

Susana ARROYO HIDALGO

(Tecnológico de Monterrey, Campus Estado de México / UAB)

Charo LACALLE ZALDUENDO

(Universitat Autònoma de Barcelona)

VINCULACIÓN DE LA SEMIÓTICA Y LA INFORMÁTICA: INDEXACIÓN DIGITAL DE TELEDIARIOS¹

Históricamente, la Semiótica ha sido considerada como una disciplina profundamente influenciada por el modelo de la Teoría de la Información (Shannon y Weaver, 1947). El carácter binario de los elementos que integran el modelo ha fundamentado procesos semióticos como la relación *competence-performance* o *presencia-ausencia* para definir las estructuras de reconocimiento de la significación, que han incidido notablemente en una buena parte de las teorías de la comunicación.

Desde la perspectiva *peirceana* consideramos que la información contenida en el texto audiovisual es susceptible de segmentación en grupos de unidades representativas de enteras secuencias (*storyboard*², figura 1). El proceso de segmentación de una noticia televisiva exige el rigor de un ‘interpretante’, entendiendo esta noción como “la expresión de una relación dialéctica entre el mundo real, un ya-instalado codificado mediante relaciones institucionales, interiorizado por los individuos (a través de la inculcación pedagógica que ejerce la sociedad "instructora") y esos mismos individuos que las actualizan en sus determinaciones concretas”. (Marty, R <http://www.univ-perp.fr/see/rch/lts/marty/preg39.htm>). El carácter cultural de los procesos interpretativos se presenta como condición necesaria para frenar el vértigo, irreductible de otro modo, de la ‘semiosis ilimitada’.

¹ Esta ponencia forma parte de un proyecto CICYT titulado “Indexación semántica para vídeo digital” realizado por un grupo de investigadores de la Facultad de Ciencias de la Comunicación y del Centro de Visión por Computador de la Universidad Autónoma de Barcelona.

² El *Storyboard* es una imagen mosaico de una noticia segmentada de modo automático.

La construcción de significados es posible porque, antes de la interpretación, las determinaciones individuales han sido traducidas y sometidas a un proceso de reconstrucción que se encuentra contenido en el pensamiento, es decir, en la mente humana: “It appears, then, that the rule for attaining the third grade of clearness of apprehension is as follows: Consider what effects, which might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object” (Peirce, 1992: 133). En el proceso de recepción de los informativos, al recuperar el efecto de la noticia televisiva por un 'interpretante' se perciben, por tanto, los tres elementos del estado de *consciousness: Single, Dual and Plural*:

In this kind of consciousness subject and object are nowise discriminated, in fact there is no discrimination, no parts, no analysis, there is no considering a thing for anything else, no relation, no representation, but just a pure indescribable quale which is gone in the twinkling of an eye and which bears no resemblance to any memory of it. It is just the quality of the immediately present, which is continually pouring through us, always here but never stopping to be examined. It is always fresh, always new, sporting in unbounded manifoldness. Dual consciousness is a sense of another, not present, a sense of hitting and getting hit, of action and reciprocal reaction, of energy. This is the most wide-awake kind of consciousness; its strenuously sets object over against subject, in place of the dreamy failure to recognize the situation which belongs to Feeling (Peirce, 1992:283).

La noticia, entonces, se convierte en un conglomerado de imágenes altamente informativas y tricotómicamente redundantes cuyo diverso –y a veces disperso– contenido, suele desviar la atención del texto indefectiblemente incoherente e irrepetiblemente ausente.

INDEXACIÓN SEMÁNTICA PARA VÍDEO DIGITAL.

Nuestra ponencia se inscribe en un proyecto en el que se plantea la vinculación de la Semiótica con la Informática mediante la recuperación digital, la segmentación y la indexación de las noticias televisivas. El corpus de referencia está constituido por 30 noticias de la edición vespertina de los telediarios de TVE1, TV3, Antena3, Tele5 y Canal+, emitidos entre enero y julio de 2002. En este corpus se incluyen noticias de los diferentes subgéneros que integran los telediarios españoles: nacional, internacional, sucesos, cultura, sanidad, economía y deportes.

