



*Perspectivas en el análisis de la multimodalidad:  
desde los inicios al estado del arte*

MARTÍN KALTENBACHER

Universidad de Salzburgo.

TRADUCCIÓN: C. GABRIELA D'ANGELO

RESUMEN - Este artículo proporciona una amplia visión histórica desde los inicios del análisis multimodal hasta los candentes temas actuales y las perspectivas futuras en el estudio de la multimodalidad. A diferencia de otras publicaciones, se considera que el comienzo histórico de las investigaciones sobre multimodalidad fue el *Tratado sobre Laoconte* de Lessing (1766), de modo que lo que actualmente se da es el resurgimiento de un debate que ya había tenido lugar, parcialmente, entre los clasicistas alemanes. Un tema que se discute fervientemente en la actualidad es cómo los modos se relacionan con otros modos en forma de sub-modos y dentro de los diferentes medios. El potencial de significación de un modo cambia de acuerdo con la preponderancia del modo dentro del constructo semiótico así como también a través de la pérdida o adquisición del significado meta-funcional en el desarrollo tecnológico de los medios masivos. Un ejemplo de esto es la tipografía. Se presentan nuevos temas y tendencias en el campo actual de la investigación de la multimodalidad y los últimos descubrimientos en una gama de estudios en áreas de investigación aplicada, por ejemplo se presentan y discuten trabajos sobre traducción en hospitales, diseño de hipertextos, diseño de museos, subtítulo de películas y traducción de tiras cómicas. Para terminar, se esbozan las futuras perspectivas sobre cómo llevar la investigación de la multimodalidad a una base más empírica, por ejemplo a través del análisis del corpus multimodal.

PALABRAS CLAVE: *multimodalidad, semiótica, tipografía, síntesis y automatismo del habla, medicina, hipertexto, aprendizaje, traducción, análisis de corpus*

ABSTRACT. This article gives a broad historical overview from the early beginnings of multimodal analysis to currently burning issues and future perspectives in the study of multimodality. In contrast to other publications, the historical beginning of multimodality research is seen with Lessing's *Treatise on Laocoon* (1766), so what we see today is the revival of a debate that was partly already discussed among the German classicists. A fervently discussed topic today is how modes relate as sub-modes to other modes and within different media. The meaning potential of a mode changes according to the salience of the mode within a semiotic construct as well as through the loss or gain of meta-functional meaning in the technological development of the mass-media. This is exemplified for typography. New topics and trends in current multimodality research are portrayed, and recent findings in a range of studies in applied research fields, such as translation in hospitals, hypertext design, museum design, film sub-titling and comics translation are presented and reviewed. Finally, future perspectives on how to put multimodality research on a more empirical basis, e.g. through multimodal corpus analysis, are outlined.

Recibido: 18 de marzo de 2007 • Aceptado: 16 de agosto de 2007.

KEY WORDS: *multimodality, semiotics, typography, speech synthesis and automation, medicine, hypertext, learning, translation, corpus analysis.*

RESUMO. Este artigo oferece uma ampla visão histórica desde os inícios da análise multimodal até os candentes temas atuais e as perspectivas futuras no estudo da multimodalidade. Ao contrário de outras publicações, considera-se que o *Tratado sobre Laoconte* de Lessing (1766) foi o começo histórico das investigações sobre multimodalidade, de modo que o que atualmente acontece é o ressurgimento de um debate que já tinha acontecido, parcialmente, entre os classicistas alemães. Um tema que é discutido ferventemente na atualidade é como os modos se relacionam com outros modos na forma de sub-modos e dentro dos diferentes meios. O potencial de significação de um modo muda segundo a preponderância do modo dentro do construto semiótico assim como também através da perda ou aquisição do significado meta-funcional no desenvolvimento tecnológico dos meios em massa. Um exemplo é a tipografia. Apresentam-se novos temas e tendências no campo atual da investigação da multimodalidade e os últimos descobrimentos em uma gama de estudos nas áreas de investigação aplicada, por exemplo, apresentam-se e discutem-se trabalhos sobre tradução em hospitais, desenho de hiper-textos, desenho de museus, subtítulagem de filmes e tradução de historietas cómicas. Para terminar, são esboçadas as futuras perspectivas sobre como levar a investigação da multimodalidade a uma base mais empírica, por exemplo, através da análise do corpus multimodal.

PALAVRAS CHAVE: *multimodalidade, semiótica, tipografia, síntese e automatismo da fala, medicina, hiper-texto, aprendizagem, tradução, análise do corpus.*

## *Introducción*

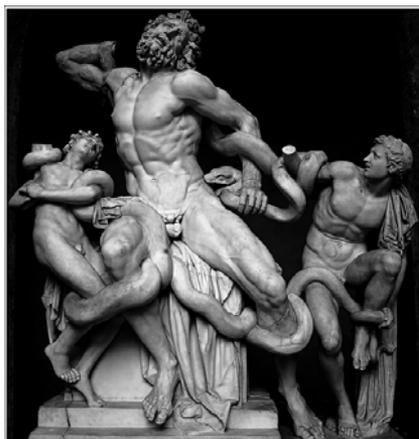
Helmut Stöckl (en prensa) ha afirmado que el florecimiento de las investigaciones sobre multimodalidad es «el descubrimiento de lo obvio». Y, de hecho, la multimodalidad debería ser obvia, no sólo para los que nos ocupamos de la investigación científica de los sistemas semióticos y sus interrelaciones, sino para todo aquel involucrado en el proceso de escritura o diseño de textos de cualquier tipo. A pesar de, o probablemente debido a su obviedad, la investigación multimodal es una disciplina muy joven o, para decirlo con más propiedad, es una disciplina que debería haber resurgido mucho antes. El presente artículo trata de esbozar algunos de los temas candentes de la investigación multimodal desde sus inicios al actual estado del arte. La parte 1 trata sobre el desarrollo histórico de la disciplina, mencionando los aportes que han hecho a este campo los clasicistas alemanes, la lingüística sistémico funcional y el análisis del discurso. En la parte 2 se discutirán preguntas teóricas referidas a las jerarquías modales y al cambio de preponderancia y función de los modos, centrándonos particularmente en la tipografía. En la parte 3 abordaremos algunos campos industriales y económicos en los que se aplica la multimodalidad con mayor dinamismo. Y en la parte 4 cerraremos el artículo con un llamado para que se realicen investigaciones de carácter más empírico, y sugeriremos algunas perspectivas para el futuro desarrollo de la investigación multimodal.

## 1. *El crecimiento de una disciplina - breve historia de la investigación multimodal*

### 1.1. LESSING Y LOS CLASICISTAS ALEMANES.

Muchos lingüistas, particularmente aquellos con antecedentes sistémicos funcionales, consideran que la investigación multimodal nació con la publicación de los dos innovadores libros de Kress & van Leeuwen: *Reading Images* (1990) y *Reading Images: The Grammar of visual design* (1996). Pero mucho antes de la publicación de estos dos influyentes libros, los académicos habían comenzado a mirar el potencial semiótico inherente a los diferentes modos y medios. En 1755, Johann Joachim Winckelmann escribió un influyente tratado titulado *Über die Nachahmung der griechischen Werke in der Malerei und Bildhauerkunst* (Pensamientos acerca de la imitación de obras griegas en la pintura y la escultura), que llevó a Gotthold Ephraim Lessing (1766) a escribir su famosa respuesta *Laokoon: oder über die Granzer der Malerei und Poesie* (Laoconte: un ensayo sobre los límites de la pintura y la poesía). Este ensayo debe ser considerado como el primer abordaje al estudio de la multimodalidad. Lessing analiza dos famosos retratos de la muerte de Laoconte y sus dos hijos. Laoconte fue un sacerdote troyano que aconsejó a su pueblo que no permitiera entrar a la ciudad de Troya el gigantesco caballo diseñado por el héroe griego Odiseo. Sin embargo, los dioses habían decidido la caída y la destrucción de Troya y enviaron dos malignas serpientes para que mataran a Laoconte como advertencia para que los troyanos no destruyeran el caballo. Este asesinato es el tema de dos famosas obras de arte clásicas. Una es la antigua escultura 'El grupo Laoconte' (ver Figura 1), hallada en Roma en 1506 y que aún hoy puede verse en el Museo del Vaticano. La otra es un pasaje de la épica nacional romana «La Eneida» de Virgilio (Eneida 2, 199-267), en la que Eneas narra cómo Laoconte fue asesinado por las serpientes.

