

Pedro SANTANA MARTÍNEZ

(Universidad de La Rioja)

CONTIGÜIDAD E HIPERTEXTO: CONTRIBUCIÓN A UNA DEFINICIÓN ESENCIAL DEL HIPERTEXTO

Se propone una definición de hipertexto que no atienda a determinaciones externas como pueda ser una implementación concreta del mismo. El énfasis se hace en la pluralidad de secuencias de símbolos que se advierte –frente a la unicidad en el caso del texto– cuando se describe el hipertexto, haciendo abstracción incluso de su descomposición en textos aislados e identificables: De una manera bastante evidente se comprueba que al poner esto en marcha se rebasa la idea de cadena de símbolos, que es central en la teoría de las gramáticas. Sin embargo, la persistencia de una cadena única como parte de la representación formal del hipertexto lleva a mantener a éste dentro de los dominios cubiertos por la citada teoría si se hace lo propio con los textos.

Algunas perspectivas que se abren de este modo tienen que ver con la gramática de los lenguajes de marcas, con la descripción del hipertexto como grafo –lo que habría de conducir a una definición de la familia de grafos que representan posiblemente hipertextos– y con la definición de una medida de la complejidad de los hipertextos, que sería vista como una complejidad sintáctica que habría de conjugarse con la complejidad textual de los mismos definida según alguno de los índices habituales.

1. CONTEXTO DE LA DISCUSIÓN.

Las definiciones habituales de ‘hipertexto’ aparecen profundamente marcadas por su implementación habitual y por el contexto cultural, tecnológico y económico en

que aquél ha llegado a ser lo que es hoy. Nos referimos –por más que aquí renunciamos a repasar las principales definiciones operativas o estructurales que se manejan¹– a que de una manera, en general poco disimulada, el hipertexto se ve como un conjunto de textos enlazados de un modo muy similar a como nos figuramos que aparecen enlazados los archivos informáticos en el sistema de ficheros de un ordenador: nodos o entidades aisladas –si bien extensas y no puntuales– que se integran en un esquema de direccionamiento y que pueden –como evidentemente sucede en los documentos hipertextuales e hipermediáticos que se han convertido en cotidianos– incluir referencias de unos a otros o, afinando más, de unas partes de unos a otras partes de sí mismos o de otros.

Una impronta más que las definiciones habituales han recibido es la de las ideas concernientes al modo en que nos enfrentamos los sujetos o agentes humanos al hipertexto, sobre todo por contraste a un modo ideal que tendríamos de enfrentarnos a los textos. Frente a una lectura canónica, lineal, del texto, el hipertexto ofrecería, junto a acercamientos aberrantes, numerosas posibilidades no más regulares unas que otras².

No se trata de negar que ambas grandes aproximaciones sean pertinentes en su respectivo lugar o complementariamente, pero sí ha de insistirse en que ambas son, por decirlo de algún modo, externalistas; queremos decir que no se basan en propiedades de la materia y la forma lingüística –habría de decirse textual o más bien hipertextual– que esencialmente los integran.

Por otro lado, la noción de texto incluye naturalmente su apertura, la referencia –semántica– a lo que no es texto y a otros textos. Identificar un texto nos procura no sólo un problema filológico: también supone negar que una implementación del acceso a ese texto y a otros sea un hipertexto. Entiéndase que existe una contradicción entre la identidad del texto y su inserción en un hipertexto. De hecho, nos tememos que sea en

¹ El lector puede recorrer desde las reflexiones de un informático en Joyanes Aguilar (1997) hasta las de críticos y teóricos de la cultura reunidos en Landow *comp.* (1997) y también hasta llegar a las teorías más características de la tercera cultura en su avatar más desatado (Aarseth 1997). Desde ese 1997 a la fecha la literatura que se ha unido a estos títulos ha sido ingente, sobre todo, nos tememos, en formato convencional. Los trabajos clásicos de Vannevar Bush y otros autores están disponibles en Internet.

² La linealidad o, mejor, la secuencialidad entre las partes del texto será aquí una nota esencial del mismo. Esto convierte al texto en una entidad ideal no identificable con un documento que puede incluir varios textos. En otras palabras, una novela como *Rayuela* sería más de un texto, si uno se mantiene apegado a la definición de texto implícita en lo que se acaba de decir. En la medida en que no se quiera llegar a tal conclusión, habrá que vigilar el alcance que se dé a definiciones más o menos ideales. Podríamos tal vez estipular que materialmente manejamos documentos –que incluyen inscripciones lingüísticas– que admiten idealizaciones en la forma de textos.

este marco –el del contraste entre un texto ideal y unas plasmaciones concretas que pueden imponer ciertas determinaciones a través de su propia materialidad– donde hayan de entenderse, si se quiere preservar un mínimo de racionalidad, no sólo las relaciones entre texto y documento a las que se aludía en la nota (2), sino igualmente multitud de ideas en torno a la relación del texto y el lenguaje con el mundo –según reza una fórmula habitual–, a los límites del texto, etc.