El retraso de la convergencia digital en Europa³, nos ha obligado a digitalizar previamente las noticias analizadas. El paso del texto analógico al sistema digital se ha llevado a cabo mediante el empleo del programa Wintur2000 TV. Dicho programa reconoce el número de *píxeles* que integran la imagen analógica y posteriormente procede a la configuración de un nuevo material que queda inscrito en un archivo. El proceso de decodificación y codificación del texto analógico al texto digital se realiza desde la instrucción dada de manera posterior a la entrada del archivo y, después de efectuarse dicha codificación, este nuevo material se inserta en la base de datos que se ha creado para tal fin.⁴

Tras haber ejecutado el programa Wintur2000 TV y colocado el material digitalizado en la base de datos, se procede a la segmentación automática de las noticias mediante el programa *Anotación*, puesto a punto por el Centro de Visión por Computador. Se trata de un *software* de recuperación y segmentación de texto cuyo resultado conforma una imagen mosaico (*storyboard*) compuesta de unidades (*key frames*, figura 1) que proceden de la selección de elementos pertinentes y de la eliminación de otros elementos no significativos. Mediante una sintaxis previamente determinada por instrucciones precisas, se obtiene una respuesta compleja de significado denotativo, cuya función es agrupar los elementos mínimos que conforman una imagen semánticamente clara y compleja para formar cada una de las mencionadas unidades (*key frames*) del *storyboard*.

³ En España, la fecha final del Plan de Migración es el 1º de enero de 2012, fecha en la que se producirá el *switchoff* de las emisiones digitales.

⁴ Esta parte de la investigación la hemos llevado a cabo en el Centro de Visión por Computador de la UAB.

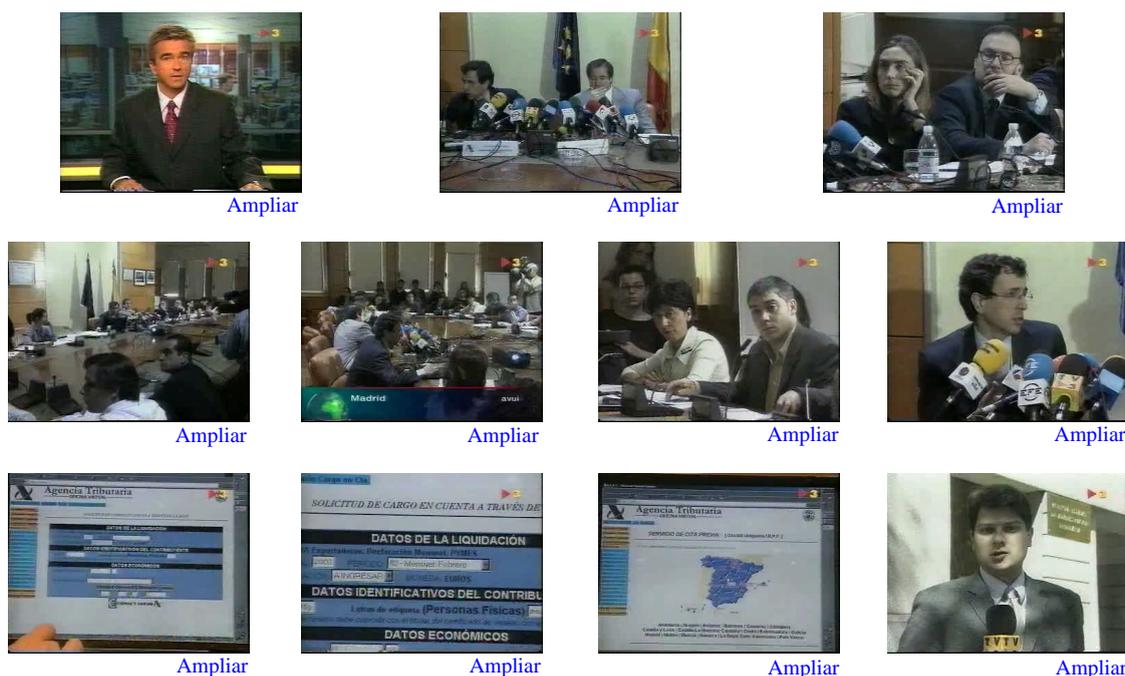


Figura 1. Storyboard de la noticia “El Ministerio d’Hisenda insisteix que fer les declaracions a través d’Internet resulta més comode” (TV3, 25 de marzo de 2002, Telediario de las 20:30 h).

Hemos segmentado de forma manual las noticias antes de someterlas a la segmentación automática, con el rigor que el 'interpretante' mantiene frente a la imagen televisiva. Por el contrario, la segmentación automática es un proceso de recuperación del texto digital, que sucesivamente fragmenta la continuidad de la imagen en movimiento para poder anotarla y conservarla como imagen de archivo.

