Section B: The Science and the Engineering*, 59, 3084.
- Novak, J.D. (1998). *Learning, creating, and using knowledge: Concepts maps as facilitative tools in schools and corporations*. Mahwah: Erlbaum.
- O'Donnell, A.M. (1994). Learning from knowledge maps: The effects of map orientation. *Contemporary Educational Psychology*, 19, 33-44.
- O'Donnell, A.M., Dansereau, D.F. y Hall, R.H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review*, 14, 71-86.
- Ortega, J.A. y Fernández, E. (1996). *Alfabetización visual y desarrollo de la Inteligencia. Programa de intervención didáctica basado en el entrenamiento de capacidades perceptivo-visuales, grafomotrices y de lectura crítico-analítica de la imagen*. Granada: Fundación Educación y Futuro.
- Ortega, J.A. (1998). Los Lenguajes no verbales y sus implicaciones tecnológico-curriculares. En *Actas del I Congreso Nacional de Didáctica de las Lenguas en el Sistema Educativo Español*. Granada: Grupo Editorial Universitario.
- Paivio, A. (1986). *Mental representation: a dual coding approach*. Oxford: Oxford University Press.

- Paivio, A. (1988). Basic puzzles in imagery research. In Michel Denis y Johannes Engelcamp (eds.), *Cognitive and neuropsychological approaches to mental imagery*. Dordrecht, The Netherlands: Martinus Nijhoff Publishing.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255-287.
- Paivio, A. y Clark, J.M. (1986). The role of topic and vehicle imagery in metaphor comprensión. *Communication and Cognition*, 19, 367-387.
- Pazzaglia, F. y Cornoldi, C. (1999). The role of distinct components of visuospatial working memory in the processing of texts. *Memory*, 7, 19-41.
- Pearson, D.G., Logie, R.H. y Gilhooly, K.J. (1999). Verbal representations and spatial manipulation during mental synthesis. *European Journal of Cognitive Psychology*, 11, 295-314.
- Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, 3, 227-238.
- Perales, J.C., Catena, A., y Maldonado, A. (2002a). Aprendizaje de relaciones de contingencia y causalidad: Hacia un análisis integral del aprendizaje causal desde una perspectiva computacional. *Revista Cognitiva*, 14, 15-41.
- Perales, J.C., Catena, A., y Maldonado, A. (2002b). Tesis, antítesis y síntesis del aprendizaje causal: Defensa de un modelo integrador. *Revista Cognitiva*, 14, 75-93.
- Perales, J.C. y Romero, J.F. (2003). *Executive strategies in text-image integration tasks*. Comunicación remitida a la IV Reunión Científica de Atención (RECA-4). Valencia.
- Pike, R. (1984). Comparison of convolution and matrix distributed memory systems for associative recall and recognition. *Psychological Review*, 91, 281-294.
- Posner, M.I. y Badgayan, R.D. (1998). Attention and neural networks. In R.W. Parks y D.S. Levine (eds.), *Fundamentals of neural network modelling: Neuropsychology and Cognitive Neuroscience*. Cambridge, MA: The MIT Press.
- Posner, M.I., y DiGirolamo, G.J. (1998). Executive attention: Conflict, target detection, and cognitive control. In Raja Parasuman (ed.), *The Attentive Brain*. Cambridge: MIT Press.
- Posner, M.I. y Raichle, M.E. (1994). *Images of Mind*. New York: Scientific American Books.
- Pylyshyn, Z.N. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 80, 1-24.
- Rieber, L.P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal Educational Computing Research*, 5, 431-444.
- Rieber, L.P. (1990). Using computer animated graphics in science instructions with children. *Journal of Educational Psychology*, 82, 135-140.
- Roodenrys, S., Hulme, C. y Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology*, 56, 431-442.
- Roesler, F., Heil, M. y Hennighausen, E. (1995). Distinct cortical activation during long-term memory retrieval of verbal, spatial and color information. *Journal of Cognitive Neuroscience*, 7, 51-65.
- Roesler, F., Heil, M. y Hennighausen, E. (1996). Exploring memory functions by means of brain electrical topography: A review. *Brain Topography*, 7, 301-313.
- Romero, J.F. (2002a). *Diseño y creación de gráficas didácticas: Estudio de las habilidades creativas de los profesores del Campus de Melilla*. Tesis Doctoral. Universidad de Granada. (Pendiente de Publicación).
