

ALGUNAS CONSIDERACIONES SOBRE LA PROPUESTA DE ANÁLISIS DE LA COHESIÓN LÉXICA DE M. HOEY

ANA DÍAZ GALÁN

Universidad de La Laguna

43

1. Introducción

En una reseña del volumen editado por Scott y Thompson *Patterns of Text: In honour of Michael Hoey* (2001) difundida a través del foro de discusión de Internet *The Linguist List*, el autor de la misma, Mario Saraceni (2002), concluía su resumen destacando cómo la contribución de M. Hoey al entendimiento de la organización del texto y del papel de la cohesión en la misma ha sido tan importante como ignorada. El modelo de análisis de la cohesión léxica de este autor, explicado en detalle en *Patterns of Lexis in Text* (1991) y que nos proponemos revisar aquí, es uno de los avances más significativos en el desarrollo de este concepto desde las publicaciones de Hasan (1984, 1989). Lejos, sin embargo, de haber tenido la resonancia que merece en los círculos teóricos del análisis del discurso, ha logrado mayor repercusión dentro de la lingüística aplicada, especialmente en el terreno de la didáctica (Díaz Galán y Fumero Pérez 1997, 2003), hecho que ha sido además fomentado por el propio autor (Hoey 2001), y, sobre todo, en el campo del estudio de la lengua por medios informáticos. La posibilidad de implementar informáticamente este análisis de la cohesión léxica ha sido ampliamente aprovechada para su aplicación en terrenos como la realización automática de resúmenes (Benbrahim y Ahmad 1994; Renouf y Collier 1995) o para la segmentación de textos pertenecientes a *corpora* (Berber Sardinha 1997, 2001).

De un modo condensado, puede decirse que la finalidad del análisis de la cohesión léxica diseñado por Michael Hoey es demostrar que las conexiones (*links* 1991: 51) entre los elementos léxicos de un texto son fundamentales para la organización global del mismo. Siguiendo a Winter (1979), el autor defiende que el texto se organiza por medio de la repetición de la información contenida en las oraciones, principalmente a través de distintos tipos de reiteración léxica y, en menor medida, por el uso de elementos gramaticales. A efectos prácticos, su propuesta de análisis se divide en dos fases que podríamos denominar, por una parte, de identificación y clasificación de conexiones y, por otra, de creación de matrices y descripción de la organización del texto.

Este artículo está basado en los resultados de una investigación previa (Díaz Galán 2004a) en la que el mencionado método se aplicó de forma manual a un corpus de noticias de prensa¹, en el que se mostró absolutamente adecuado para la descripción de su organización e, incluso, en contra de la opinión de Hoey (1991: 29) al respecto, de su estructuración (Díaz Galán 2002, 2004b). Su uso deparó, no obstante, ciertas dificultades prácticas, especialmente en la primera fase, que merecerían alguna revisión o que, cuando menos, parecen ser dignas de discusión en tanto que pueden causar algunos problemas de aplicación especialmente relevantes para aquellos interesados en la implementación informática de este análisis o, simplemente, para quienes pretendan profundizar en el mismo. Intentaremos aquí ponerlas de relieve y, para ello, en las siguientes secciones, a la par que repasamos los pasos de esta fase inicial para aquellos no familiarizados con este análisis, indicaremos los problemas surgidos de su utilización. Aunque es nuestra intención tratar fundamentalmente los aspectos prácticos del método por medio de las propias afirmaciones del autor, especialmente las vertidas en el volumen de 1991, haremos, en ocasiones, referencia a ciertas cuestiones teóricas que se abordarán, al ser el propio Hoey un funcionalista, desde la perspectiva de dos de las escuelas funcionalistas europeas más relevantes: la Gramática Sistemico-Funcional (GSF) de Michael Halliday, con reconocida trayectoria en la investigación del discurso, y la Gramática Funcional (GF) de Simon Dik, cuyo empuje investigador más reciente se centra también en este campo.

44

2. La definición ortográfica de oraciones y elementos léxicos

Hoey concibe la organización del texto como una red de conexiones entre sus oraciones. Al poseer éstas un significado independiente y, en ciertos casos, clave para el texto, el autor (1991: 78), no sólo las define como “packages of information, a miniature text”, sino que afirma acerca de ellas que son la base de su propuesta.

Siendo tan relevantes no es de extrañar, por tanto, que el primer paso de su análisis consista en la división del texto en oraciones; sí es en cambio más chocante, sobre todo para aquellos no familiarizados con las teorías de la llamada Escuela de Birmingham, que su concepción de la oración no sea sintáctica o semántica, sino, literalmente, ortográfica, esto es, una oración se comprende entre una mayúscula y un punto.

A pesar de que, inicialmente, uno pueda acercarse a este concepto ortográfico de la oración con cierta reticencia por sus implicaciones teóricas, lo cierto es que resultó plenamente adecuado tanto en la fase inicial del análisis como en la segunda de creación de matrices y descripción de la organización textual. Según el autor predice, las oraciones ortográficas demuestran tener, por el hecho de presentar una información completa, un valor textual que en el caso de la prensa se comprobó (Díaz Galán 2002, 2004b) al coincidir las oraciones más relevantes con las categorías más significativas de la noticia (generalmente *resumen* y *evento principal*, en términos de la clasificación de van Dijk 1980, 1986).

No podemos decir lo mismo, sin embargo, de la definición ortográfica de los elementos léxicos que propugna el autor (1991: 266). La relación entre las oraciones del texto se materializa a través de la conexión entre sus elementos léxicos, conexión que se realiza, principalmente, a través de distintos tipos de repetición. Distinguir, por tanto, cuáles son los elementos léxicos del texto analizado es fundamental para no incurrir en errores que afectarían decisivamente a los resultados de la segunda fase, la organización textual, basada precisamente en la cuantificación numérica de estas conexiones. Hoey, sin embargo, es bastante impreciso a la hora de delimitar qué es un elemento léxico, limitándose a indicar que éstos son de clase abierta y que pueden estar representados por una o más palabras ortográficas, siempre y cuando sean unidades “estables” y tengan un mismo referente. Sin embargo, la dificultad para distinguir entre elementos léxicos, sintácticos o unidades referenciales no se supera acudiendo meramente a la ortografía, como el mismo autor (1991: 266) prevé:

Lexical items have been treated as co-terminus with words defined orthographically. Current research into idioms and collocations [...] suggests this is not an entirely safe practice, and there are certainly going to be problems on occasions deciding whether a particular two-word phrase represents one or two repetitions, depending on whether it is treated as two separate lexical items or as comprising a single lexical item made of two orthographically defined words.