En esta comunicación intentamos aproximarnos a una que llamamos definición esencial del hipertexto –y ciertamente lo haremos sobre el fondo que proporciona lo que sabemos que es el texto. No se trata de olvidar que el hipertexto ha llegado a ser lo que es por el desarrollo de unas tecnologías que ciertamente lo hacen desarrollarse y lo determinan. Sí se trata de abstraer lo que en común pueda tener el hipertexto con otros depósitos de información y conocimiento o con agrupaciones intencionales de textos, que presentan un carácter distinto (por ejemplo, porque se basan en la actividad de los sujetos que manejan textos o documentos, como es el caso en muchas ocasiones de la llamada *intertextualidad*).

Notemos simplemente, a efectos ilustrativos, el carácter extrínseco de la distinción entre enlace externo e interno³. Del mismo modo que podríamos romper un texto en dos (aunque sólo fuera en el sentido de separar sus soportes actuales), podríamos recomponerlo y habríamos de transformar así enlaces externos en internos y viceversa: no es una hipotética pluralidad de textos componentes lo que define al hipertexto, dicho sea sin perjuicio de que un hipertexto pueda incluir –y esto es algo usual– varios textos.

Con todo, hay que señalar algo importante. La accesibilidad de la totalidad de un hipertexto puede venir posibilitada o no por saltos hipertextuales. Esto depende de la implementación concreta para las tecnologías usuales. Como veremos, esto será importante para la ordenación de las secuencias del hipertexto y para una medida de la complejidad. No se debe confundir, sin embargo, un hipertexto con múltiples partes

³ La materialidad del hipertexto se hace patente aquí, pero es esa misma materialidad la que nos marca los límites de un hipertexto. Sigue habiendo enlaces externos hacia documentos que, por decir así, residen en otro lugar. Por otro lado, un hipertexto ideal corresponde a unas implementaciones materiales concretas y posiblemente alternativas. Sin embargo, un enlace puede ser interno o externo con respecto a un archivo, o con respecto a un hipertexto. La producción material del hipertexto nos señala sus límites y su realidad. Cuestión distinta es que uno puede ir desde ese hipertexto a otro texto o hipertexto sin, por así decir, “levantarse de la silla”.

sólo accesibles mediante saltos, con un hipertexto con enlaces externos a otros hipertextos, algo posibilitado por una red.

Comoquiera que llegaremos a nuestra definición de hipertexto a través del texto, comenzaremos por definir unas relaciones entre los componentes sintagmáticos del texto y veremos a qué restricciones están sujetas y de qué modo, en cambio, se dan en el hipertexto.

2. RELACIONES DE CONTIGÜIDAD BÁSICAS.

En un sistema de escritura alfabético –aunque naturalmente no sólo en él, y también en la lengua hablada–, si dejamos de lado recursos como las notas al pie de página, un texto consiste en una sucesión de instancias de símbolos pertenecientes a un inventario finito que se halla naturalmente sujeta a numerosas restricciones, las que se separan un texto de una cadena aleatoria de símbolos⁴.

Lo dicho es verdadero en lo que corresponde al nivel alfabético (cercano al fonológico, al menos en algunos casos), lo es a nivel léxico y, casi lo es en lo que se refiere a las oraciones: si tomamos un texto como una sucesión de oraciones, éstas no son ya las unidades que se dan en un inventario finito sino elementos de un conjunto infinito⁵.

Como se ha apuntado, no cualquier sucesión de grafemas, sílabas, palabras u oraciones ha de aceptarse como un texto, pero ése es otro problema, y además arduo,

⁴ Es interesante señalar que, aunque pudiera concederse un valor, por ejemplo artístico, a ésta última, la misma no aportaría en principio ninguna información útil expresada en una lengua dada, por más que cualquier sucesión de instancias de símbolos transmitiría intrínsecamente alguna información en la medida en que seamos capaces de medir su probabilidad y de disponer de un código. Hay que reconocer, sin embargo, que la información no es el objetivo último de ciertos usos del lenguaje y, sobre todo, que el conocimiento –como diferente de la información– transmitido por un texto procede de la integración de la cadena textual en una estructura de conocimiento aportada por el sujeto lector.