En el momento de etiquetar lingüísticamente las imágenes segmentadas, recurrimos a la transcripción de lo que hemos denominado la 'frase clave', es decir, los elementos léxicos y semánticos representativos constituyentes de la unidad de sentido de cada noticia, tal y como el informador ha querido o ha tratado de ilustrar la imagen. No obstante, el formato y las características de cada noticia sitúan la frase clave en precedencia de las imágenes puesto que, al enunciarla el conductor de manera introductoria, se convierte en un elemento de carácter paratextual (Genette, 1982) y, por tanto, externo a la imagen (“Libertad sin fianza a un antiguo empleado de una funeraria malagueña que guardaba una veintena de cadáveres en su domicilio”)⁵.

⁵ Hilario Pino, Canal+, Martes 19 de marzo de 2002, Telediario de las 22:00 h.

En todos los ejemplos de la muestra analizada, la segmentación de la imagen arroja resultados diferentes a los obtenidos en la segmentación manual, debido principalmente a la tendencia sincrética del diseño del programa. La imagen segmentada, por consiguiente, nos brinda signos icónicos complejos y generalmente desestructurados, pero de asequible factura. Esto es debido a un lento proceso inferencial, que nos ha permitido ir nutriendo el desarrollo del programa con constantes aportaciones y demostrar a los informáticos la necesidad de determinar imágenes clave en el reconocimiento del material digitalizado, a fin de que el periodista pueda trabajar de manera formal y consistente con *key frames* semánticamente completos.

De manera inversa, de acuerdo con nuestras apreciaciones, el espectador que recibe la noticia o el documentalista que archiva las imágenes para su posterior recuperación, pretenden establecer un vínculo –la mayoría de las veces, casi inexistente– con la información que aporta la imagen, proveyendo de esta forma, un efecto de sentido global, abigarrado e inconcluso pero significativo. La precipitación y voluptuosidad de las imágenes impide el cualquier intento de análisis sistemático de la noticia dado que la fuerza perlocutiva e icónica del texto visual no permite el acceso virtual al significado.

Las diferencias de la segmentación manual y automática se ponen de manifiesto en todas las noticias de la muestra utilizada⁶. Mientras que la segmentación manual recupera la vinculación de la imagen visual con el texto sonoro que intencionalmente concibe el 'interpretante', la segmentación automática modifica la cantidad de imágenes que percibe el espectador. Debido a los condicionantes a los que aún está sometido el proceso de digitalización, la segmentación automática identifica como imágenes significativas diferentes partes de una misma secuencia que tan sólo constituyen simples variaciones de luz (ejemplo: furgoneta de la noticia "Talibán estadounidense", figura 2), contrastes cromáticos (ejemplo: Garzón con una parte de la carátula de Tele 5, figura 3) y superposiciones (ejemplo: Noticia sobre la posición de José Luis Rodríguez Zapatero respecto a la isla Perejil, figura 4).

⁶ La muestra contiene 30 noticias, de las cuales las siguientes han sido segmentadas manualmente: "Declaración de Hacienda a través de Internet" y "Detenciones ETA" (TV3, 25-3-02); "Homenaje víctimas de Atocha" y "Talibán estadounidense" (TVE1, 24-1-02), "Congelación salarial de funcionarios" y "Incidente Metro Bilbao" (Tele5, 26-2-02); "Reforma de la ley de enjuiciamiento criminal" y "Conflicto árabe-israelí" (Antena 3, 12-3-02); "Apertura al público de la Casa Batlló" y "Falsas incineraciones" (CNN+, 19-3-02).



Ampliar



Ampliar



Ampliar



Ampliar

Figura 2

Figura 3

Figura 4

EL PROGRAMA ANOTACIÓN.

La indexación automática es un proceso que se inicia a partir del empleo del programa *Anotación*. Después de haber segmentado el texto se procede a la descripción semántica de cada uno de los *Key frames*. Se acciona el programa y se mantiene la indicación para grabar en el recuadro situado debajo de la pantalla de vídeo, en la parte superior izquierda de la aplicación. Se activa el botón *New session* y después se elige el material que se encuentra en la base de datos correspondiente a la URL (*Uniform Resource Locator*) del proyecto. Se diseña una nueva base de datos por medio de la elección de un título para la noticia y se agregan los comentarios pertinentes dentro de los recuadros de la parte inferior izquierda de la aplicación (*Title* y *Comments*), lo que permitirá su posterior recuperación en la *web*.