Este problema, patente para cualquier analista desde el primer momento, conduce, en muchas ocasiones, a tener que tomar al respecto decisiones basadas en la intuición. En algunos casos, por ejemplo, no es posible establecer los elementos léxicos integrantes de las repeticiones atendiendo a su pertenencia o no a una clase

abierta de palabra, sino según criterios sintácticos, es decir, contabilizando grupos o sintagmas; un ejemplo de ello es la conexión entre *hear* y *was told* en las oraciones uno y dos del texto 7c “*a court heard yesterday*” y “*a Worcester Crown Court was told*”, donde *hear* y *told* individualmente no estarían, de acuerdo con los criterios de Hoey, relacionados y es sólo la presencia del auxiliar la que los conecta en ese contexto. Otra muestra de este problema son los nombres propios comunes, ya que los casos en los que encontramos compuestos como *Gulf War* (textos 1a,1b), *Grantham Hospital* (texto 6a) o *Professional Pool Players Association* (texto 8a) repetidos posteriormente por sólo uno de los elementos del grupo inicial (*war*, *Grantham* o *professionals*) nos obligan a decidir si estamos ante uno o dos elementos. Ante este tipo de ejemplos, y a falta de indicaciones claras del autor, no parece haber más opción que recurrir a decisiones individuales, utilizando la identidad de referencia o la similitud contextual o colocacional para contabilizarlos o no como repeticiones y confiar en que el establecimiento de este tipo de repeticiones, en muchos casos *ad hoc*, no afecte al resultado final del análisis, merced al mecanismo de agrupamiento de conexiones, es decir, por el hecho enfatizado por Hoey de que una conexión individual no es significativa a menos que la presencia de otras, como mínimo dos más, la convierta en un *vínculo, bond* (Hoey 1991: 91), con valor textual. Este criterio, aunque de gran utilidad en la mayoría de las ocasiones, no deja de ser sólo una solución parcial a un problema que no por conocido deja de ser relevante.

46

3. Detección y clasificación de repeticiones

Una vez delimitados los elementos léxicos, el siguiente paso del análisis consiste en la detección de las repeticiones y de sus tipos contrastando todos los elementos léxicos del texto entre sí de forma anafórica. La tabla que aparece más adelante (véase Tabla 1) resume y ejemplifica los distintos tipos de repetición propuestos por Hoey (1991: 51-75).

Debe notarse que cuando dentro de una misma oración más de un elemento repite a otro previo se contabilizará sólo una conexión entre las oraciones, lo que implica decidir cuál de estas repeticiones es más relevante. Este proceso ha de hacerse siguiendo la llamada *jerarquía de repeticiones* (véase Tabla 3), una escala que da prioridad a las conexiones léxicas sobre las no estrictamente léxicas² (correferencia, sustitución o elipsis) y, dentro de las primeras, a las más diáfanos (como la repetición idéntica o repetición simple).

De la estadística de los textos que hemos analizado (véase Tabla 2) parece deducirse que la conexión entre elementos, es decir la cohesión de un texto, desde un punto de vista general, se sustenta mayoritariamente, según predice también el

Algunas consideraciones sobre la propuesta de análisis de ...

Tipo de repetición	Abreviatura	Ejemplo
Repetición Léxica Simple o Repetición Simple	rs	i.e. bear-bear
Repetición Léxica Compleja o Repetición Compleja	rc	i.e. human (nombre) human (adjetivo)
Paráfrasis Simple Mutua (aplicable en los 2 sentidos)	psm	i.e. produce-cause
Paráfrasis Simple Parcial (aplicable en 1 sentido)	psp	i.e. volume-book
Paráfrasis Compleja	pc	i.e. author-writings
Repetición Correferencial (+ superordinados e hipónimos)	cor	i.e. Augustus- The Emperor
Sustitución	s	i.e. author-he
Elipsis	el	i.e. author-∅

TABLA 1: Tipos de repetición del modelo de Hoey, abreviaturas y ejemplos

autor, en la repetición de elementos léxicos en cualquiera de sus variantes (79% de las repeticiones), o sea, en la cohesión léxica, y no en recursos asociados a ella como la correferencia, la sustitución o la elipsis (21% de las repeticiones). Más específicamente, parece ser la denominada repetición léxica simple la que articula la unidad de los textos.

Aun cuando es evidente que Hoey favorece la denominada repetición simple frente a otros recursos cohesivos, según demuestra la jerarquía de repeticiones, no podemos obviar el hecho de que éste es el tipo más utilizado en todos los textos, habiendo sido sólo superado en una ocasión por la correferencia (texto 3c) e igualado en otras dos con la sustitución (textos 10b y 10c).