Figura 1: El grupo Laoconte en los Museos del Vaticano



Ambas obras describen el mismo hecho mítico, es decir, tratan de construir el mismo significado. Lessing escribe acerca de cómo este significado se expresa de manera diferente de acuerdo con el modo en el que se narra la historia. Una gran diferencia en la semiótica de la narración épica y la escultura es el hecho de que la escultura es estática, mientras que el poema es dinámico (Lessing, 1766/1990, pp. 35f.). Las esculturas pueden mostrar sólo un momento de una historia compleja y, por lo tanto, deben concentrarse en el clímax del argumento. El poema puede mostrar la historia completa como si se desarrollara a través del tiempo. Una de las diferencias más notorias en estas dos obras es la forma en que se presenta el grito de Laoconte. Lessing (1766/1990, p.29) sostiene que la escultura griega muestra a un Laoconte que suspira más que grita, ya que un grito afearía su rostro; los escultores griegos tenían que retratar a sus héroes como poseedores de la más pura belleza, incluso si padecían un dolor mortal. El poema de Virgilio no padece de este tipo de restricciones, ya que, como Lessing (1766/1990, p. 35) sostiene «cuando el Laoconte de Virgilio grita, ¿quién va a pensar que para gritar se necesita una boca grande y que una boca de gran tamaño nos hace feos?» Por lo tanto tiene sentido que Virgilio permita que su héroe trágico grite ('clamores simul horrendos ad sidera tollit / al mismo tiempo manda horribles gritos a las estrellas'; Virg. Eneida 2, 222), ya que esto no viola ninguna convención de género sino que contribuye al desarrollo dinámico de la trama.

Lessing continúa trabajando sobre las diferencias entre las dos obras y relaciona las características semióticas generales de los respectivos géneros, escultura y escritura. En otras palabras, explora las modalidades en las que se basan estas formas artísticas diferentes y describe su potencial y sus limitaciones semióticas. Su ensayo puede ser considerado como uno de los primeros estudios multimodales de las humanidades europeas. Este hecho ha sido ampliamente referido, particularmente en filología alemana y en el período clásico de la literatura alemana, por autores como Schiller (*Über das Pathetische*, Tratado sobre lo patético, 1793) y Goethe (*Über Laokoon*, Un tratado sobre Laoconte, 1798). Como tal, el discurso sobre Laoconte se centra en los mismos temas que el famoso libro de Kress & van Leeuwen (1990, 1996) y, en particular, el de O'Toole (1994) que aborda y profundiza el nuevo marco de análisis. Sin embargo, resulta sorprendente que hasta el momento este hecho no se haya reflejado en los trabajos modernos sobre multimodalidad.

## 1.2. HALLIDAY Y EL RECONOCIMIENTO DE LA SEMIOSIS COMO HECHO SOCIAL

Como se mostró anteriormente, el actual interés en investigar los procesos multimodales no puede verse completamente como una nueva disciplina sino como el resurgimiento o redescubrimiento de un importante e interesante campo de investigación. La principal diferencia entre los nuevos y los antiguos abordajes es que los clasicistas alemanes veían las obras que estudiaban

con ojos de críticos de arte en general y con ojos de críticos literarios en particular. En la actualidad, los académicos que trabajan en el campo de la multimodalidad a menudo vienen de la lingüística y, en particular, de la escuela de lingüística sistémico funcional desarrollada por Halliday (por ejemplo, 1961, 1975). Esta rama de la lingüística ve el lenguaje como un sistema semiótico que explotamos para funcionar en un amplio contexto social y cultural. Es este énfasis en el contexto social, la principal diferencia respecto de los abordajes lingüísticos tradicionales, estructuralistas y generativos, en los que las oraciones se ven y se analizan en sí mismas, desprovistas de todo contexto y contexto situacional que pueda rodearlas. La visión sistémica considera que este contexto situacional constituye los factores determinantes de las estructuras semióticas que elegiremos en nuestras interacciones sociales con los demás (Halliday, 1978, p.110). Estas estructuras semióticas incluyen tres dimensiones que constituyen tres tipos diferentes de significados:

- campo (el tema del intercambio lingüístico) le da forma al significado ideativo;
- tenor (el rol de la relación entre las personas que participan en el intercambio) le da forma al significado interpersonal; y
- modo (el canal a través del cual se transmite el intercambio) modela el significado textual.

Las tres dimensiones proporcionan el «marco conceptual para representar el contexto social como el medio ambiente semiótico en el que las personas intercambian significados» (ibid.).

### 1.3. KRESS Y VAN LEEUWEN Y LA SEMIÓTICA DE LA IMAGEN

Los primeros que extendieron las tres dimensiones metafuncionales a un modo semiótico que no fuera el lenguaje fueron Kress & van Leeuwen (1990, 1996) y O’Toole (1994).

Kress & van Leeuwen (1990, 1996) comenzaron a explorar la semiótica de las imágenes (dibujos, pinturas, gráficos, bosquejos, mapas, retratos, fotografías, etc.) en un profundo análisis de lo que dieron en llamar:

- ‘patrones de representación’ (1996, p. 13), es decir, cómo visualizamos nuestra experiencia del mundo (significado ideativo);
- los ‘patrones de interacción’ (ibid.), es decir lo que sucede entre los que realizan las imágenes y aquellos que las observan (significado interpersonal);
- los ‘principios de composición’ (1996, pp. 181 ff), es decir, cómo las imágenes están visualmente organizadas por su valor informativo, preponderancia contextual (saliencia) y su marco de referencia (framing).

En trabajos más recientes, Kress & van Leeuwen (2001) han ampliado su visión de los textos multimodales como articulaciones múltiples que pueden crear significado en y a través de los dominios del discurso, el diseño, la producción y la distribución.

#### 1.4. O'TOOLE Y LA SEMIÓTICA DEL ARTE EN EXHIBICIÓN

Al mismo tiempo que Kress y van Leeuwen desarrollaban su marco semiótico para abordar el significado de lo visual, O'Toole (1994) adaptaba la visión sistémica funcional del lenguaje a la semiótica del arte en exhibición, y en ese marco, principalmente para el análisis de la escultura, la arquitectura y la pintura. En analogía con el lenguaje, él sostiene que las tres metafunciones son, de hecho, las funciones de todo tipo de semiosis y, de acuerdo con esto, una obra de arte sólo puede ser analizada a la luz de estas tres funciones. En su teoría de la semiótica, las tres metafunciones son:

- la 'función modal' (interpersonal), que refiere a cómo la atención, los pensamientos y las emociones del observador están unidas y relacionadas a la obra (O'Toole, 1994, pp. 5ff). En el retrato de la muerte de Laoconte descrito anteriormente, la función modal se relaciona exactamente con cómo los elementos que emocionan, como el suspiro, el dolor o la tristeza, influyen sobre la recepción subjetiva del observador de esta obra de arte;
- la 'función representacional' (ideativa), es la que presenta a un observador lo que la obra de arte de hecho muestra (O'Toole, 1994, pp. 12ff). Esta transmite información básica sobre las personas, los objetos, las acciones, los hechos, las escenas, los paisajes, etc. que se muestran en una pintura o en una escultura, en nuestro caso, dos serpientes que estrangulan a tres personas;
- la 'función composicional' (textual) (O'Toole, 1994, pp. 22ff.), que representa las decisiones del artista relacionadas con la organización del espacio disponible, la distribución de las formas, las líneas, los colores, etc.. En nuestro ejemplo, esto se relaciona con la distribución de los cuerpos en relación el uno con el otro, por ejemplo, colocar a Laoconte como la figura más grande en el centro de la escultura, presentándolo por lo tanto también como tema.

Al igual que las metafunciones del lenguaje, "las tres funciones siempre trabajan juntas. Son inseparables, y sólo las separamos en nuestra descripción por la conveniencia de concentrarnos en cada uno de los sistemas particulares que operan en cada función" (O'Toole, 1994, pp.22f.).

## 1.5. ANÁLISIS DEL DISCURSO MULTIMODAL

Mientras la lingüística sistémico funcional proporcionó la teoría lingüística a la luz de la cual muchos académicos se orientaron hacia la multimodalidad, la disciplina madre de aplicación fue el análisis del discurso (Iedema, 2003, pp. 29ff.). El análisis del discurso tradicional, o la lingüística textual como se la llamó en un primer momento (por ejemplo, Halliday & Hasan, 1976; De Beaugrande & Dressler, 1981), comenzaron a explorar las estructuras gramaticales que funcionaban más allá de los límites de la oración, es decir, a nivel textual, e investigaron los fenómenos que unían oraciones individuales para formar un todo cohesivo y coherente. El objeto de estudio original fue, primero, el texto escrito y, después, el texto hablado, los que eran considerados como fenómenos más o menos monomodales.