A continuación se procede a la grabación del material seleccionado de la información digitalizada mientras observamos la segmentación automática del vídeo (*Play*). Durante el proceso de indexación, cada uno de los *Key frames* agrupados automáticamente en décimas de segundo, de acuerdo con el tiempo de aparición en el vídeo, se coloca en una marcación especial situada en la columna de la derecha y es entonces cuando elegimos la imagen clave de la noticia (*Main key frame*), la cual puede eliminarse para su modificación (*Delete main key frame*). Tras la indexación automática se debe cancelar manualmente la indicación del recuadro que permitió grabar la noticia. Se efectúa, ahora, la anotación de las imágenes mediante los descriptores que se han diseñado para cada noticia o para el grupo de noticias que se pretende indexar (“objetos”, “personas”, “situaciones”, “lugares”, “tiempo”, etc.), a los que se accede situándonos en el recuadro que se encuentra en el centro de la pantalla y que permite anotar uno o varios descriptores por cada imagen.

El audio se anota de acuerdo con la selección previa de cada segmento, el cual puede revisarse las veces que sea necesario siempre y cuando se mantenga sin accionar

la indicación del recuadro del vídeo. El vídeo aparecerá en la pantalla al ser solicitado. La constatación del audio permitirá realizar anotaciones que deberán colocarse en el recuadro inferior izquierdo denominado TC (*Time counter*), seleccionando la décima de segundo la cual se colocará automáticamente en el recuadro inferior derecho para que, a continuación, se proceda a su anotación, la cual puede eliminarse (*Delete audio*) o modificarse (*Modify*). Después de la anotación de imagen y vídeo se inserta el material indexado (*Insert*), se espera el tiempo necesario de su colocación en la base y se obtiene un archivo constituido por una nueva base de datos. Todo esto forma el *Storyboard* de cada una de las bases de datos, es decir, de cada noticia.

SEGMENTACIÓN AUTOMÁTICA VS. SEGMENTACIÓN MANUAL.

En algunas noticias no hay apenas coincidencias entre la imagen mostrada y el discurso enunciado y, en consecuencia, la falta de cohesión entre la imagen y el sonido impiden que el espectador se forme una idea medianamente aceptable del contenido de la información o, cuando menos, dificulta la comprensión de la noticia. A fin de sortear estos obstáculos de la interpretación, el espectador construye *superestructuras* (van Dijk, 1977 y 1983) o se remite a *esquemas* cognitivos (Livingston, 1990) y *supertemas* (Jensen, 1988), que le permiten afrontar la entropía de su propio proceso interpretativo. La sorpresa, el alto contenido informativo de las imágenes *per se*, el conocimiento previo que el espectador tiene de los antecedentes de la noticia o sus propios intereses personales, constituyen factores claves a la hora de fijar la noticia en la memoria del receptor con una eficacia que, de lo contrario, le negarían en todo momento las condiciones *sine quan non* del telediario.

En los diferentes procesos de producción, recepción y conservación de la noticia televisiva, el periodista, el documentalista y el espectador, construyen significaciones culturalmente compartidas que, a pesar de fundamentarse en un mismo ‘objeto’ (el acontecimiento mediático), tan sólo se superponen parcialmente entre sí, ya que la heterogeneidad de sus respectivos marcos de interpretación configura otros tantos ‘interpretantes’.

El periodista parte de una significación global (el conjunto de ítems que le permiten construir una noticia) mucho más amplia que la que facilita al espectador quien, de manera necesaria, se remite a la imagen para construir su propia significación.

De modo inverso el documentalista intenta aferrar, mediante el texto lingüístico que (re)construye, un significado primigenio que ya se ha ido diluyendo en buena parte a lo largo del proceso de interpretación. Imágenes pobres (ejemplo: puerta y edificio en la noticia sobre ETA, figura 5) respecto de la información. Vana quimera. Alquimia



[Ampliar](#)

Figura 5

imposible que difícilmente llegaría a configurar un verdadero proceso de significación sino fuera por el conductor-demiurgo quien, en la presentación de la noticia enuncia, determina y valora el sentido global de la noticia de la misma.