A la luz de los porcentajes globales que hemos sintetizado en la Tabla 2 se deduce, por tanto, que la repetición simple supera a cualquier otro tipo de relación léxica, pero, también, que no existen indicios claros para establecer cuál es el segundo tipo más frecuente, ya que, si bien hay una diferencia de más del treinta y cinco por ciento entre la repetición léxica simple y el resto, es igualmente cierto que la

Tipo de repetición		Léxica					Otros			
todos	rs	rc	ps	psm	psp	pc	cor	s	el	
N1	3689	1756	461	443	162	281	254	464	296	15
%	100	47,6	12,4	12	4,4	7,6	6,9	12,5	8	0,5

TABLA 2: Número y porcentaje global de los tipos de repetición léxica del corpus

correferencia, la paráfrasis simple (mutua o parcial) y la repetición compleja sólo están separadas por un estrecho margen porcentual. Es evidente, también, la menor frecuencia de la paráfrasis compleja, la sustitución y, sobre todo, la elipsis. Comparando los resultados arriba expuestos con la jerarquía de repeticiones propuesta por Hoey obtenemos un orden paralelo en el que llama la atención la posición de la correferencia y de la sustitución, las cuales, individualmente, parecerían tener en el análisis de la cohesión léxica de este autor un papel más destacado que el que se les quiere reconocer:

	Jerarquía de repeticiones propuesta por Hoey	Tipos de repetición más frecuentes en el análisis del corpus
1	Repetición Léxica Simple	Repetición Léxica Simple
2	Repetición Léxica Compleja	Correferencia
3	Paráfrasis Simple Mutua	Repetición Léxica Compleja
4	Paráfrasis Simple Parcial	Sustitución
5	Paráfrasis Compleja	Paráfrasis Simple Parcial
6	Sustitución	Paráfrasis Compleja
7	Correferencia	Paráfrasis Simple Mutua
8	Elipsis	Elipsis

TABLA 3: Comparación entre la jerarquía de repeticiones y los tipos de repetición más frecuentes en el análisis del corpus.

4. Correferencia, sustitución y elipsis

Precisamente la detección de la correferencia y la sustitución, además de la elipsis, motiva muchos de los problemas analíticos derivados de la aplicación manual de este método. Al centrar el autor sus intereses en las repeticiones de tipo léxico, estas otras clases de reiteración, que para su modelo son categorías menores³, parecen presentar algunos aspectos inacabados. Por ejemplo, la barrera que separa la correferencia de la sustitución es, en muchos casos, permeable y, así, en la práctica y debido a que las conexiones se contabilizan sólo de forma anafórica, una relación que inicialmente es una sustitución puede convertirse en menciones posteriores en una correferencia (Hoey 1991: 86). Esto puede comprobarse en las siguientes conexiones del texto 3b:

- (1) *Calderwood* claimed she listened as Herron and Green talked about making a hoax call to Burger King (or.8)

“This young lady lost her job for a stupid act to which *she* would say no” (or.13)
Calderwood, admitted making the hoax bomb call last December (or.15)

conexiones:

s: Calderwood - she (ors.8-13)
cor: she - Calderwood (ors.13-15)

De la misma manera que este hecho no parece un demérito, sino, por el contrario, una adaptación al contexto en el que tiene lugar la conexión, creemos que es posible realizar, en estas dos categorías, algunos cambios que facilitarían el análisis práctico sin restarle importancia al aspecto léxico de la cohesión.

Hoey incluye dentro de la correferencia relaciones como la hiponimia o superordinación que otros autores como Halliday y Hasan (1976), Hasan (1984, 1989) o Halliday (1994) consideran no directamente dependientes de la correferencia. Al contemplar sólo los casos en los que existe correferencia, el análisis de Hoey se ve obligado a descartar muchas conexiones léxicas basadas en la composición (i.e. *cheek-ears*, texto 2b) o, simplemente, en la igualdad de campo semántico, (i.e. *magistrates-trial*, texto 6c; *dustmen-rubbish*, texto 7b, etc.). Aplicando el análisis llegamos a la conclusión de que algunos de estos casos podrían ser contemplados en el recuento de repeticiones léxicas como una paráfrasis simple (o sea, de forma similar a un sinónimo), en su variante mutua o parcial, sin alterar lo que Hoey entiende por tales. Como ejemplo podemos utilizar tres oraciones del texto 2b:

- (2) TV ad scrapped after *school* slapping craze (or. título)

A MILLION-POUND TV advertising campaign has been dropped because it led to *playground* violence. (or.1, entradilla)

But 28 complaints received from around the country protested to the Independent Television Commission that children were copying the antics, shouting the advertising slogan as they slapped their *playmates*. (or. 5)

La relación entre el todo y la parte de los elementos *school* y *playground* del título y la entradilla, obvia para cualquier lector, no puede, sin embargo, contabilizarse como una conexión al no compartir los elementos el mismo referente. Evitar la conexión por esta razón no parece lógico, puesto que es evidente que funcionan como sinónimos, es decir, como paráfrasis simple, en este caso, mutua. Esta alternativa al procedimiento de Hoey se aplicó con resultados positivos a otros ejemplos: *drugs-LSD* texto 7a, ors. tí-10; *court-jury* texto 7b, ors.1-7; *paper-sheets* texto 7b, ors.2-13 o *cue-pool* texto 8c, ors. tí-1.

Esta posibilidad, además de tener en cuenta el aspecto léxico de estas relaciones (en línea con Hasan 1984, 1989 o con Martin 1992), nos ofrece, también, la oportunidad de contabilizar terceras conexiones, que de otro modo serían imposibles

de considerar, como en el ejemplo que acabamos de ver (2), la existente entre *school* (or. título) y *playmates* (or. 5). Este último procedimiento se llevaría a cabo por medio del denominado “triángulo de conexiones”, un mecanismo diseñado por el autor para dilucidar si dos elementos que parecen subjetivamente conectados lo están realmente. El funcionamiento de este artefacto, que se utiliza de forma bastante restringida y del que nos volveremos a ocupar, consiste en la conexión de dos elementos por mediación de un tercero común a ambos. Es decir, si al encontrarnos ante tres elementos *a*, *b* y *c* podemos establecer una conexión (únicamente de repetición compleja o de paráfrasis simple) entre *a* y *b* y otra entre *a* y *c*, existirá una tercera entre *b* y *c*. En el ejemplo que tratamos (2), la paráfrasis simple entre *school* y *playground* y la repetición compleja entre este último y *playmates* nos daría pie a reconocer la paráfrasis compleja entre *school* y *playmates*. Debemos decir, sin embargo, que, aunque este procedimiento es de gran utilidad, sigue siendo, de nuevo, una solución parcial, puesto que no logra capturar muchas relaciones léxicas evidentes desde un punto de vista subjetivo (muchos casos de lo que Halliday y Hasan en su obra de 1976 habrían clasificado como “colocación”: *cheek* y *ears*, texto 2b; *school* y *teachers* texto 2b, etc.) y es, por tanto, un aspecto susceptible también de mejora.