Mientras tanto, el trabajo pionero de Kress & van Leeuwen y O'Toole ha servido de inspiración para muchos académicos y lingüistas para profundizar la teoría sistémico funcional y los métodos de análisis del discurso aplicándolos a otros tipos más complejos de semiosis. Los lingüistas se han dado cuenta de que no existe el discurso monomodal puro. Que el lenguaje hablado no existe sin gestos que lo acompañen, muecas, postura corporal y movimiento de la vista, del mismo modo que los textos escritos no existen sin marcos, espacios, tipografía ni color. De modo que lo que vimos no fue, de hecho, un cambio de atención de mono a multimodalidad sino el reconocimiento de la naturaleza completamente multimodal de todo texto (Cf., p.e., Scollon & Levine, 2004: 1).

Una de las tendencias más recientes en análisis del discurso multimodal es explorar el despliegue multimodal del discurso en los nuevos medios que proporcionan las tecnologías de información y comunicación. Parece que hasta el momento hemos subestimado la gran influencia que tienen las limitaciones tecnológicas de los nuevos medios sobre el discurso que en ellos tiene lugar. La Mesa Redonda de Georgetown (Georgetown Round Table, GURT) lanzó este tema a la palestra y lo hizo centro de sus discusiones en 2002, y muchas de esas discusiones están actualmente disponibles en un libro nuevo titulado *Discourse & Technology - Multimodal Discourse Analysis*, editado por Levine & Scollon (2004). Algunas de las preguntas que surgen en esa publicación son:

- Cómo el concepto de 'contexto' para discursos en comunicación mediada por computadora, los mensajes de correo electrónico y los chats necesitan adaptarse al cambiante medio físico y virtual en el que se desarrolla la comunicación (Jones, 2004);
- Cómo se aplican los rituales retóricos tales como preguntar, aludir y citar en videoconferencias a fin de establecer autoridad, identidad y diseño del destinatario (Davis & Mason, 2004);

- Cómo las relaciones entre las tecnologías de la representación (multimodalidad) y las tecnologías de la diseminación (multimedialidad) en softwares educativos da forma al educando, al medio de aprendizaje y al contenido (Jewitt, 2004);
- Cómo las videograbaciones han abierto nuevas posibilidades para el análisis del discurso que incorpora formas de comunicación no verbal, como el diseño espacial y arquitectónico de un desván que limpiaron voluntarios vocacionales (de Saint-Georges, 2004), el retiro y la distribución de las mercaderías como parte de un 'discurso de atención al público' en una tienda de departamentos (Filliettaz, 2004), el contacto visual y el diseño gráfico de una entrevista con solicitante de Cédula de Residencia (Johnston, 2004) y modos de compensar la falta de contacto visual en interacciones entre videntes y no videntes (Everts, 2004).

## 2. *Cambio de jerarquías y cambios de funciones*

### 2.1. ¿MODOS, VARIANTES, SUBMODOS O CARACTERÍSTICAS?

Algunos de los puntos cruciales que se están debatiendo en la actualidad en relación a esta nueva disciplina son: qué es exactamente un modo, cuál es la relación entre modo y medio, cuántos modos diferentes existen, y cómo se los puede categorizar tipológicamente. Algunos de estos temas surgen en el actual debate acerca de la jerarquía de la tipografía como recurso semiótico. ¿La tipografía es un modo en sí mismo y, de serlo, cuál es la relación entre tipografías y entidades tales como tipo, tamaño, negrita, itálica, letra negra versus roja, etc. de las letras? Van Leeuwen sostiene que

la nueva tipografía es completamente multimodal. Comunica no sólo a través de las formas de las letras en sí sino también a través del color, la textura, la perspectiva, los marcos, el movimiento. Todos estos son modos que no son particulares de la tipografía sino que forman parte de un lenguaje visual multifacético que la tipografía tiene en común con las imágenes, el diseño de productos, la decoración de interiores y la arquitectura entre muchos otros. (Van Leeuwen, en prensa).

Cuando leemos la lista anterior de van Leeuwen, vemos que algunos de estos modos son más dominantes o sobresalientes, mientras que otros lo son menos. Por lo tanto, necesitamos categorizar los recursos semióticos en una jerarquía de modos primarios, secundarios, terciarios, etc. Lo están intentando, por ejemplo, Stöck (en prensa), que considera que la tipografía está ubicada en un continuum de recursos semióticos y distingue entre modos centrales (el lenguaje), variantes mediales (la escritura), modos periféricos (tipografía), submodos (color) y características (saturación, pureza). Contrariamente, Lim (en prensa) sostiene que la tipografía no debería ser vista como un modo semiótico sino como una red de sistemas que operan en el plano de

la expresión del lenguaje. De modo que, al igual que las líneas, sombras y formas, los componentes tipográficos poseen potencial sólo como subsistemas a través de los que se expresa el significado.

Un problema del que se debe dar cuenta es que ciertos modos, como el color, operan como submodos en diferentes niveles en una jerarquía de modos, dependiendo del medio o de la construcción semiótica en la que tienen lugar. En una pintura, el color tendrá más preponderancia y contribuirá más al significado global que el color de ciertas palabras o de ciertas letras en un texto. Esto significa que el color en un texto puede ser un submodo que conlleva parte del significado pero no significados esenciales, mientras que en una pintura, particularmente en la pintura abstracta, el color puede ser un modo central a través del cual se expresa la mayoría de la significación.

## 2.2. MODOS DINÁMICOS CON FUNCIONES CAMBIANTES

Además de la preponderancia de los modos, pueden cambiar también algunas de las funciones originales que cumplían ciertos modos o incluso pueden desaparecer completamente con el paso del tiempo, de manera que el potencial de significación que se aloja en un modo puede también experimentar algunos cambios. Un ejemplo de este tipo de pérdida de potencial de significado es el diseño gráfico de las mayúsculas al comienzo de artículos o párrafos de revistas, como por ejemplo en *Newsweek*. Esas iniciales normalmente se imprimen en letras más grandes, negritas y mayúsculas. Históricamente esta técnica fue desarrollada por un cúmulo de razones. En la época medieval los monasterios recurrían a monjes particularmente eficientes para diseñar e iluminar las letras iniciales de los folios simples de un códice con oro, plata y colores que se obtenían combinando otros materiales preciosos. Las ‘mayúsculas ilustradas’ contenían las llamadas ‘miniaturas’, que eran dibujos muy pequeños que ilustraban parte del contenido del texto, a veces en forma de pequeños dibujos de historias completas. Estos ornamentos cumplían las tres metafunciones descritas anteriormente. Desde el punto de vista ideativo, representaban las escenas, acciones, figuras que se describían en el texto y de este modo facilitaban la comprensión. En cuanto a la función interpersonal, cumplían una función estética apelando al lector y poniendo de manifiesto la habilidad del artista. A menudo los monasterios también deseaban exaltar su riqueza e incrementar el valor de los libros mediante el uso prolífico de oro o de otros materiales raros o de gran valor en el diseño caligráfico del libro. Por supuesto, la rica ornamentación de los ‘libros sagrados’ también se consideraba parte de la alabanza a Dios. Textualmente las iniciales servían para estructurar el texto en unidades jerárquicamente organizadas separándolo en capítulos y párrafos. De modo que las mayúsculas de los códices medievales eran multifuncionales y multimodales. El libro de Umberto Eco (1980, 1983) *El nombre de la rosa* nos puede aportar una idea ‘lúdica’ del trabajo de un calígra-

fo y miniaturista medieval; para obtener una historia más profunda de la caligrafía ver, por ejemplo, Mazal (1985) o Jakobi-Mirwald (1997).

Al contrario de lo que ocurría en tiempos medievales, en los libros y en las revistas modernas las iniciales han perdido bastante de su anterior potencial de construcción de significado. Tanto su significado interpersonal así como también su significado ideativo se reduce a un mínimo. Ya no se expanden en forma de miniaturas porque han sido reemplazadas por otras ilustraciones gráficas que, normalmente, no están unidas a la tipografía verbal, por ejemplo, fotografías o dibujos. En vez de elaborar un libro más valioso usando sustancias preciosas, en la actualidad se utilizan materiales y diseños de bajo costo a fin de mantener los valores de producción de los libros y revistas tan bajos como sea posible. La única función que queda de las mayúsculas iniciales es textual, dado que marcan el comienzo de nuevos artículos y secciones. Incluso esta información es bastante redundante, ya que también se expresa mediante otros recursos tipográficos tales como los encabezados, la tipografía, o incluso otros modos, tales como el espaciado o los enmarcados. Por lo tanto, la mayoría de las antiguas funciones de las mayúsculas caligráficas ya no se explotan y actualmente los lectores han perdido el inmenso potencial de significación que solían tener.