⁵ Naturalmente, la gramática nos aporta el criterio para decidir si una cadena determinada es una oración o no, lo cual no es demasiado problemático, al menos, si mantenemos que la cardinalidad de ese conjunto no es mayor que la de los números naturales, al contrario de lo que parecen postular Langedoen y Postal (1984). Nos limitamos aquí a observar que su argumento se basa en oraciones de longitud actualmente infinita, lo que no sería el caso contemplado por la gramática. Es cierto sin embargo, que la gramática podría decirnos algo acerca de una cadena infinita si tuviéramos una regla de formación de la misma, lo que parece que nos vuelve a situar en el terreno de la cardinalidad de los naturales porque se admitiría una descripción finita de la misma en los términos propios de la gramática. Esta cuestión ha dado lugar a cierta literatura. En cualquier caso, la idea de la infinitud aparece entremezclada con la idea de recursividad. Como se sabe, la gramática generativa se basa en el carácter recursivo del conjunto de las oraciones gramaticales de una lengua. Aun aceptando ese postulado, no se sigue que el conjunto de los textos sea recursivo. Pese a eso, ciertas gramáticas del texto se construyeron sobre una inspiración generativista.

porque aunque se acepte que contamos con una gramática de las oraciones, no parece que ése sea el caso para con los textos en general, aunque sí, a cierto nivel de descripción para muy determinados tipos de textos⁶.

Notemos que el hecho de que un texto, una vez fijado, esté perfectamente ordenado, en cuanto secuencia de instancias de símbolos, significa a efectos prácticos que no hay ambigüedad en cuanto a qué unidades preceden o siguen a una dada para todas las unidades. Esto es, contamos con una aplicación, para decirlo en términos técnicos y de manera no exenta de ociosa pedantería, del conjunto de los enteros positivos al de las unidades que estemos considerando en el texto: no debe haber duda en esta cuestión para ningún texto. También con la dosis de pedantería correspondiente, cabría decir que la edición de textos consiste en alguna de sus partes en definir esa aplicación.

Otra manera de referir el asunto se basaría en la definición de unas relaciones entre las unidades del texto que llamaremos de contigüidad. De una manera bastante obvia se definirán unas relaciones de precedencia y de consecuencia: una unidad (sintagmática) de un texto se halla relacionada con otras por la primera y con otras más por la segunda. Se trata, en fin, de las relaciones sintagmáticas más primitivas y menos restringidas que pueda imaginarse y, por eso mismo, pueden definirse para cualquier nivel sintagmático de descripción. Nótese que a ciertos niveles, estas relaciones dejan de ser una vacía repetición de lo anterior, porque permiten precisamente aclarar cómo es la aplicación de referencia. E, incluso, son las relaciones no determinables de antemano las que nos señalan algunos de los límites de una hipotética gramática textual.

Con todo, si introducimos una relación de precedencia y de consecuencia inmediatas, la situación es un tanto diferente. Si nos referimos a las unidades inmediatamente contiguas a una dada, habrá de reconocerse, que con las excepciones de la primera y última unidades de la sucesión, cada una de ellas únicamente se relaciona con una sola por cada una de las relaciones: llegamos al portentoso hecho de que en un texto sólo pasamos de una unidad a otra con el salto más pequeño que podemos dar. Es más, con sólo una y siempre al menos con una, salvo en los casos señalados. Sin embargo, es en esta evidencia en lo que basaremos nuestra argumentación. Y, en fin,

⁶ Nos abstenemos de justificar este aserto mediante el repaso histórico de la lingüística del texto. Obsérvese, sin embargo, que los programas informáticos que generan textos, generan textos de un cierto tipo bien especificado.

dado que con la excepción del último, a cada instancia de las que componen el texto le corresponde sólo una mediante la dicha relación, nos encontramos con algo muy parecido a una aplicación o función del conjunto de las instancias de símbolos numeradas en sí mismo.

Suele hablarse de la unidimensionalidad del lenguaje escrito y hablado. Parecería que lo dicho supone una reformulación de tal cosa. Sin embargo, notemos que las relaciones de que hablamos implican además un carácter discreto del signo lingüístico.