50

4.1. Los elementos textuales

También los llamados *elementos textuales*, aquellos cuya misión consiste exclusivamente en repetir una parte más o menos extensa del texto, parecen encontrarse, en nuestra opinión, a medio camino entre la correferencia y la sustitución. Hoey (1991: 71), por su parte, incluye dichas conexiones dentro de la sustitución, argumentando lo siguiente:

Textual items [...] entirely depend on other items for their interpretation and have no definitional meaning in themselves. They are grammatical members of closed systems and not lexical items.

Al definirlos de esta manera el autor descarta, por ejemplo, el valor cohesivo de los nombres o sustantivos anafóricos, *a-nouns* (Francis 1986, 1994), del tipo *ultimatum* (texto 1b), *explanation* (texto 2a) o *behaviour* (texto 2c). Estos sustantivos anafóricamente cohesivos, según Francis (1986: 31), tienen un uso metadiscursivo y están caracterizados por funcionar de una forma similar a las proformas, decir, resumiendo el contenido de una o más oraciones anteriores, pero sirviendo, a la vez, como información dada a partir de la cual se hará el engarce de una nueva información. El siguiente es uno de los ejemplos de nuestro corpus:

- (3) *Britain and the US warned Iraq of a new war unless it eliminates its remaining weapons of destruction.* (texto 1b, or. 1)

The *ultimatum* came as Saddam Hussein's envoy Tariq Aziz was given a dressing down over his country's continued defiance of the terms of the Gulf War ceasefire. (texto 1b, or. 2)

El tratamiento de Hoey, en casos como este, ignoraría el propio sustantivo (aquí *ultimatum*) en favor del determinante que suele acompañarlos (*the*), y prima así, bien a la deixis, bien a la sustitución, por encima de la repetición léxica. De esta manera, el método obvia no sólo la función anafórica de este tipo de sustantivos, sino también su naturaleza de elementos léxicos, por lo que, según se deduce de los propios presupuestos del autor (recuérdese la jerarquía de repeticiones), se contradice doblemente.

Por estas razones y teniendo en cuenta, además, que los sustantivos anafóricos cumplen los requisitos impuestos por Dik (1997a,b) y Keizer (1991) para ser considerados elementos con referencia⁴, creemos que podrían incluirse dentro del grupo de las repeticiones correferenciales. Basándonos en las categorías cohesivas, tanto las ampliamente conocidas de la GSF (Halliday y Hasan 1976, 1989, Hasan 1985, Martin 1992, 2001), como las que está empezando a proponer la GF (Hengeveld 1997, 2004), consideramos que estaría fundamentado el tratamiento de los elementos textuales, léxicos o no, como casos de correferencia y no de sustitución.

En contraposición con la opinión de estudiosos de la cohesión desde un punto de vista textual, como Halliday y Hasan (1976, 1989), o de la referencia desde un enfoque gramatical, como Keizer (1991) o Dik (1997a), otras posturas no especializadas no parecen el modelo más idóneo a seguir en este punto, por lo que debemos disentir del parecer de Hoey (1991: 71) expresado en este párrafo:

Their [textual items] function is to stand in, or substitute, for lexical items. For this reason we will follow Quirk *et al.* (1972) in terming them *substitution* links. Halliday and Hasan (1976) use the label *substitution* to apply to a narrower class of items, using the label *reference* to describe most of the items we list here.

Si se siguiera el criterio que proponemos, los elementos con referencia textual se considerarían casos de repeticiones correferenciales, y se podría, por tanto, detraer del grupo de sustitutos de Hoey, no sólo tales nombres o sustantivos anafóricos, sino también los pronombres demostrativos y la combinación del verbo *do* con elementos como *that* o *it*, con idéntica función textual. En este punto seguiríamos, además, la opinión al respecto de Halliday y Hasan: “the expressions *do that*, *do it*, in fact function as reference items [...] an occurrence of *do that* constitutes a single cohesive tie, not two” (1976: 125-126). En nuestro corpus sólo encontramos un caso inter-oracional de este tipo de conexión (en cursiva):

(4) The caller told Burger King bosses *that a bomb would be going off and then hung up*. (texto 3b, or.11)

She *did it* because she was dominated by Herron and was so pressured that she found it difficult to refuse. (texto 3b, or.14)

4.2. El ámbito de influencia de las conexiones

La correferencia trae aparejado un segundo problema: su ámbito de influencia. Las repeticiones con referencia textual, por ejemplo los demostrativos, conllevan, por su propia naturaleza, la repetición de una parte, más o menos extensa y más o menos próxima, del texto. Es decir, que su influencia puede ser tanto uno de los constituyentes de la oración anterior como un párrafo entero donde se incluyan varias oraciones (en los ejemplos anteriores (3) y (4) una oración independiente y una subordinada, respectivamente). Esta variedad en la extensión y naturaleza del referente de los elementos no es tenida en cuenta, sin embargo, por Hoey, que, en estos casos, independientemente del número de oraciones repetidas, sugiere anotar una única conexión y siempre con el elemento más cercano. Tal tratamiento lleva a unos resultados incongruentes desde un punto de vista no sólo semántico sino también teórico (téngase en cuenta que el método enfatiza el valor de la redes de relaciones y no de las conexiones individuales) que nos inducen a sugerir la posibilidad de que se anoten individualmente cada una de las oraciones repetidas por un mismo elemento textual. El resultado de este procedimiento, en nuestro corpus, no afectó a la segunda fase del análisis que mencionamos anteriormente, es decir, no repercutió en la descripción de la organización de los textos, de lo que deducimos que no sería completamente desacertado. Reflejar, sin embargo, el referente de los elementos cuyo ámbito es menor que el de la oración por medio de los tipos de reiteración establecidos por Hoey no es posible y, puesto que el autor sólo contempla la vinculación entre oraciones completas, no podemos censurarle en este punto.