### *3. Análisis multimodal en investigación aplicada*

Aparte de los modos estudiados en el trabajo ya mencionado, los recursos semióticos que han sido más investigados incluyen:

- todo tipo de visualización, como por ejemplo imágenes, películas, dibujos animados, etc., por ejemplo, Echkrammer (en prensa), Hodge & Kress (1988), Kaindl (en prensa), Stöckl (1998, en prensa), O'Halloran (2004a);
- arquitectura, por ejemplo O'Toole (2004), Alias (2004);
- modos acústicos tales como el habla, la música, el sonido, por ejemplo, van Leeuwen (1999);
- color, por ejemplo, Kress y van Leeuwen (2002);
- gestualidad, por ejemplo, Muntigl (en prensa), Kendon (1990), McNeill (1992, 2001);
- y, por supuesto, varias combinaciones multisemióticas de diferentes modos, por ejemplo, Hofinger & Ventola (en prensa), Kaltenbacher (en prensa), Taylor (en prensa).

Algunas de las ramas industriales y económicas hacia las que se ha orientado la investigación multimodal son:

- síntesis del habla y habla automática (sistemas telefónicos, robótica),
- medicina (hospitales, textos de auto ayuda);
- diseño de hipertextos (www, CD);
- la industria del entretenimiento (cine, caricaturas);
- y el sector educativo (materiales educativos, diseño de museos, aprendizaje por internet).

En la siguiente sección intentaré brindar un resumen amplio de algunos los últimos desarrollos en estos campos.

Sin embargo, debido a la amplia gama de investigaciones multimodales y particularmente a su inmenso espectro de aplicación, un resumen de esta naturaleza debe necesariamente ser de alguna manera rudimentario. Muchas de las contribuciones que aquí se discuten se originan en una perspectiva sistémico funcional y se publican en dos nuevos libros sobre discurso multimodal, es decir, O'Halloran (2004b) y Ventola, Charles & Kaltenbacher (en prensa).

### 3.1. SÍNTESIS ORAL Y AUTOMATIZACIÓN DEL HABLA

La investigación en este área siempre ha tenido una fuerte base tecnológica pero también ha comenzado a reconocer en la multimodalidad un elemento clave para el desarrollo de nuevos sistemas, en particular sistemas que permitan la interacción hombre-computadora. Los últimos trabajos en este campo se han orientado hacia, por ejemplo:

- multimodalidad e inteligencia artificial en interfaces multimedia (Lee, 1997; Nijhold, 2003);
- terminales de computadoras de pantallas táctiles interactivas (Raisamo, 1999);
- unidades automáticas de generación de voz, por ejemplo asistentes de teléfonos celulares en automóviles o centros de atención de llamadas de servicio al consumidor automatizadas (Colby, 2002; Sajor, 2002);
- modelos animados de rostros bidimensionales y cabezas tridimensionales que mueven labios y boca de acuerdo a la cualidad del fonema y fruncen el ceño, elevan el entrecejo, asienten y niegan con la cabeza de acuerdo con el contenido (Beskow, 2003).

Esta rama depende tanto del desarrollo de la micro-electrónica, de la mecánica de precisión y de la ingeniería de software como de la cabal comprensión de los procesos semióticos subyacentes. El problema crucial a superarse es que solo puede predecirse el intercambio lingüístico de manera limitada. Los sistemas de síntesis del lenguaje existentes sólo son capaces de generar segmentos de significados recurrentes, por ejemplo, un reloj automático con

una unidad de reconocimiento de voz puede responder a preguntas verbales del tipo ‘¿qué hora es?’ y generar la respuesta correcta porque este tipo de intercambio es altamente estandarizado. Sin embargo, si alguien le pregunta al reloj ‘¿ya es hora de almorzar?’, es poco probable que esa persona obtenga una respuesta apropiada.

Por supuesto que existen sistemas de respuesta de llamadas que logran funcionar más allá de este nivel extremadamente básico, como un evaluador de este artículo me ha señalado. Muchos sistemas de servicios al cliente en aeropuertos, terminales ferroviarias, servicios telefónicos pero también servicios de asistencia técnica al cliente de empresas de computadoras pueden comprender y reaccionar a ciertas solicitudes. Sin embargo, todos estos sistemas se limitan a un dominio muy específico que es posible programar y ni bien se abandona este dominio la interacción se rompe. Además, muchos usuarios probablemente hayan experimentado la cansadora y frustrante experiencia de que la frase más repetida de muchos sistemas de respuesta automatizada sea ‘¡Por favor, repita!’.

Por lo tanto es correcto sostener que los sistemas de generación de voz que existen hasta el momento sólo pueden reproducir cierta cantidad de habla grabada previamente o combinar frases en un contexto muy limitado dentro de un conjunto muy acotado de estructuras opcionales predeterminadas. Si estos sistemas alguna vez serán capaces de participar en una negociación de significados creativa que no se limite a lo verbal sino que incluya lo visual por medio de expresiones faciales y gestos, está muy en duda.

### 3.2. MEDICINA

La medicina es una ciencia muy diversa y compleja y naturalmente también se alimenta de los avances y desarrollos de una vasta cantidad de ciencias auxiliares. La investigación multimodal se vincula con la medicina en una cantidad de formas, tanto en el campo tecnológico como en aspectos que dependen de la comunicación exitosa y la ilustración de conocimiento médico experto. Algunas de las instituciones médicas que recientemente han centrado su atención en el análisis de discursos multimodales son los hospitales, los centros de rehabilitación y los hogares para enfermos y ancianos; géneros clásicos, como la traducción, también se han volcado a la perspectiva de discurso multimodal. Bührig (en prensa), por ejemplo, ha analizado el discurso médico-paciente en hospitales en casos en los que el paciente no es hablante nativo de una lengua y requiere de la asistencia de un intérprete. Ella sugiere que uno de los principales obstáculos en este tipo de comunicación es la barrera lingüística que surge cuando los expertos utilizan un lenguaje especializado que las otras partes involucradas no conocen. En estas situaciones, médicos, pacientes e intérpretes a menudo tratan de superar la barrera lingüística recurriendo a modos semióticos alternativos a la comunicación verbal, y Bührig

muestra cómo los médicos utilizan, por ejemplo, materiales visuales, tales como diagramas y esquemas, a fin de proporcionar a sus pacientes instrucciones acerca de las operaciones pendientes.

Eckkrammer (en prensa) ha examinado hipertextos médicos de auto-ayuda a la luz de lo que ella denomina ‘capas inter-semióticas’ de textos multimodales. La autora muestra que los elementos verbales y pictóricos de un texto nunca desarrollan su significado en forma separada sino a través de un proceso de interacción entre las capas semióticas involucradas. Por ejemplo, la Figura 2.

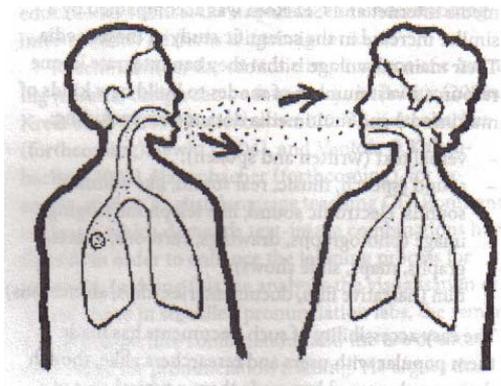


Figura 2: Diseminación de gérmenes de tuberculosos a través del aire

Figura 2: ejemplo de capas intersemióticas del tipo ‘juxtaposición’, reproducidas de Eckkrammer (en prensa), (originalmente reproducidas de <http://www.cpmc.columbia.edu/tbcpp/abbouttb.html>).

Eckkramer sugiere cuatro dimensiones de capas inter-semióticas:

- la transposición tiene lugar cuando las imágenes se reproducen en textos verbales o viceversa;
- la juxtaposición tiene lugar cuando el texto visual y el verbal coexisten sin mezclarse;
- la combinación ocurre cuando los elementos visuales y verbales se combinan en un texto;
- la fusión tiene lugar cuando las imágenes y el texto verbal se funden en una nueva forma textual.

Además de su contribución a la investigación del discurso médico, Eckkrammer también toca otras áreas clave de la investigación sobre la multimodalidad, la del diseño de hipertexto.

### 3.3. DISEÑO DE HIPERTEXTO

El gran éxito y la rápida difusión de dos medios relacionados, internet y el CD-Rom, coincidieron con un aumento similar en el estudio científico de estos medios. Su principal ventaja es que pueden integrar en un mismo recurso una gran variedad de modos para construir nuevas clases de documentos multimodales y multimediales, incluyendo:

- textos verbales (escritos y orales);
- sonido (habla, música, sonido real, como sonidos de animales, sonidos electrónicos, como el sonido de un teléfono sonando);
- imágenes (fotografías, dibujos, caricaturas, tablas, gráficos, mapas, diapositivas);
- películas (películas narrativas, documentales, clips, animaciones).