3. CONTIGÜIDAD INMEDIATA NO UNITARIA.

¿Qué sucede si rompemos la condición de la unicidad de los elementos que relacionan a través de la relación que hemos establecido? –La respuesta es ya nuestro punto de llegada. Si se acepta que hemos numerado todas las instancias del hipertexto, de manera que la contigüidad inmediata que pueda existir se refleje en la sucesividad de la numeración, entonces el hipertexto es precisamente eso. Con las restricciones debidas y dadas a todos los niveles, incluidos algunos que le serían propios, el hipertexto sería una sucesión de unidades lingüísticas que ya no es tal, pues se ha roto la unicidad de la contigüidad inmediata: sería un conjunto de sucesiones, por decirlo de algún modo, que conforman una unidad, describible además como tal unidad en los términos genéricos de las relaciones de contigüidad. De una relación de orden total pasamos a un orden parcial en que no se cumplen ciertos supuestos habituales en el caso del texto.

En este último caso tendríamos que todas las instancias salvo la última tendrían una y sólo una instancia siguiente. En el caso del hipertexto tendríamos que algunas instancias tendrían dos siguientes⁷ y posiblemente habría más de una instancia que no tuviera siguiente.

Notemos que los segmentos que formarían el hipertexto de enlace a enlace, es decir, donde la contigüidad inmediata fuera unaria, serán textos en un sentido lato impropio, pero no sería ya sólo el caso que éste se compusiera de textos enlazados, ni siquiera de textos enlazados de cierta manera, sino que sería una misma familia de relaciones las que definirían a texto e hipertexto.

⁷ Supondremos que de cada origen se llega a un solo destino por un salto hipertextual. La argumentación no se resiente, sin embargo, si se considera que determinados orígenes presentan un “menú” de destinos.

Entonces, a cada instancia le correspondería un sucesor y, a algunas de ellas, dos sucesores. Para representar el hipertexto habrían de añadirse una serie de pares en que se reunieran los elementos origen de un enlace con los elementos destino.

Podría argüirse, sin embargo, que este planteamiento, que adoptamos, presenta la debilidad de que se ha hecho una ordenación de las instancias que puede ser arbitraria. A tal conclusión puede llevar la perspectiva de asimilar los enlaces externos a enlaces internos, a la que hicimos referencia más arriba. Sin embargo, es justamente la misma materialidad del hipertexto la que determina una numeración de las instancias, al menos en el caso en que sólo haya una primera instancia.

Nos encontramos en una situación parecida a la de un documento –por ejemplo, un anuncio que combine diversos bloques de texto e imagen– en que es difícil establecer el orden de las instancias de los símbolos alfabéticos. En este caso sería la misma materialidad del documento, la disposición del texto y la tipografía, lo que induciría un orden. Más claramente, en el caso del hipertexto podríamos numerar consecutivamente desde la entrada y mantenernos así hasta alcanzar la última instancia accesible sin hacer uso de los enlaces. Tal procedimiento podría cubrir o no la totalidad del hipertexto, dependiendo de su estructuración. Si no es así, habría que continuar numerando las instancias accesibles desde el primer enlace, etc. La descripción del hipertexto constaría de esa numeración más la lista de los pares ordenados que definiesen los enlaces hipertextuales.

La siguiente cuestión que puede suscitarse tiene que ver con la paradoja de la representación lineal del hipertexto y, en particular, con la reducción de una representación o implementación del hipertexto a una cadena, con las consecuencias que ello pueda tener para una definición de complejidad en el sentido de que se hablará más abajo.

4. LA REPRESENTACIÓN GEOMÉTRICA.

Desde el punto de vista de la teoría de lenguajes y gramáticas, la novedad del hipertexto queda por dilucidar y a ello dedicaremos la siguiente sección. Aquí diremos algo, por más que sólo aproximativamente, sobre la representación geométrica del hipertexto. Una oración gramatical es una cadena a la que le podemos, según

determinadas gramáticas, asignar un árbol⁸. Los elementos de aquella cadena serían hojas del correspondiente árbol. Un texto sería representable también por una cadena, y quizá por un árbol, pero no parece algo simplemente aceptable (véase nota 4) que éste constituya una descripción estructural del mismo de acuerdo con una gramática de los textos.

Resulta, en cualquier caso, tentadora la posibilidad de representar el hipertexto mediante un grafo. Ciertamente, en lugar de una cadena lineal, tendríamos una más compleja con una maraña de eslabones. Más en serio, lo más adecuado a la intuición es tomar las instancias de símbolos como arcos que conectan nodos a identificar con los espacios entre ellos. En el caso de un texto, un solo arco saldría de cada nodo. Ese no sería el caso para todos los nodos de un hipertexto. Ahora bien, si se quiere obtener una estructura más parecida a la habitual, podría optarse por una cadena que representase la numeración estipulada en la sección anterior para todas las instancias del hipertexto, a la que se añadirían los arcos que fuera preciso para representar el hipertexto. Alternativamente, se podrían separar las partes del hipertexto accesibles sin hacer uso de los enlaces y proceder de manera semejante. Obtendríamos una serie de cadenas que habrían de unirse entre sí y que podrían tener también arcos internos.