No son estos, de cualquier modo, los únicos problemas concernientes al ámbito de influencia de las conexiones, ya que si la referencia a más de una oración no se tiene en cuenta adecuadamente, tampoco son completamente convincentes las conexiones de correferencia o de sustitución de más de un elemento, esta vez dentro de una misma oración. Como norma general, para evitar un recuento doble, cada elemento debe establecer una única conexión por oración, aplicándose, en aquellos casos en que haya varias opciones, la jerarquía de repeticiones que conocemos. Este procedimiento, generalmente útil, no es válido, sin embargo, para la correferencia y la sustitución, ya que, con frecuencia, encontramos casos como el del sustantivo *pair* (3a, or. 7) que, implícitamente, hace referencia a más de un elemento, por ejemplo, *Herron* y *Green* en el texto 3a (or. 13), y lo mismo puede decirse de los demostrativos o de los pronombres plurales. Al tener los elementos a los que sustituyen o hacen referencia el mismo estatus en la jerarquía, no podemos

elegir entre uno u otro, ya que resultarían conexiones parciales del tipo *cor: pair-Herron*, o bien, *cor: pair-Green*. Si, para solucionarlo, anotamos, en cambio, una única conexión entre los tres elementos, i.e. *cor: pair-Herron, Green* (3a, ors. 8-13), estaríamos dando cuenta de la referencia real sin incurrir en un recuento doble. De acuerdo con la misma premisa, el procedimiento podría ser utilizado también a la inversa, es decir, dos elementos podrían establecer una única conexión con un tercero. Ambos procesos se pueden ver a continuación:

- (5) The court was told that Helen Calderwood [...] heard the manageress Patima Herron and the assistant Lance Green bemoaning the loss of a 12 per cent bonus [...] (texto 3b, or.7)
The pair had been geed up by a regional supervisor earlier that day [...] (texto 3b, or.8)
Ms Herron and Mr Green are charged with making a hoax bomb call but have not entered a plea (texto 3b, or.12)

conexiones sugeridas:

- cor: Herron, Green - pair* (texto 3b, ors.7-8)
cor: pair- Ms Herron and Mr Green (texto 3b, ors. 8-12)

4.3. Conexiones de sustitución

Las conexiones de sustitución, además de por su relación con la correferencia, y por su ámbito de influencia, conllevan sus propios problemas analíticos, relacionados, sobre todo con los pronombres posesivos. Según se deduce del tratamiento de los determinantes demostrativos o del hecho de que el autor restringe los sustitutos del sistema pronominal fundamentalmente a los de tercera persona con función sujeto, este otro tipo de sustitutos no mencionado explícitamente (aunque sí utilizado por el autor en su análisis, Hoey 1991: 87) no debería ser tenido en cuenta para establecer conexiones ya que supone un recuento doble en el que el determinante y el nombre al que acompaña reiteran un mismo elemento léxico.

Este tratamiento se aplicaría, por ejemplo, al un grupo o sintagma como *those gays* (texto 5a or. 7), en el que sólo *gays* podría establecer relaciones léxicas (por ejemplo, una paráfrasis con *homosexuals*, or. 10) al tener el demostrativo que funciona como determinante la misma referencia que el núcleo. Nótese, sin embargo, que no es la función de determinante la que impide la conexión, ya que, en casos como *Carey's actions* (5b, or. 9), cada uno de los elementos, por su naturaleza léxica y por su distinto referente, y aun a pesar de que la función del genitivo es idéntica a la de *those*, puede establecer sus propias repeticiones. Si, con estos antecedentes y ninguna indicación específica del autor, nos enfrentamos al análisis de ejemplos

del mismo texto como *his actions* (or. 10), referido al mismo hecho (esto es, *Carey's actions*) se verá claramente la razón por la cual sería posible incluir los pronombres posesivos en particular, y los determinantes en general, como sustitutos de pleno derecho. La equivalencia sintáctica, y, en esta ocasión particular, semántica, entre el pronombre posesivo y el genitivo y la distinta referencia entre ambos y el núcleo parecen justificar que se traten de igual modo, es decir, como elementos que establecen sus conexiones independientemente.

4.4. Elipsis

El tratamiento de la elipsis también merecería ser revisado. El texto 5a nos proporcionó varias muestras, al ser uno de sus rasgos más sobresalientes la aparición de varias repeticiones elípticas. El primer caso se detectó en la oración diecisiete, donde la palabra *couples* hace referencia a parejas del mismo sexo, tal y como aparece expresamente en el antetítulo, sobreentendiéndose, entonces, los modificadores *same-sex* o *homosexual*:

54

- (6) Homophobia claim after archbishop acts to halt publication of prayers for *same-sex couples* and Aids victims (texto 5a, or. antetítulo)

A section of the book proposes [se sobreentiende “*same sex*” o “*homosexual*”] *couples* imitate the ancient Israelites by marking their door with “a sign of blessedness” (texto 5a, or. 17)

Establecimos, así, por tanto, la conexión elíptica de estas oraciones, pero, inmediatamente, surgió la duda de si la palabra omitida podría relacionarse con otras del mismo campo, también presentes en la noticia como *homophobia*, *gay* o *lesbian*. A falta de indicaciones específicas en la obra de Hoey (1991), decidimos que los casos de elipsis se restringieran a las repeticiones exactas de la misma palabra, evitando las paráfrasis o cualquier otro tipo de conexión, y, de la misma forma, no consideramos que la elipsis pudiera establecer sus propias conexiones con elementos posteriores. De esta manera creemos respetar la idea del autor, quien considera la elipsis como el último recurso a tener en cuenta, precisamente por dar prioridad a la repetición de elementos léxicos “tangibles”, por así decirlo. En principio, Hoey (1991: 74) parece bastante tajante respecto al tratamiento de la elipsis:

ellipsis, the absence of some required stretch of language that has to be supplied by the listener or reader to make sense of the sentence encountered. It seems only ever to function as a type of short-term repetition and is only countenanced in our analysis as a means of creating repetition in those cases where a sentence is grammatically incomplete unless something is supplied from earlier in the text.