La fácil accesibilidad a este tipo de documentos los ha hecho muy populares tanto entre usuarios como investigadores, aunque su rápida difusión los ha convertido en el típico caso de una disciplina en la que la comprensión de los fundamentos semióticos de base está muy retrasada respecto de la producción y distribución de los productos. Hasta hace muy poco tiempo prácticamente no se había propuesto ninguna didáctica semióticamente orientada para guiar el proceso de construcción de significado en documentos conectados de esta manera. Esto no es sorprendente debido a que el foco en lo semiótico (o multimodal) incluso en el diseño de documentos más tradicionales es bastante nuevo. Para el diseño de hipertextos en Internet, Lutz (1998, p. 151f.) mostró que una búsqueda de ‘autoría de hipertexto’ (‘Hypertext Authoring’) con Yahoo o Altavista arrojaba miles de sitios con instrucciones de carácter tecnológico pero prácticamente ninguna de ayuda que estuviera fundada en la lingüística (o en la semiótica). De la misma manera, van der Geest y Spyridakis (2000, p. 301) afirmaron: «Una mirada por los estantes de una librería con un buen stock muestra toda una gama de libros sobre los aspectos técnicos del diseño de portales de internet y un puñado de libros sobre diseño gráfico, pero falta algo sobre el abordaje del diseño de portales de internet desde una perspectiva comunicacional abarcadora». No resulta llamativo que en la actualidad muchos diseñadores de hipertextos reconozcan que el arte de diseñar sitios web ha sido abordado directamente desde un aprendizaje de ensayo y error orientado sólo por la tecnología, en el que poco o nada tuvo que ver un abordaje teórico desde la semiosis o desde la comunicación (Lee 1997, p. 1, Lutz 1998; p. 152, de Jong & van der Geest 2000, p. 311).

Ha comenzado a producirse un cambio de las cuestiones exclusivamente tecnológicas bajo el término clave «usabilidad» (uso), ligado a Jakob Nielsen. El énfasis en este tipo de didáctica está puesto en diseñar sitios que sean «a la vez útiles y utilizables para la audiencia a la que están dirigidos» (Murray & Cos-

tanza 1999). Si bien el énfasis en las reglas de uso parece estar en la selección de contenidos relevantes que el usuario puede recuperar con facilidad, estas reglas incluyen sugerencias mínimas acerca de cómo destacar el texto integrándolo de manera razonable con otros modos. Por ejemplo, [www.usability.gov](http://www.usability.gov), sugiere cómo usar el espacio en blanco, tablas, listas, íconos, ilustraciones, etc, por ejemplo. «Utilice dibujos, líneas artísticas, tablas y otros gráficos de manera funcional. Los gráficos deben contribuir al mensaje, no sólo ser bonitos». ([http://www.usability.gov/methods/collecting\\_writing.html](http://www.usability.gov/methods/collecting_writing.html)).

En los últimos años, académicos de diferentes disciplinas (particularmente de la ciencia de la comunicación y la lingüística) han comenzado a incorporar a sus trabajos sobre diseño de hipertextos un área de atención dedicada a las cuestiones semióticas. Van der Geest (2001, p. 7f.) propone una gama de preguntas nuevas relacionadas con aspectos de la multimodalidad, tales como:

- qué modo de presentación se seleccionará y con qué propósito;
- qué aspecto tiene una presentación efectiva;
- cómo lograr coherencia entre los modos de un sitio web.

*Technical Communication* (3/2000) ha dedicado un número especial a desarrollar un conjunto de herramientas heurísticas para el diseño web dentro del marco de la ciencia de la comunicación. En este número, por lo menos dos trabajos abordan aspectos que también emergen del análisis lingüístico multimodal, ya que trabajan sobre:

- aspectos de la navegación en la web a través de links y nodos (Farkas & Farkas, 2000);
- estructuración de la información mediante tipografías diferentes, tamaños, mayúsculas, alineados, encabezamientos, ilustraciones, íconos, etc. (Williams, 2000).

Además de lo anterior, un creciente número de lingüistas ha comenzado a describir la semiótica que subyace en los hipertextos existentes, aunque sin lograr abarcar todos los aspectos tal como lo solicitaran van der Geest y Spyridakis (2000). Las últimas contribuciones incluyen:

- Lemke (2002), que ha proporcionado algunas nociones fundamentales de la semántica del hipertexto en un análisis de cómo los elementos textuales y visuales se combinan en sitios web informativos y educativos de la NASA;
- Kok (2004) explora cómo diferentes presentaciones semióticas se combinan y disponen en un hipertexto a fin de captar la atención del observador mientras representan la realidad;
- Eckkrammer (en prensa, para mayor detalle véase 3.2.);

- Kaltenbacher (en prensa, para mayor detalle véase 3.4.);
- y muchos otros artículos recopilados en Baldry (2000a).

### 3.4. EDUCACIÓN

También íntimamente interrelacionada a la cuestión del diseño de hipertextos están los actuales avances en la industria educativa, que parece ser un campo particularmente apropiado para explotar las tecnologías de la información y la comunicación. Actualmente podemos distinguir dos tendencias en las disciplinas educativas, ambas basadas en nuevas formas de representaciones multisemióticas. La primera, una tendencia que está bastante de moda, puede incluirse mejor bajo el término *e-learning* (aprendizaje electrónico). Esta rama atraviesa un período de desarrollo en el cual el contenido de las interacciones en las aulas y los materiales de clase tradicionales (libros, casetes, videos) se adaptan y transfieren a los nuevos medios. A fin de satisfacer los desafíos que constituye el diseño de una lección electrónica efectiva, Clark & Mayer (2003, p. 25) requieren conocimiento experto en psicología de la enseñanza, producción multimedia, gráficos, programación y diseño de interfaces. Pero mientras ellos claramente reconocen la necesidad de las ciencias de la educación, la importancia de comprender las relaciones inter-semióticas permanece nuevamente ignorada.

Trabajos recientes sobre las implicancias semióticas de los materiales de aprendizaje electrónico se pueden encontrar en Baldry (2000a), Kress et al, (2001), Kaltenbacher (2003), Kaltenbacher (en prensa), Jewitt (2004), y Ventola & Kaltenbacher (2003). Kaltenbacher (en prensa) estudia los CDs de enseñanza de inglés y evalúa qué demandas deben satisfacer las combinaciones texto-imagen a fin de profundizar el proceso de aprendizaje de los usuarios. Al hacer esto, analiza la visualización de las ondas sonoras en los llamados laboratorios de pronunciación, la semántica de las combinaciones texto-imagen y el uso de video clips de corta duración para practicar pronunciación. Sostiene que las visualizaciones deben satisfacer ciertos requisitos semióticos: tienen que ser fácilmente interpretables, deben ayudar al alumno a desambiguar el significado que están aprendiendo, y deben tener una representación exacta del significado que se está enseñando. En imágenes estos requisitos se cumplen mejor mediante el uso de presentaciones icónicas, mientras que la utilización de visualizaciones complejas a menudo impide, en vez de favorecer, la comprensión y el aprendizaje.

La otra tendencia en el sector educativo es el museo multisensorial y el diseño de exhibiciones. Los diseñadores que trabajan en este campo han comenzado a integrar modos de grandes dimensiones físicas, como por ejemplo, objetos que se pueden tocar, esculturas, muebles, diseño de interior, arquitectura, paisajismo, etc. para crear nuevas experiencias multisemióticas que desafíen a los visitantes a utilizar sus cinco sentidos y no sólo la vista y el

oído. Trabajos muy recientes centrados en el análisis de la multimodalidad de estos museos han sido llevado a cabo, por ejemplo, por Pang (2004) y Hofinger & Ventola (en prensa). Pang (2004) analiza una cantidad de muestras en una exhibición en el Museo de Historia de Singapur centrándose en cómo la exhibición establece roles sociales e ideológicos particulares entre su tema (*Desde la colonia hasta la nación*), lo que se exhibe en sí y los visitantes. Hofinger & Ventola (en prensa) analizan los procesos semióticos dinámicos involucrados en la visita a un museo dedicado a la vida familiar de Wolfgang Amadeus Mozart en Salzburgo, Austria. Analizan el interior de una sala del museo concentrándose en la forma en que la pintura de la familia de Mozart y la disposición de algunos muebles próximos a la pintura se interrelacionan con el audiotexto que lo acompaña. También demuestran que la particular explotación de diferentes recursos ideativos, interpersonales y organizativos en la sala pueden decirnos algo acerca del contexto socio-cultural en el que el museo se encuentra.