Nos parece, en cambio, más fructífero explorar las potencialidades de un grafo desarrollado a partir de un árbol, aceptando que éste es adecuado para representar la supuesta estructura jerárquica de un texto y su linealidad. Supóngase, pues, que un árbol representa algo parecido a una posible ordenación del hipertexto del tipo de las propuestas más arriba.

En el caso del texto o del hipertexto, frente a lo que ocurre en la gramática oracional, las relaciones sintácticas serían ampliamente rebasadas, como lo serían las de cualquier semántica restringida: las conexiones expresarían relaciones pertenecientes, por decirlo rápidamente, a la semántica enciclopédica. Con todo, algunas restricciones de carácter descriptivo podrían apuntarse. Así, en consonancia con lo enunciado más arriba, de cada hoja sólo habría de salir un arco extra, el cual podría tomarse como apuntando a una hoja más lejana. Si se concede que tenemos una gramática para generar los árboles sin arcos de esta última clase y que describiesen no ya oraciones sino también textos, una gramática del hipertexto debería dar cuenta de esos arcos extra y de

⁸ O un conjunto ordenado de ellos, que serían los marcadores de frase que representarían una derivación.

nada más. Sin embargo, esos arcos –que equivalen a seleccionar pares ordenados de hojas– no contienen información que no pudiera aparecer en un texto, según razonaremos más abajo. Si las relaciones de contigüidad inmediata vienen dictadas por una gramática de los textos, esto deberá ser válido también para un hipertexto. Formalmente, puede encontrarse un análogo en la operación *mover* de la gramática generativa y hemos de suponer que, en ese caso, la información que determina la ejecución de la operación se contiene en las hojas del árbol, esto es, en las piezas léxicas que lo van componiendo.

5. LENGUAJES DE MARCAS E HIPERTEXTO.

Un hipertexto aparece linearizado a cierto nivel de descripción (al menos aquél en que se almacena informáticamente si se forma un único archivo y, también, aquél que hemos propuesto aquí). Unos marcadores especiales señalan los puntos en que la contigüidad inmediata no es unívoca. Esto es posible porque sólo hay dos destinos y no más en estos casos.

Desde el punto de vista gramatical, un lenguaje de marcas constituye un caso especial de un lenguaje de paréntesis. Sin embargo, las restricciones que deben darse en un hipertexto de “interés humano”, no parecen abarcables por una gramática que sólo dé cuenta léxica y sintácticamente de la corrección de los paréntesis.

Tenemos que cada marca es un tipo de paréntesis. Sin embargo, origen y destino de los enlaces no suponen una suerte de paréntesis compuesto, sino pares de elementos que pueden aparecer entrelazados.

La pregunta es si estos pares pueden aparecer en una gramática que vaya más allá de la corrección de un correcto balance de paréntesis y pares; cuestión que formulamos simplemente para tratar de mostrar el grado de similitud entre un texto y un hipertexto. La tesis es que si hay una gramática del texto, entonces hay una gramática del hipertexto a este respecto. Naturalmente, lo que pedimos es algo superior a una gramática –que sí existe– en la que los elementos entre las marcas hipertextuales sean léxicamente correctos y la corrección de los paréntesis se determine por una gramática independiente de contexto bastante simple.

$$(1) \alpha(\beta)\gamma$$

En efecto, una estructura como la mostrada en (1), donde los paréntesis señalan dos partes no contiguas de una estructura, implica que existan en la gramática reglas que categoricen en forma de no terminales α , β y γ . Ahora bien esto es algo factible en una oración. Por ejemplo, para una estructura *si-entonces*, que asimilamos a un par de paréntesis, tenemos:

$$(2) \Sigma \rightarrow \text{si } \Sigma \text{ entonces } \Sigma$$

Un objetivo análogo no es accesible en un texto en general, salvo que se defina en el grado de abstracción y al nivel de descripción que sean precisos un tipo de texto (por ejemplo, los pertinentes en la Poética).

Ahora bien, si esto no es factible en el texto tampoco lo será en el hipertexto. A la inversa, si es factible en el texto, lo será en el hipertexto porque la información contenida en la regla correspondiente para el texto, puede aparecer en forma de regla para el hipertexto.