De este párrafo, sin embargo, se desprende ya una contradicción respecto al propio concepto que, aunque se define inicialmente en términos semánticos (“make sense”), ha de tenerse en cuenta, a efectos de repeticiones, de acuerdo con criterios sintácticos que invalidarían, por ejemplo, la conexión del texto 5a que acabamos de ejemplificar (6). Aún más sorprendente es que, a renglón seguido, el autor (1991: 74) se desdiga e introduzca un tercer criterio, en este caso textual, afirmando:

ellipsis has to be accounted for, whether or not it gives rise to grammatical incompleteness, when we consider the way in which we may produce sub-texts from the non-narrative texts we analyse.

Para comprender esta cita debemos recordar que una de las hipótesis de Hoey (validada también por nuestro corpus) es que la unión de cualesquiera oraciones relacionadas por un número significativo de conexiones, generalmente tres o más, deparará un sub-texto coherente (véanse los ejemplos (4) y (5) anteriores) merced a que la vinculación léxica entre oraciones sirve para organizar la información del texto (por ejemplo, en patrones de problema-solución, Hoey 1983, 2001) y tiene, por tanto, valor textual. La afirmación de Hoey reproducida arriba implica que encontraremos casos en que será necesario recurrir a la elipsis para justificar la vinculación de las oraciones y que ésta puede tenerse en cuenta cuando sea necesario desde el punto de vista de la organización textual. Esta posibilidad nos daría pie, no sólo a anotar efectivamente los casos anteriores, sino también a considerar otros descubiertos en la fase organización del texto posterior al análisis de las conexiones.

Aunque consideramos que la postura más sensata, en la aplicación de un método destinado a identificar las conexiones léxicas, es utilizar la elipsis de forma restrictiva, creemos, también, que limitarla con criterios sintácticos, como se deduce de las palabras del autor, es poco práctico. De esta manera, parece más lógico anotar las conexiones meridianamente claras desde un punto de vista semántico, como en el ejemplo (6), o aquellas que se demuestren necesarias para la organización del texto, como da a entender el autor. Esta última circunstancia es la que motivó la inclusión de otros casos de elipsis detectados en el mismo artículo (5a), así como en otros (5c, 7a, 7b, etc.). Tal tratamiento, según se deduce del pobre porcentaje de las repeticiones elípticas en el cómputo general que acabamos de ver en la Tabla 1, no parece afectar al resultado general del análisis de los tipos de repetición.

5. Las conexiones catafóricas

Como ya adelantamos, los tres tipos de conexión discutidos, correferencia, sustitución y elipsis, probablemente porque Hoey les presta menos atención, son los que

parecen requerir más y más frecuente adaptación. El resto de los instrumentos formales, y no queremos dejar de resaltarlos, han sido fácilmente aplicables, aunque no quisiéramos, sin embargo, acabar sin matizar algunos aspectos que consideramos inconclusos o mejorables.

Uno de ellos es el tratamiento de las conexiones catafóricas. El análisis de Hoey restringe la detección de las conexiones estrictamente a las relaciones anafóricas. A pesar de las apariencias, ignorar las relaciones catafóricas no supone un gran problema para el método debido a que la mayoría de ellas suelen darse de forma introracional o, cuando menos, dentro de una misma oración ortográfica, no contribuyendo ejemplos como el siguiente, por tanto, y de acuerdo con las premisas de Hoey, a la cohesión del texto:

(7) Yesterday I went to the movies with *him* - with *John* I mean. (Maes 1987: 134)

Evitando los supuestos mencionados, el método de Hoey no se ve, en apariencia, afectado por ignorar las conexiones catafóricas, lo cual le conviene, además, para la implementación informática del mismo. Sin embargo, quedarían ciertos cabos sueltos, como el que presentaría un ejemplo de este tipo:

(8) I would have never believed *it*. *They have accepted the whole scheme*. (Halliday y Hasan 1976: 56)

en el que el pronombre *it* sustituye catafóricamente a la oración (ortográfica y sintácticamente definida) que le sucede; o en casos como el que hemos hallado en el análisis del texto 4a, un elemento como *handling* (4a, or.1), correferente catafóricamente con ocho oraciones posteriores, que no reproduciremos por falta de espacio. Este sustantivo con referencia textual condensa el contenido de las oraciones siguientes, hecho reforzado por encontrarse, además, en la entradilla, cuya misión es precisamente la de resumir el artículo. De este modo, no se nos escapa su similitud con los sustantivos o nombres anafóricos que conocemos y que parece respaldada por las palabras de Francis (1986: 103) en esta cita: “A-nouns [...] function as structural signposts which the writer uses to enable the reader to grasp the outlines of his/her argument more efficiently” y ratificada (1986: 104) por esta otra: “the same lexical items which can function as A-nouns can also function as *cataphoric* structural signals within a discourse”.

No parece, por tanto, haber razones para descartar ni estos elementos léxicos cohesivos ni otros textuales claramente catafóricos. Siendo consistentes, su tratamiento debería ser el mismo que el de casos similares (los mencionados sustantivos anafóricos u otros elementos con referencia textual), o sea, como conexiones correferenciales cuyo ámbito, por supuesto, puede ser superior al de la oración más cercana.

6. La implementación informática

Sugeríamos en el punto anterior que el tratamiento que de la catáfora hace Hoey podía estar influido por el aspecto informático del método. Tras haber encontrado a lo largo de todo el análisis un gran número de conexiones que, a nuestro entender, dificultarían en gran manera la aplicación informática de este método, confirmamos nuestra impresión inicial de que muchas de las restricciones impuestas por el autor para la clasificación de conexiones parecen tener una razón práctica, el análisis informático, más que un fundamento estrictamente teórico.