### 3.5. INDUSTRIA DEL ENTRETENIMIENTO

Otra rama económica que actualmente atrae mucha atención de lingüistas y semióticos es la de la industria del entretenimiento. Esta rama siempre ha sido de algún modo dependiente de la experiencia lingüística dado que las películas, caricaturas, video-juegos, etc. necesitan traducirse a diferentes idiomas para su distribución y recepción internacional. No resulta sorprendente que en la actualidad particularmente la disciplina de estudios de traducción comience a integrar la multimodalidad a sus investigaciones. Algunos aspectos del lenguaje siempre han sido más difíciles de traducir que otros, especialmente aquellos en los que ciertas formas de los textos de origen, como las expresiones idiomáticas y los juegos de palabras, están intrínsecamente ligadas y son altamente dependientes de la semiótica de otros modos, como las imágenes. Por lo tanto, los lingüistas han comenzado a mirar más de cerca los aspectos semióticos del humor y su traducción en largometrajes y caricaturas.

Taylor (en prensa) ha llamado la atención sobre el papel que pueden jugar los análisis de textos multimodales en el proceso de subtítulo de un largometraje a un idioma extranjero. Se concentra en instancias en las que el texto de origen contiene un juego de palabras específico de la lengua fuente y que también es visible en la película. En este caso, el juego de palabras puede sufrir una considerable pérdida de significado en el proceso de traducción, dado que el texto de origen y el de llegada involucran muchas modalidades que interactúan y que requieren lo que Taylor denomina una «traducción intersemiótica». Un factor interesante del subtítulo es que no reemplaza el texto original, como en el caso del doblaje, sino que agrega un nuevo canal que tiene que corresponder tanto al texto visual como al texto verbal original. Las estrategias más frecuentes que se emplean en subtitulación son las de

condensar, cortar o elidir elementos del texto original. Taylor sostiene que el lugar y la medida en que estas estrategias deben aplicarse depende de si el espectador puede o no recuperar el sentido original a través de otros modos semióticos (imágenes, música, sonido) de manera que el texto en su conjunto no sufra ninguna pérdida de significado.

Kaindl (en prensa) se interesa por la traducción de humor expresado de manera multimodal en las historietas francesas *Astérix* y *Tintin*. Sus premisas son que la traducción es un proceso más cultural que lingüístico y que el humor de las historietas está codificado en las ilustraciones y no en el texto. Este último aspecto ha sido hasta el momento especialmente dejado de lado en los estudios de traducción. Kaindl identifica una cantidad de estrategias diferentes para la traducción del humor, que van desde el borrado del humor, cambiando la modalidad del humor, a inventar un humor ausente en el original (ver Tabla 1). Además, investiga la relación entre textos verbales e imágenes en la creación de juegos entre lengua y signos. Kaindl explícitamente desafía la ampliamente difundida premisa de que las imágenes poseen una semiótica comprensible universalmente. Lo que crea un problema particular para el proceso de traducción es cuando un juego visual o sígnico depende de la presentación de un objeto que pertenece a una cultura en particular, por ejemplo una pelota de rugby. Otro fenómeno que depende del conocimiento cultural específico es la intertextualidad pictórica y la alusión pictórica.

Tabla 1: Ocho estrategias de traducción de Kaindl reproducidas de su trabajo (en prensa).

	Texto de partida: Tipo semiótico	Texto de llegada: Tipo semiótico	Texto de partida vs. Texto de llegada: técnica humorística		Texto de partida vs. Texto de llegada: Relación lengua-imagen	
A	humor monomodal	humor monomodal	similar	diferente	similar	diferente
B	humor monomodal	humor multimodal	similar	diferente	diferente	
C	humor monomodal	sin humor	--		--	
D	sin humor	humor monomodal	diferente		diferente	
E	humor multimodal	humor multimodal	similar	diferente	similar	diferente
F	humor multimodal	humor monomodal	similar	diferente	diferente	
G	humor multimodal	sin humor	--	--	--	--
H	sin humor	humor monomodal	diferente		diferente	

Aparte de los temas de traductología discutidos anteriormente, en las películas la multimodalidad actualmente está siendo abordada a través de un análisis basado en la tecnología y desde el punto de vista de los estudios culturales. O'Halloran (2004a) explora la manera en que el software de edición *Adobe Premiere 6.0* puede utilizarse para analizar el desarrollo temporal de las elecciones semióticas en el largometraje *Barrio Chino* (1974) de Roman Polanski. Y Markis Rheindorf (en prensa) emplea una perspectiva de estudios culturales para investigar cómo la danza, la música y el vestuario se utilizan

para establecer la clase social y los roles de género de los personajes en el largometraje *Dirty Dancing* (1987) de Emile Ardolino.

#### 4. *Perspectivas en el análisis de la multimodalidad*

Si hacemos una búsqueda en Google™ del término *multimodalidad* (*multimodality*) encontraremos una respetable cantidad de más de 50.000 entradas. Muchas de estas menciones aparecen en artículos publicados en revistas tecnológicas *on-line* o en los portales de internet de empresas dedicadas a desarrollar tecnología de habla o salud. En una de estas revistas Michael Sajor, vice-presidente de una empresa dedicada a desarrollar infraestructuras multimodales, ha afirmado recientemente: «parece que el mundo está hablando sobre la multimodalidad» (Sajor, 2002). Parece que muchos expertos en industrias tecnológicas han reconocido la necesidad de que entender la multimodalidad es algo vital para comunicar mensajes rutinarios a diario así como para ilustrar complejos conocimientos especializados. Pero aún así, la investigación de la multimodalidad todavía carece de prestigio para muchos lingüistas y expertos en comunicación. Uno de los motivos de que esto ocurra es la falta de fuerte evidencia empírica que sostenga muchas de las afirmaciones que se hacen. De hecho, «hay algunos rincones del campo que tildan al análisis multimodal de investigación débil - afirmando que sólo cuenta lo empírico», como me señaló un evaluador.

Tabla 2: Transcripción fílmica multimodal de un cuadro del documental *Echo of the Elephants*, reproducida de Taylor (en prensa).

Tiempo 1	Cuadro visual TOMA 1	Imagen visual	Acción kinética	Banda de sonido fases y subfases	Interpretación metafuncional
		CP: estacionario HP: frontal VP: mediano D: MLS: VS: elefantes VF: cerca CR: gris, verde, azul CO: naturalista	animales pastando, moviendo colas y orejas, pájaros caminando	Sonido de la sabana: canto de pájaros, piar de las cicadas	Fase 1 Subfase (a) Exp. Actores: elefantes reunidos. Int. Espectador posicionado como observador del ámbito de los elefantes

Esta falta de empirismo es por cierto un punto débil. Sin embargo esta debilidad no sólo ha sido reconocida sino abordada por muchos académicos que se encuentran trabajando en la disciplina. Casi a diario se presentan consideraciones teóricas originales así como los análisis cualitativos de documentos multimodales individuales de todo tipo que están muy bien preparados para ser validados o refutados por medio de rigurosas evaluaciones empíricas. El tema candente que necesita abordarse es ¿cómo llevar a cabo esas pruebas! En Italia, un grupo de investigadores pugna por proporcionarle a la investiga-

ción multimodal una base más empírica y cuantitativa (Anthony Baldry, Paul Thibault, Christopher Taylor). Baldry (2000a, 2000b, 2004), Thibault (2000), Taylor (2000, en prensa) y Baldry & Thibault (2004) han mejorado métodos de transcripción de documentos multimodales «al punto de que se han convertido en herramientas para el desarrollo de lingüística de corpus multimodal» (Baldry, 2000a, p. 31). En la Tabla 2 se presenta un ejemplo de transcripción multimodal. Mientras tanto, Baldry (2004) ha extendido el alcance de esta transcripción al desarrollo de un concordador multimodal en línea, el sistema Multimodal Corpus Authoring (MCA) que brinda la posibilidad de rastrear patrones visuales y tipos en películas y videotextos.

Un segundo proyecto, el Modelo GeM (Género y Multimodalidad) está a cargo de las universidades de Bremen, Nottingham Trent, Stirling y Bradford (John Bateman, Judy Dein, Renate Henschel y Patrick Allen). En el proyecto GeM (para mayor detalles véase <http://purl.org/net/gem> y Bateman et al. en prensa), un equipo de lingüistas computacionales y sistémico funcionales está desarrollando un corpus multimodal anotado en una base XML que integra marcas de lingüística de corpus tradicional con nuevos procedimientos de anotación. El corpus proporciona información sobre el diseño multimodal y las estructuras jerárquicas de documentos individuales, tales como la estructura gráfica de oraciones, encabezados, fotografías, dibujos, figuras, leyendas, textos en imágenes, íconos, celdas, encabezados de listas, numeración de página, notas al pie, etc.. La Figura 3 muestra un ejemplo. Bateman et al. (en prensa) sostiene que el Modelo GeM puede, en última instancia, proporcionar «la capacidad de hacer preguntas acerca de nuestro corpus de manera suficientemente flexible y poderosa como para promover la construcción y la evaluación de una teoría».

```
<text xref="lay-02" font-family="sans-serif"
font-size="18" font-style="normal"
font-weight="bold" case="mixed"
justification="left" color="white"
background-color="grey"/>
```

Figura 3: anotación de tipografía inicial, reproducida de Bateman et al. (en prensa)

Este tipo de proyectos está bien equipado para satisfacer la exigencia de colocar la investigación multimodal en una sólida base empírica. Sin embargo, estos primeros intentos no son suficientes. Además de estas investigaciones empíricas necesitamos investigaciones sobre el alcance semiótico de documentos multimodales orientadas a las siguientes cuestiones:

- cómo los receptores ‘leen’ los documentos multimodales;
- qué combinaciones de diferentes tipos de modos promueven la comprensión y una comunicación exitosa, y cuáles inhiben dicho proceso;

- si los alumnos que trabajan con complejos materiales educativos multimodales aprenden mejor y más rápido; y sobre todo:
- si los sofisticados documentos multimodales comunican cosas que los documentos tradicionales no pueden comunicar.