Sea (3) una regla para un hipertexto, donde los paréntesis indican el destino y el origen de un salto y las letras griegas símbolos no terminales:

$$(3) \rightarrow \alpha(\beta)\gamma$$

En términos de texto sería una regla como la siguiente, donde la barra vertical indica opcionalidad de los elementos ente corchetes:

$$(4) \Sigma \rightarrow \alpha[v_1|v_2]\delta$$

La transformación entre una y otra consistiría, al menos, en unir la regla (4) con la que dictase la selección de qué alternativa sería la contigua y cuál no. Ahora bien, lo dicho es válido para establecer la existencia de una conexión hipertextual. Sin embargo, su generación local concreta dependería de la presencia de rasgos compartidos por cuatro nodos superiores: los tres que se consignan en la regla, y el cuarto que correspondería al anclaje diferido del fragmento de texto conectado por un salto hipertextual:

$$(5) \alpha[v_1|v_2]\delta \rightarrow \alpha(v_1\beta)v_2\gamma$$

Ello se daría con la siguiente situación, en la que representamos la compartición de un rasgo ϕ_i :

$$(6) (\phi_i \in \alpha) \& (\phi_i \in v_1) \& (\phi_i \in v_2) \& (\phi_i \in \beta)$$

Ahora bien, tales rasgos no serían otros que los que definiesen relaciones textuales *ad hoc*. Se ve así cómo una gramática del hipertexto de dejaría reducir a una gramática del texto, siempre supuesta ésta. El interés del razonamiento no reside en

apuntar hacia la prueba de algo bastante intuitivo, sino más bien en mostrar la homología a este respecto entre texto e hipertexto.

En general, puede pensarse que en términos de gramáticas el problema reside en la necesidad de un número infinito y abierto de no terminales, con la consiguiente infinitud de la gramática. La apuesta de una gramática oracional generativa sería, precisamente, que el número de tales símbolos sería finito y que podría desconectarse la generación de una lengua de valores semánticos de carácter inabarcable o no computable. Lo que se prueba, paradójicamente, en la tesis de la semántica cognitiva según la cual las reglas de la sintaxis derivan de la semántica, porque lo que se hace precisamente es reducir a unas cuantas situaciones la complejidad de la semántica.

6. UNA NOTA SOBRE COMPLEJIDAD.

Esta sección y la siguiente poseen un carácter tentativo que no disimulamos. Se trata aquí de comenzar a definir una complejidad puramente hipertextual. Dejamos de lado, medidas de complejidad que se refieren al texto, como puedan ser las bien conocidas que se expresan por el índice de Tuldava u otros similares. El lector puede acudir como referencia a Hřebíček y Altmann *eds.* (1993) y a Bernárdez (1995). Tales medidas no suelen precisamente considerar los aspectos del texto que se refieren a conexiones no locales y que serían precisamente las que están detrás de los razonamientos contenidos en la sección anterior.

En cuanto a las medidas de complejidad utilizadas en otros campos, existen diversas posibilidades en la literatura. Por ejemplo, las representaciones gráficas pueden inducir a utilizar el número ciclomático⁹ como medida de la complejidad de un hipertexto representado como grafo según la descripción apuntada más arriba. A la fórmula usual consignada en (7), que presentamos aquí sin ningún coeficiente, replicamos, en una primera aproximación, con la expresada en (8), donde W es el número de unidades no accesibles sin mediación de salto hipertextual y que, por tanto, caracteriza a un hipertexto en su plasmación concreta para una tecnología dada:

$$(7) C = \text{número de arcos} + \text{número de unidades} - \text{número de nodos}$$

⁹ El número ciclomático se propone medir la complejidad de un programa informático. Hacemos aquí una traslación quizá no demasiado adecuada.

$$(8) C = \text{número de saltos hipertextuales} + W - \text{longitud total del hipertexto}$$

Puede verse que (8) da importancia a la cuestión de la accesibilidad y renuncia a representar el hipertexto como una única unidad. Al proceder así parece que se devalúa la complejidad de las conexiones entre las unidades aisladas. Si se hace $W = 1$, el número de enlaces aumentaría al menos en $W-1$. En consecuencia, (9), en donde hemos añadido unos coeficientes sin especificar a los términos, parece mejor:

$$(9) C = k_1 * \text{número de saltos hipertextuales} + k_2 * \text{número de saltos entre unidades} - k_3 * \text{longitud total}$$