El autor, consciente de ello, menciona el problema que implica detectar los casos de conexión correferencial basados en el conocimiento contextual o del mundo (del tipo *BBC-The Corporation* texto 9b, ors. 2-4), y propone (1991: 70), como solución, dotar al programa informático de conocimientos enciclopédicos. Sin embargo, tales conocimientos enciclopédicos, aun actualizados, no garantizarían la detección de los casos en los que la correferencia depende del contexto generado por el propio evento comunicativo, más concretamente, del conocimiento textual (Dik 1997b: 412). Es este conocimiento textual el que nos capacita para reconocer conexiones del tipo *cor: chiefs-watchdogs* (texto 2c, ors. título-5), u otras como *cor: weapons-things* (texto 1a, ors. 1-5), que involucran, además, a los nombres o sustantivos genéricos de Halliday y Hasan (1976) o a la hiponimia o meronimia relacional de Martin (1992).

Estas conexiones que dependen del conocimiento textual no se limitan a la correferencia, sino que pueden concernir a otros tipos de repetición, véase, si no, la paráfrasis *psp: ban-pull the plug* (texto 2c, ors. 1-5), que, además, nos da pie a mencionar aquellas conexiones donde se hace un uso metafórico o con doble sentido del lenguaje.

A pesar de la evidencia para cualquier analista humano, la mayoría de estos usos tuvieron que ser descartados en el recuento de repeticiones al no encajar en el contexto colocalacional, requisito indispensable para Hoey. Al ser los juegos de palabras un rasgo típico del lenguaje de los titulares de las noticias en inglés, el llamado *headlines*, muchas de las conexiones metafóricas se perdieron en esta parte de la noticia, hecho que afectó en algunos casos, a la segunda fase de análisis, la de creación de matrices y organización del texto. Como muestra de ello podemos mencionar la relación no contabilizada entre *fire first salvos* (texto 9a, or. título) y *break out* (texto 9a, or. 1) o entre *battle* (texto 9b, or. título) y *weapon* (texto 9b, or. 2), donde, además, intervendrían relaciones semánticas de composición o, al menos, de campo semántico. La utilización intencionada de la homonimia o de la polisemia, conocida como *puns* (Leech 1969), fue, sin embargo, más fácil de integrar debido a que al menos una de las acepciones de los elementos conectados no tiene un uso figurado. Para ilustrar este punto podemos utilizar el título del ar-

título 8a, *Woman sinks pool opponents*, donde el uso homonímico de *pool* no impidió su conexión con repeticiones no ambiguas de la misma palabra, y donde, incluso, *sinks* pudo ser conectado por medio de una paráfrasis a *beaten* (or. 3):

- (9) *Woman sinks pool opponents* (texto 8a, or. título)
 Ms Thompson, of Rucorn, Chesire, has won 200 women's titles and beaten eight of the game's 40 male professionals including Britain's number two, Greg Farran (texto 8a, or.3).

Conexiones:

psp: sinks-beaten (8a, ors. tít-3)

psp: pool-game's (8a, ors. tít-3)

Estos ejemplos no deben, sin embargo, llevarnos a engaño, pues buena parte de los usos no explícitos del lenguaje fueron imposibles de relacionar y, para dar una última muestra de ello, podemos comparar el uso metonímico que se hace de la palabra *cue* en los titulares del texto 8c y 8b:

- (10) *Cue ace Sue wins battle to turn pro.* (texto 8b, or. tít.)
 Susan wins the right to *cue* up for pool room. (texto 8c, or. tít.)

Aunque en ambos casos una de las acepciones del vocablo es “palo de billar”, sólo fue posible establecer conexiones en el texto 8c, utilizando *cue* como paráfrasis de *pool* (nótese, a su vez, el tratamiento de la meronimia); en el segundo caso, sin embargo, las conexiones de *cue* quedaron descartadas, tanto en su sentido literal por aparecer la palabra *pool* en el título, como en su uso figurado como “entrar” por no encajar en el contexto colocacional de posteriores repeticiones.

Igualmente achacables al análisis informático creemos que son las restricciones del método respecto a alguno de sus instrumentos, como es el caso del triángulo de conexiones. Este mecanismo parece basarse en la afinidad del contexto y en la idea de que la cohesión léxica se sustenta en la interrelación de los elementos léxicos. Sin embargo, en la práctica, como hemos explicado, el autor restringe su uso a los casos en que un elemento parafrasea de forma simple y repite complejamente a otros dos, un tratamiento que creemos desaprovecha muchas de sus posibilidades para detectar conexiones que ahora no pueden ser tenidas en cuenta. Por esta razón, provisionalmente, en el texto 1a y en 4a lo utilizamos para establecer un tipo de paráfrasis compleja derivada de dos paráfrasis simples.

Ciertos elementos relacionados semánticamente no pueden ser considerados paráfrasis porque el método de Hoey les exige que encajen en el mismo contexto colocacional. De esta forma, la relación entre, pongamos por caso, *attacked* y *accusers* (4a, ors. 1 y 7), a pesar de la evidencia contextual, no puede ser incluida en

ninguna de las categorías descritas por él. Cada uno de estos elementos, sin embargo, es individualmente una paráfrasis simple de *denounced* (texto 4a, or. 3) y, por lo tanto, pueden aparecer en su lugar:

- (11) An all-party committee of MPs has *attacked* the Bank of England's handling of the BCCI collapse [...] (texto 4a, or. 1)
This [report] *denounced* the Bank for taking so long to close BCCI [...] (texto 4a, or. 3)
A minority report *accuses* the Bank's governor of grossest negligence [...] (texto 4a, or. 7)

Es decir, que esta última palabra, *denounced*, es el nexa y la prueba tangible de que existe una relación entre las dos primeras. Al no ser tal relación, según sus criterios, una paráfrasis simple, podría, en cambio, ser una paráfrasis compleja (nótese como la relación entre *attacked* y *accuses* es mucho más clara que la de otros elementos que son paráfrasis complejas indiscutibles, por ejemplo, el sustantivo *production* y el verbo *making*, texto 7b, ors. 3-13). De esta manera, además del propio triángulo de conexiones, se modificaría, de paso, el tratamiento de la paráfrasis compleja.