Este artículo ha tratado de mostrar cómo la comprensión teórica de los procesos multisemióticos sumada a un análisis profundo de documentos existentes en una disciplina puede contribuir a detectar y resolver problemas en la construcción y comprensión de significado que a menudo han pasado desapercibidos. Este artículo también ha señalado áreas que requieren mayor investigación empírica a fin de aportar hipótesis basadas en la multimodalidad con una base más sólida. Entre aquellos que pugnan por lograr alguna de las ambiciosas tareas esbozadas en este artículo existe el consenso general de que es necesario el desarrollo de métodos sistemáticos para que sea posible analizar teórica y sistemáticamente los datos, y evaluar empíricamente el despliegue semiótico de recursos y modos y sus combinaciones en todos los aspectos de nuestra vida cotidiana. Algunos de estos aspectos pueden parecer más triviales, como la comprensión de largometrajes e historietas. Otros pueden facilitar tareas más importantes de nuestras vidas, tales como estudiar y aprender. Algunos, sin embargo, pueden ser cruciales para nuestro bienestar, por ejemplo, cuando nuestra vida depende de la exitosa comunicación entre paciente, médico y máquinas. El potencial para la integración de especialistas en multisemiótica en campos sociales, económicos e industriales es grande. En qué medida lo es, es algo que recién comenzamos a comprender.

## NOTAS

1. Desafortunadamente, reproducir mayúsculas ilustradas a modo de ejemplo es difícil debido a que estos materiales están fuertemente protegidos por derechos de autor. Pero en Internet abundan las imágenes de iniciales ilustradas. Sólo tienen que buscar 'historiated initials' en cualquier buscador.

## REFERENCIAS BIBLIOGRAFICAS

- ALIAS, S. (2004). "A semiotic study of Singapore's Orchard and Marriott Hotel", en K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 55-79. Londres: Continuum.
- BALDRY, A. (Ed.) (2000a). *Multimodality and multimodality in the distant learning age*. Compopasso: Palladino Editore.

- BALDRY, A. (2000b). "English in a visual society: Comparative mediality". en A. Baldry (Ed.), *Multimodality and multimodality in the distant learning age*, pp. 41-89. Campopasso: Palladino Editore.
- BALDRY, A. (2004). 'Phase and transition, type and instance: Patterns in media texts as seen through a multimodal concordancer', en K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 83-108. Londres: Continuum.
- BALDRY, A. & Thibault, P. (2004). *Multimodal transcription and text analysis*. Londres: Equinox Publishing.
- BATEMAN, J., DELIN, J. & HENSCHER, R. (en prensa). "Multimodality and empiricism: Preparing for a corpus-based approach to the study of multimodal meaning-making", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 65-88. Amsterdam & Philadelphia: John Benjamins.
- BESKOW, J. (2003). *Talking heads - Models and applications for multimodal speech synthesis*. Tesis doctoral no publicada. Department of Speech, Music and Hearing, KTH, Stockholm. <http://www.speech.kth.se/~berkow/thesis>
- BÜHRIG, K. (en prensa). 'On the «multimodality» of interpreting in medical briefings for informed consent: using diagrams to impart knowledge' en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 227-214. Amsterdam & Philadelphia: John Benjamins.
- CLARK, R.C. & MAYER, R. (2003). *E-learning and the science of instruction. Proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: Jossey-Bass/Pfeiffer.
- COLBY, J. (2002). "Multimodality: The new wave of mobile interaction", en *Speech Technology Magazine*, November/December 2002. <http://www.speechtechmag.com>
- DAVIS, G. & MASON, P. (2004). "Trying on voices: Using questions to establish authority, identity, and recipient design in electronic discourse", en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp. 47-58. Georgetown, Washington, D.C.: Georgetown University Press.
- DE BEAUGRANDE, R.A. & DRESLER W.U. (1981). *Einführung in die Textlinguistik*. Tübingen: Niemeyer.
- DE JONG, M. & VAN DER GEEST, T. (2000). "Characterizing web heuristics", en *Technical Communications 3*, 311-324.
- DE SAINT-GEORGES, I. (2004). "Materiality in discourse: The influence of space and layout in meaning making", en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp. 71-87. Georgetown, Washington, D.C.: Georgetown University Press.
- ECO, U. (1980). *Il nome della rosa*. Milano: Bompiani.
- ECO, U. (1983). *The name of the rose*. New York: Harcourt.
- ECHKRAMMER, E.M. (en prensa). "Drawing on theories of inter-semiotic layering to analyse multimodality in medical self-counselling texts and hypertexts",

- en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on Multimodality*, pp. 211-226. Amsterdam & Philadelphia: John Benjamins.
- EVERTS, E. (2004). 'Modalities of turn-taking in blind/sighted interaction: better to be seen and not heard?', en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp.128-145. Georgetown, Washington, D.C.: Georgetown University Press.
- FARKAS, D. & J. FARKAS (2000). "Guidelines for designing web navigation", en *Technical Communication*, 3, 341-358.
- FILLIETTAZ, K. (2004). "The multimodal negotiation of service encounters", en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp. 88-100. Georgetown, Washington, D.C.: Georgetown University Press.
- GOETHE, J. W. (1798/1998). "Über Laokoon", en F. Apel (Ed.), *Goethe - Ästhetische Schriften 1771 - 1805*, pp. 489-500. Frankfurt a. Main: Deutscher Klassiker Verlag.
- HALLIDAY, M.A.K. (1961). "Categories of the theory of grammar", en *Word*, 17, 3, 242-292.
- HALLIDAY, M.A.K. (1975). *Language as social semiotic: towards a general sociolinguistic theory*. Columbia: Hornbeam Press.
- HALLIDAY, M.A.K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Arnold.
- HALLIDAY, M.A.K. & HASAN, R. (1976). *Cohesion in English*. London: Longman.
- HODGE, R. & KRESS, G. (1988). *Social semiotics*. Cambridge: Polity Press.
- HOFINGER, A., & VENTOLA, E. (en prensa). "Multimodality in operation: Language and Picture in a Museum", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 193-2210. Amsterdam & Philadelphia: John Benjamins.
- IEDEMA, R. (2003). "Multimodality, resemiotization: Extending the analysis of discourse as multi-semiotic practice", en *Visual Communication*, 2, (1): 29-57.
- JAKOVI-MIRWALD, C. (1997). *Buchmalerei; ihre Terminologie in der Kunstgeschichte*. Berlin: Reimer.
- JEWITT, C. (2004). "Multimodality and new communication technologies", en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp. 184-195. Georgetown, Washington, D.C.: Georgetown University Press.
- JOHNSTON, A. (2004). "Files, forms, and fonts: Mediational means and identity negotiation in immigration interviews", en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp. 116-127. Georgetown, Washington, D.C.: Georgetown University Press.
- JONES, R. (2004). "The problem of context in computer-mediated communication", en P. Levine & R. Scollon (Eds.), *Discourse & technology. Multimodal discourse analysis*, pp.20-23. Georgetown, Washington, D.C.: Georgetown University Press.