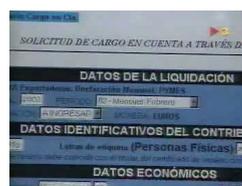
La tipología que presentamos a continuación se remite a la propuesta *peirceana* de clasificación de los signos a partir de su relación con el objeto representado. Peirce señala que el signo mantiene una relación con su objeto en tres niveles: icónico, indicial y simbólico. El resultado de esta tripartición nos permite clasificar aquellas imágenes que nos remiten a una conciencia socialmente compartida dentro de los siguientes parámetros:

1. Imágenes *icónicas*: son imágenes ilustrativas que realmente muestran lo que se está diciendo; se caracterizan por su singularidad. Ejemplos: la noticia del material incautado a ETA, las páginas de *internet* para realizar la declaración de la renta.



[Ampliar](#)

Figura 6



[Ampliar](#)

Figura 7

2. Imágenes *indiciales*: son imágenes contextuales; tienen carácter metonímico. Por ejemplo, marcadores espacio-temporales. Ejemplos: placa y pancarta de conmemoración de los 25 años de la matanza de Atocha.



[Ampliar](#)

Figura 8



[Ampliar](#)

Figura 9

3. Imágenes *simbólicas*: son imágenes cuya significación (inmotivada) ha sido socialmente aceptada; son imágenes metafóricas. Ejemplos: ventana y lámpara como elementos característicos del estilo de Gaudí y, por extensión del Modernismo, mostrados en la noticia sobre la Casa Batlló.



Figura 10 [Ampliar](#)

La mayor dificultad que afrontamos en el momento de la indexación es explicar los signos visuales que no tienen ningún contenido social. Se trata de *seudoindicios* puesto que han sido codificados como ‘ripios’⁷ de manera *ad hoc* para



Figura 11 [Ampliar](#) incineración de cadáveres.

La abundancia y la importancia visual que pueden llegar a adquirir los *seudoindicios* en cada noticia representan hoy por hoy uno de los mayores desafíos al que se enfrenta el programa de indexación.



Figura 12 [Ampliar](#)

Otro de los problemas es que más dificultan la automatización del proceso, consiste en la altamente frecuentada disociación entre imágenes y texto. Como consecuencia, la persona encargada de realizar la indexación se enfrenta a la disyuntiva siguiente:

1. Diseñar un *tesaurus* léxico-semántico tan restringido que sea incapaz de definir puntualmente las particularidades de cada imagen, o
2. Permitir una conceptualización *ad hoc* que expandiese la categorización léxico-semántica hasta extremos inabarcables para el usuario.

⁷ **Ripio:** Conjunto de palabras inútiles o con que se expresan cosas vanas o insustanciales en cualquier clase de discursos o escritos. Fig. Introducir en escritos o discursos o en composiciones artísticas especies o cosas inútiles o insustanciales. (RAE, 1986)

Cualquier decisión que se tome al respecto, para poder continuar el proceso de recuperación de las imágenes indexadas, dificulta extraordinariamente una descripción sistemática y puntual de la significación global de cada noticia anotada automáticamente. Por lo tanto, en el diseño de materiales que desarrolla el Centro de Visión por Computador de la UAB y conscientes de las limitaciones actuales, se trabajan prototipos para que en un futuro no muy lejano se procure el establecimiento de vínculos reales entre la imagen detectada por el programa y el texto en *off* que la avala.

Quizá el día que los periodistas y los documentalistas encuentren imágenes y texto convergentes, será cuando la informática podrá dar soluciones a los diferentes problemas a los que nos enfrentamos hasta el momento en la indexación y recuperación de imágenes digitales.

Referencias bibliográficas

- DIJK, T. A. van (1983). «Discourse analysis. Its development and Application to the structure of news», *Journal of Communication*, 33, 2.
- DIJK, T. A. van (1977). *Text and context*, London: Longman.
- GENETTE, G. (1982). *Palimpsestes. La littérature au second degré*. Paris: Seuil.
- KLAUS B. J., (1988). «News as a social resource. A qualitative empirical study of the reception of Danish television news», *European Journal of Communication*, 3, 3.
- LACALLE, Ch. y ARROYO, S. (2002). «Methodological notes on the image/text of digital archives in TV News». En *Actas del Congreso de la International Association for Media And Communication IAMCR/AIERI/AISC*, Barcelona, Intercultural Communication, 22-26 de julio de 2002.
- LIVINGSTONE, S. (1990). *Making sense of television*, London: Pergamon.
- PEIRCE, Ch. S. (1992a). «How to make our ideas clear», en *Selected Philosophical Writings*, Vol. I.
- ___ (1992b) «Trichotomic», en *Selected Philosophical Writings*, Vol. I.