- Romero, J.F. (2002b). Evaluando el pensamiento del profesor universitario relacionado con el proceso de elaboración de gráficas didácticas. En M. Lorenzo et al. (coor.), *La organización de los nuevos procesos de institucionalización de la educación (CD-Rom Interactivo)*. Granada: Grupo de Investigación AREA/Innovación y cualificación.
- Rossano, M.J., y Morrison, T.T. (1996). Learning from maps: General processes and map-structure influences. *Cognition and Instruction*, 14, 109-137.
- Rubman, C.N. y Waters, H.S. (2000). A, B seeing: The role of reconstructive processes in children's comprehension monitoring. *Journal of Educational Psychology*, 92, 503-514.
- Rumelhart, D.E. y Zipser, D. (1986). Feature discovery by competitive learning. *Cognitive Science*, 9, 75-112.
- Samarapungavan, A., Vosniadou, S. y Brewer, W.F. (1996). Mental models of the Earth, Sun and Moon: Indian children's cosmologies. *Cognitive Development*, 11, 491-521.
- Schmalhofer, F. y Glavanov, D. (1986). Three components of understanding a programmer's manual. *Journal of Memory and Language*, 25, 279-294.
- Schnotz, W. (1993). Some remarks on the commentary: On the relation of dual coding and mental models in graphics comprehension. *Learning and Instruction*, 3, 247-249.
- Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14, 101-120.
- Schnotz, W., Picard, E. y Hron, A. (1994). How do successful and unsuccessful learners use text and graphics?. *Learning and Instruction*, 3, 181-199.
- Schofield, N.J. y Kirby, J.R. (1994). Position location on topographical maps: Effects of task factors, training and strategies. *Cognition and Instruction*, 12, 35-60.
- Schwartz, N.H. y Kulhavy, R.W. (1981). Map features and recall of discourse. *Contemporary Educational Psychology*, 6, 151-158.
- Schwartz, N.H. y Philippe, A.E. (1991). Individual differences in the retention of maps. *Contemporary Educational Psychology*, 16, 1-12.
- Shah, P., y Carpenter, P.A. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General*, 124, 43-61.
- Shah, P. y Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14, 47-69.
- Shallice, T. (1994). Multiple levels of control processes. In C. Umiltà y M. Moscovitch (eds.), *Attention and Performance XV*. Cambridge, MA: MIT Press.
- Siegel, L. (1994). Working memory and reading: A lifespan perspective. *International Journal of Behavioral Development*, 17, 109-124.
- Simon, H.A. y Chase, W.G. (1973). Skill in chess. *American Scientist*, 61, 394-403.
- Sims, V.K. y Hegarty, M. (1997). Mental animation in the visual sketchpad: Evidence from dual-task studies. *Memory and Cognition*, 25, 321-333.
- Smith, E.E., Jonides, J. y Koeppel, R.A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, 6, 11-20.
- Stenning, K. y Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*, 19, 97-140.
- Titone, R. (1981). *Psicodidáctica*. Madrid: Narcea.
- Vaidya, C.J. Zhao, M., Desmond, J.E. y Gabrieli, J.D.E. (2002). Evidence for the cortical encoding specificity in episodic memory: Memory induced reactivation of picture processing areas. *Neuropsychologia*, 40, 2136-2143.
- Van Essen, D.C., Anderson, C.H. y Felleman, D.J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science*, 25, 419-423.
- Verdi, M.P., Johnson, J.T., Stock, W.A., Kulhavy, R.W. y Whitman-Ahern, P. (1997). Organized spatial displays and texts. Effects of presentation order and display type on learning outcomes. *Journal of Experimental Education*, 4, 303-317.
- Vilches, L. (2002). *La lectura de la imagen*. Barcelona: Paidós.
- Vosniadou, S. (1993). Knowledge acquisition and conceptual change. *Applied Psychology: An International Review*, 41, 347-357.
- Vosniadou, S. (1996). Towards a revised Cognitive Psychology for new advances in learning and instruction. *Learning and Instruction*, 6, 95-109.
- Wallace, D.S., West, S.W.C., Ware, A. y Dansereau, D.F. (1998). The effect of knowledge maps that incorporate gestalt principles on learning. *Journal of Experimental Education*, 67, 5-16.
- Warrington, E. y Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain: A Journal of Neurology*, 92, 885-896.

(Artículo recibido: 19-6-2003, aceptado: 6-10-2004)