- KAINDL, K. (en prensa). "Multimodality in the translation of humour in comics", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 173-192. Amsterdam & Philadelphia: John Benjamins.
- KALTENBACHER, M. 2003. "Language learning via CD-Rom - Old wine in new bottles" en D. Newby (Ed.), *Mediating between theory and practice in the context of different learning cultures and languages*, pp. 171-175. Strasbourg: Council of Europe Publishing.
- KALTENBACHER, M. (en prensa). "Multimodality in language teaching CD-ROMs", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 119-136. Amsterdam & Philadelphia: John Benjamins.
- KENDON, A. (1990). *Conducting interaction: Patterns of behavior in focussed encounters*. Cambridge: Cambridge University Press.
- KOK, A. (2004). "Multisemiotic mediation in hypertext", en K. O'Halloran (Ed), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 131-159. London: Continuum.
- KRESS, G., JEWITT, C., OGBORN, J. & TSATSARELIS, C. (2001). *Multimodal teaching and learning. The rhetorics of the science classroom*. London and New York: Continuum.
- KRESS, G. & VAN LEEUWEN, T. (1990). *Reading images*. Geelong: Deaking University Press.
- KRESS, G. & VAN LEEUWEN, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- KRESS, G. & VAN LEEUWEN, T. (2001). *Multimodal discourse. The modes and media of contemporary communication*. London: Arnold.
- KRESS, G. & VAN LEEUWEN, T. (2002). "Colour as a semiotic mode: Notes for a grammar of colour", en *Visual Communication*, 1, 3, pp. 343-368.
- LEE, J. (ED.) (1997). *Intelligence and multimodality in multimedia interfaces: Research and applications*. Menlo Park, CA: AAAI Press.
- LEMKE, J. (2002). "Travels in hypermodality", en *Visual Communication*, 1 (3), pp. 299-325.
- LESSING. G.E. (1766/1990). "Laokoon: Oder über die Grenzan der Melerei und Poesie", en W. Barner (Ed.), *Gotthold Ephraim Lessing - Werke und Briefe*, 5/2, pp. 11-321. Frankfurt a. Main: Deutscher Klassiker Verlag.
- LEVINE, P. & SCOLLON, R. (Eds.) (2004). *Discourse & technology. Multimodal discourse analysis*. Georgetown, Washington, D.C.: Georgetown University Press.
- LIM, V.F. (en prensa). "Problematising 'semiotic resource'", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 51-63. Amsterdam & Philadelphia: John Benjamins.
- LUTZ, B. (1998). "Hypertext: das Medium und der Autor. Ein Erfahrungsbericht", en B. Kettman, M. Stegu, & H. Stöckl (Eds.), *Mediendiskurse. Verbal-Workshop Graz 1996*, pp. 151-162. Frankfurt am Main: Peter Lang Verlag.
- MAZAL, O. (1985). *Das BUch der Initialen*. Wein: Edition Tusch.

- MCNEILL, D. (1992). *Hand and mind. What gestures reveal about thought*. Chicago: University of Chicago Press.
- MCNEILL, D. (Ed.) (2001). *Language and gesture*. Cambridge: Cambridge University Press.
- MUNTIGL, P. (en prensa). "Modelling multiple semiotic systems", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 31-59. Amsterdam & Philadelphia: John Benjamins.
- MURRAY, G. & COSTANZO, T. (1999) "Usability and the web: An overview", en *Network Notes 61. Information Technology Services*. National Library of Canada. <http://www.collectionscanada.ca/9/1/pl-260-e.html>
- NIJHOLT, A. (2003). "Multimodality and ambient intelligence", en W.F.J. Verhaegh, E.H.L. Aarts & J. Korst (Eds.), *Algorithms and ambient intelligence*. Dordrecht et al.: Kluwer Academic Publishers. (Disponible en: <http://doc.utwente.nl/fid/1144>).
- O'HALLORAN, K. (2004a). "Visual semiosis in film", en K. O'Halloran (Eds.), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 109-130. London: Continuum.
- O'HALLORAN, K. (Ed.) (2004b). *Multimodal discourse analysis: Systemic functional perspectives*. London: Coninium.
- O'TOOLE, M. (1994). "Opera Ludentes. The Sydney Opera House at work and play", en K. O'Halloran (Ed.l), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 11-27. London: Continuum.
- PANG, A. (2004). "Making history in 'From Colony to Nation': A multimodal analysis of a museum exhibition in Singapore", en K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 28-54. London: Continuum.
- RAISAMO, R. (1999). *Multimodal human-computer interaction: a constructive and empirical study*. Tesis doctoral no publicada. Department of Computer Science, University of Tampere. <http://acta.uta.fi/pdf/951-44-4701-6.pdf>
- RHEINDORF, M. (en prensa). "The multimple modes of dirty dancing: a cultural studies approach to multimodal discourse analysis", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.). Amsterdam & Philadelphia: John Benjamins.
- SAJOR, M. (2002). Multimodality: The coming revolution. *Wireless Future Magazine*, September/October 2002. <http://www.wirelessfuturemagazine.com>
- SCHILLER, F. (1793-1962). "Über das Pathetische", en B. von Wiese (Ed.), *Schillers Werke, 20 - Philosophische Schriften, 1. Teil*, pp. 196-221. Weimar: Böhlau.
- SCOLLON, R. & LEVINE, P. (2004). "Multimodal discourse analysis as the confluence of discourse and technology", en P. Levine & R. Scollon (Eds.), *Discourse & technology. multimodal discourse analysis*, pp. 1-6. Georgetown, Washington, D.C.: Georgetown University Press.
- STÖCKL, H. (1998). "Bild-Text-Verknüpfung", en B. Kettemann, M. Stegu & H. Stöckl (Eds.), *Mediendiskurse. Vergal-Workshop Graz 1996*, pp. 73-92. Frankfurt am Main; Peter Lang Verlag.

- STÖCKL, H. (en prensa). "In between modes: Language and image in printed media", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 9-30. Amsterdam & Philadelphia: John Benjamins.
- TAYLOR, C. (2000). "Text analysis and translation: An interactive, self-access computer application incorporating a functional approach", en A. Baldry (Ed.), *Multimodality and multimediality in the distant learning age*, pp. 295-309. Compopasso: Palladino Editore.
- TAYLOR, C. (en prensa). "Multimodal text analysis and subtitling", en E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on multimodality*, pp. 153-172. Amsterdam & Philadelphia: John Benjamins.
- THIBAUT, P. (2000). "The multimodal transcription of a television advertisement: Theory and practice", en A. Baldry (Ed.), *Multimodality and multimediality in the distant learning age*, pp. 311-385. Campopasso: Palladino Editore.
- USABILITY.GOV (2004). *Methods for designing usable web-sites. Collecting, writing and revising content*. US Department of Health and Human Services (Ed.). <http://www.usability.gov/methods>
- VAN DER GEEST, T. (2001). *Web site design in communication design*. Amsterdam: Benjamins.
- VAN DER GEEST, T. & SPYRIDAKIS, J. (2000). "Developing heuristics for web communication", en *Technical Communication 3*, pp. 301-310.
- VAN LEEUWEN, T. (1999). *Speech, music, sound*. London: Macmillan.
- VAN LEEUWEN, T. (en prensa). *Introducing social semiotics. An introductory textbook*. London: Routledge.
- VENTOLA, E. & KALTENBACHER, M. (2003). "Lexicogrammar and language teaching materials - A social semiotic and discourse perspective", en J.E. Joyce (Ed.), *Grammar in the language classroom*, pp. 158-201. Singapore: SEAMEO Regional Language Centre.
- VENTOLA, E., CHARLES, C. & KALTENBACHER, M. (Eds.). (en prensa). *Perspectives on multimodality*. Amsterdam: Benjamins.
- VERGILIUS, P. M. (1988). "Aenis", en J. Götte (Ed.), München & Zürich: Artemis Verlag.
- WILLIAMS, T. (2000). "Guidelines for designing and evaluation the display of information on the web" en *Technical Communication, 3*, pp. 383-396.
- WINCKELMANN, J.J. (1755-1969). *Gedanken über die Nachahmung der griechischen Werke in der Malerei und Bildhauerkunst*. Stuttgart: Reclam.

MARTIN KALTENBACHER estudió inglés y los clásicos en Salzburgo y Oxford y actualmente tiene un puesto postdoctoral en el Departamento de Inglés de la Universidad de Salzburgo, Austria. Sus áreas de interés en investigación son la sintaxis generativa y la lingüística sistémico-funcional y se concentra en la adquisición del lenguaje, el discurso del turismo y el discurso

multimodal. Ha publicado numerosos artículos y el libro *Universal grammar and parameter resetting in second language acquisition* (2001).

Correo electrónico: martin.kaltenbacher@sbg.ac.at

Acknowledgement: From on «Perspectives on Multimodality: from the early beginnings to the state of the art». In *Information Design Journal + Document Design*, Vol. 12:3, 2004, pp 190-207. With kind permission by John Benjamins Publishing Company, Amsterdam/Philadelphia. [www.benjamins.com](http://www.benjamins.com).

Agradecimiento: Traducción de «Perspectives on Multimodality: from the early beginnings to the state of the art», en *Information Design Journal + Document Design*, Vol. 12: 3, 2004, pp 190-207. Publicado con permiso de John Benjamins Publishing Company, Amsterdam/Philadelphia. [www.benjamins.com](http://www.benjamins.com)