La utilidad de este índice queda por probar. Vamos a proponer también la definición de otro conjunto de índices de un modo tentativo. En principio, parece pertinente considerar al respecto cuestiones tales como:

- a. La distribución de las clases de símbolos. El número de enlaces promediado por el número total de instancias proporciona una información sencilla de obtener y que se relaciona con esta medida.
- b. La ubicación de los mismos en la cadena del hipertexto. Podíamos asimilar el grado de máxima complejidad a que la distribución de los enlaces en el texto sea uniforme.
- c. La distribución de los incrementos que suponen los saltos hipertextuales.
- d. La correlación de los dos últimos.
- e. La accesibilidad sin enlaces del hipertexto.
- f. Número de caminos alternativos para llegar a un punto dado, que está relacionado con el número ciclomático.

Si ponemos que hay determinada probabilidad de que una instancia sea un enlace, vemos que (a) tiene que ver con la entropía de Shannon del hipertexto, mientras que (b) con la complejidad algorítmica. De cualquier modo, en el caso de la entropía de Shannon, si f_i es la frecuencia de instancias que representan enlaces y f_j la frecuencia de la clase de todos los otros símbolos, esta entropía S es:

$$(10) S = -(f_i \log_2 f_i + f_j \log_2 f_j)$$

S tiene un máximo para $f_i = f_j = 1/2$. Es de notar que los hipertextos interesantes deberán tener una proporción de enlaces ni muy alta ni muy baja, pero es

evidente que esto implica $f_i \ll f_j$. En este sentido, cabe en la práctica, fijar un número de “unidades temáticas” en el texto en el grado de detalle que se considere adecuado. Si son estas f_k , puede utilizarse una fórmula como:

$$(11) S = -(f_i \log_2 f_i + f_k \log_2 f_k).$$

La estimación de la complejidad algorítmica es cuestión bastante más complicada. Es siempre posible aplicar fórmulas como las presentadas en Balasubrahmanyan y Naranan (1996) Sin embargo, nos abstenemos aquí de investigar esta posibilidad. En cualquier caso, dado que estamos distinguiendo dos clases de símbolos, lo que nos interesa es la descripción de una secuencia de las dos clases de símbolos. Si los símbolos que corresponden a un salto hipertextual están concentrados en alguna parte de toda la cadena, o aparecen a intervalos regulares deberíamos adjudicarle una baja complejidad. El hipertexto podría representarse, entonces, por un vector de números que representasen las rachas de la primera clase de símbolos. Su complejidad podría ahora ser medida según los criterios prácticos que se juzgasen más oportunos. Por ejemplo, sería razonable categorizar tales números en un número más bien discreto de categorías de tal manera que el número de vectores posibles fuera finito. La agrupación de enlaces correspondería a una racha de, por ejemplo, ceros. Una distribución regular a una racha del mismo valor. Esos serían casos de baja complejidad.

En (c) se introduce una consideración capital. De hecho, la longitud de los saltos, positiva o negativa con respecto a su aparición en la cadena, ha de ser, intuitivamente, tan o más importante que el número de los mismos. Los incrementos serían un vector de números positivos o negativos. El caso de mayor complejidad sería aquél en que no se registrase ninguna tendencia temporal, salvo las obligadas por los límites impuestos por la proximidad del comienzo o el fin del propio hipertexto. Una aproximación factible sería dividir los símbolos de enlace o salto en varias categorías discretas según la longitud del salto. La información sobre este particular podría así integrarse en las fórmulas de la entropía. Información similar se obtendría de (d).

El punto (e) sería como W de la fórmula (8). Si consideramos, en lo que respecta a (f) el número de caminos posibles para pasar de una posición x a una posición y , es razonable considerar los caminos que lleven sólo por arcos representados en el hipertexto. Para un hipertexto, podemos postular el sistema de ir siempre símbolo a símbolo hacia delante en el texto aunque pudiendo ir hacia atrás en los saltos hipertextuales, y añadiendo un arco de los últimos símbolos al primero. Un texto habría

de ser de complejidad nula. Si sólo se permiten los movimientos hacia delante de manera que sólo se pase una vez como máximo por un símbolo sin siguiente y por cada salto, tendremos que, para un par de símbolos cualesquiera el número de caminos se calcularía sobre la combinatoria de los puntos de salto. Tendríamos un máximo de $n!$ desde 0 a n caminos para hipertextos con n enlaces. Como, en cualquier caso, obtendríamos para cada posición tantos valores como sea la longitud del texto menos uno, la información consiguiente se habría de representar en una matriz de proporciones notables. Ahora bien, para símbolos no separados por símbolos de salto, el valor sería el mismo en cualquier caso. Parece deducirse que la información que se consignase de tal forma sería redundante, en cuanto a medida de complejidad, con la expresada en los puntos anteriores y sería fácilmente recuperable del correspondiente grafo. Sin embargo, no incluiría el tipo de información que se presenta en a, b y c.

7. COMPLEJIDAD DE UN HIPERTEXTO PARA UN MODELO DE HIPERTEXTO DADO.

Por último, trataremos de integrar lo dicho en las dos secciones anteriores en una propuesta más práctica. Tomemos que un tipo de hipertexto se deja describir a cierto nivel de representación como un árbol más unos arcos que unen pares de hoja. Nótese que esto no presupone la existencia de una gramática. Se señala que se puede asignar una descripción estructural a un hipertexto, pero no un procedimiento para hacerlo. Supóngase que ese tipo de hipertexto incluye un número de saltos hipertextuales en un subconjunto especificado de sus hojas, o de sus pseudohojas, pues el tipo es abstracto. Serían precisamente esas hojas las que lo convertirían en un hipertexto. Si se describe la complejidad de esa distribución se estaría caracterizando la complejidad del tipo, la cual tendrá interés en particular si se comparan dos tipos de hipertextos. De cualquier manera, puede darse una medida de esa complejidad típica según la idea que se expone a continuación. Los lugares donde aparecen saltos hipertextuales serán no sólo hojas, sino también nodos intermedios que dominen tales hojas. Obsérvese que para un nivel determinado un nodo dominará 0, 1, o más enlaces. Sin embargo, existirá o no la posibilidad de describir el hipertexto sintácticamente de manera que esa descripción sea económica y prescriba en qué lugares se hallan los enlaces. Entonces, la complejidad del tipo dependerá de la existencia o no de una

descripción especialmente breve. Notemos en que al hablar de tipo hacemos abstracción del encaje que un hipertexto concreto tenga en esa estructura. Como ejemplo de situaciones de baja complejidad, puede pensarse en un árbol en que todas las hojas que contengan saltos hipertextuales estén dominadas por un mismo nodo intermedio, o bien en un árbol en que a cierto nivel intermedio encontremos repetida una misma estructura.

Esta medida de complejidad posibilita comparar la de un hipertexto individual con la del tipo al que se adscribe. Como suponemos tener la caracterización de la estructura típica, lo único que queda es establecer las adecuadas medidas de distancia entre los valores que se tomen en un caso y en otro¹⁰.

Esta estrategia presenta la ventaja de que podría utilizarse para estimar la complejidad en un marco en que se neutralizase la complejidad conocida (la del tipo), de manera que el resultado fuera un indicativo de la dificultad subjetiva del hipertexto para un usuario familiarizado con el tipo de referencia.

Por otro lado, fijar el modelo puede hacerse empíricamente a la vista de la distribución de los valores hallados en una muestra que se juzgue representativa, al menos si esa distribución cumple determinados presupuestos.

¹⁰ Lo cual no ha de presentar excesiva dificultad. Hacemos gracia al lector de mencionar la amplísima literatura sobre estas cuestiones, la cual en la mayoría de los casos alcanza una dificultad técnica notable.

Referencias bibliográficas

- AARSETH, Espen, J. (1997). *Cybertext. Perspectives on Ergodic Literature*. Baltimore: The Johns Hopkins University Press.
- BALASUBRAHMANANYAN, V.K. y NARANAN, S. (1996). «Quantitative Linguistics and Complex System Studies», *Journal of Quantitative Linguistics*, Vol. 3, No. 3, 177-228.
- BERNÁRDEZ, E. (1995). *Teoría y epistemología del texto*. Madrid: Cátedra.
- EDMONDS, B. (1999). *Syntactic Measures of Complexity*. Ph. D. Thesis submitted to the Department of Philosophy, Faculty of Arts, University of Manchester, <http://bruce.edmonds.name/thesis/>
- HREBÍCEK, L. y ALTMANN, G. eds. (1993). *Quantitative Text Analysis*. Trier: WVT Wissenschaftlicher Verlag.
- JOYANES AGUILAR, L. (1997). *Cibersociedad: los retos sociales ante un nuevo mundo digital*. Madrid: McGraw-Hill/Interamericana de España.
- LANDOW, G. P. comp. (1997). *Teoría del Hipertexto*. Barcelona: Paidós.
- LANGEDOEN, D. T. y POSTAL, P.M. (1984). *The Vastness of Natural Languages*. Oxford: Basil BlackWell.