Como último argumento respecto a los efectos de la implementación informática del modelo queremos hacer notar que los problemas no se producen sólo por el descarte de ciertas conexiones, sino también por la inclusión indiscriminada de otras. Al haber utilizado el método en las noticias de la prensa, hemos encontrado elementos léxicos típicos de este medio como los verbos de habla *say* o *tell*, que, en principio, y de la misma manera que cualquier verbo léxico, deberían incluirse en el recuento de conexiones. Sin embargo, al ser tan frecuente su uso, en algunas ocasiones, la repetición idéntica del mismo verbo, aunque referido a un asunto diferente y expresado por un agente distinto, estableció una tercera conexión, es decir un vínculo, entre dos oraciones, que, quizá, si se obviase tal verbo, no hubiera tenido lugar. Al haber tenido que recurrir en varios textos a aumentar el umbral de conexiones a cuatro (i.e. 8b, ors. 6-8), siguiendo la posibilidad sugerida por el autor, este problema no fue excesivamente importante, aunque sí nos hizo plantearnos la posibilidad de que esta clase de verbos no se incluyera automáticamente en el recuento de conexiones a menos que el sujeto o el agente fuera el mismo, para, de esta manera, evitar una repetición que parece casual. De aquí se deriva, entonces, nuestra referencia al análisis informático, ya que este debería ser capaz de reconocer las funciones sintácticas o semánticas, además de las simples repeticiones. Este procedimiento, sin embargo, no fue seguido en el análisis debido a que no lo consideramos una simple modificación; si bien en los casos que acabamos de ver se sugieren variaciones valiéndonos de las propias afirmaciones del autor o, al menos, intentando seguir sus principios generales, esta otra representaría establecer un criterio completamente ajeno al análisis de Hoey y, por ello, nos limitamos aquí a señalarla.

7. Conclusión

En este artículo hemos tratado de realizar un repaso de algunos de los elementos formales del análisis de Michael Hoey que se han revelado como más problemáticos tras la aplicación del mismo a un corpus de textos periodísticos. A pesar de la perfecta replicabilidad del mismo, hemos pretendido hacer notar la existencia de algunos aspectos que merecerían cierta revisión. Destacamos entre ellos la definición ortográfica de los elementos léxicos, el tratamiento de las conexiones de sustitución y de correferencia (con especial énfasis en los elementos textuales), la detección de las paráfrasis, la utilización del triángulo de conexiones, o bien, las dificultades de integrar los usos figurados del lenguaje. Las soluciones y variaciones que se sugieren a estos problemas intentan valerse, más que de una discusión teórica extensa acerca de cada uno de los temas abordados, de las propias afirmaciones del autor y de los principios generales de su análisis para que, de esta forma, resulten de utilidad para aquellos que desde un punto de vista, bien exclusivamente discursivo, bien computacional, traten de explorar sus muchas posibilidades de aplicación. Las apuntamos, entonces, como sugerencias auspiciadas por el ánimo que, según se desprende de sus propias palabras, parece mover a Hoey (1991: 77): “It is hoped that readers will feel to be able to test the claims made for themselves”.

60

Notas

1. En el apéndice de este trabajo puede consultarse la enumeración de los textos integrantes del corpus y la codificación con la que en adelante se hará referencia a ellos.

2. De la misma manera que Halliday y Hasan (1976) o Hasan (1984, 1989), Hoey, aunque tratándolas como categorías menores, se ve obligado a incluir la correferencia, la sustitución y la elipsis, en su recuento de conexiones cohesivas.

3. El tratamiento de estos tipos de repetición como categorías menores se refleja en el hecho de que Hoey no proporcione una definición explícita de los mismos. En el caso de la correferencia se limita a afirmar lo siguiente: “if, and only if, two items are interpreted as having identical referents they will be treated as repetitions – even though they are either unrelated as lexical items (co-refer-

ence, as in Augustus and the Emperor)” (Hoey 1991: 70). Respecto a la sustitución, ésta se incluye en un apartado final, denominado significativamente “other ways of repeating” (Hoey 1991: 70-74), que comprende algunos pronombres personales y posesivos, así como los demostrativos.

4. Dik (1997a: 131) delimita los elementos que pueden tener función referencial arguyendo lo siguiente: “I restrict the function of reference to terms which act as arguments or satellites within the clause. In using a term S either sets up a referent or helps A to retrieve a referent in order to predicate something about it”. En la misma línea, Keizer (1991: 21) afirma: “A referring expression is an expression by means of which in a given predication, a speaker pinpoints some entity (or entities) about which s/he wishes to predicate something”.

Obras citadas

- BENBRAHIM, Mohamed y Khurshid AHMAD. 1994. "Computer-aided lexical cohesion and text abridgement". *Technical report. Knowledge Processing*, 18. University of Surrey. Segunda edición revisada por Kees Hengeveld. Berlín: Mouton de Gruyter.
- BERBER SARDINHA, Tony. 1997. *Automatic Identification of Segments in Written Texts*. Tesis inédita. English Department University of Surrey. Descargada en Mayo de 2004 desde <http://www2.lael.pucsp.br/~tony/theses.html>
- 2001. "Lexical Segments in Text". En Scott, Mike y Geoff Thompson. (eds.). *Patterns of Text. In Honour of Michael Hoey*. Amsterdam: John Benjamins: 213-238.
- DÍAZ GALÁN, Ana. 2002. "La cohesión léxica como indicador de las categorías esquemáticas del artículo de prensa". *Revista Canaria de Estudios Ingleses*, 44: 245-265.
- 2004a. *Estudio de la Cohesión Léxica en la Prensa Británica*. La Laguna: Servicio de publicaciones de la Universidad de La Laguna.
- 2004b. "Una caracterización genérica de las noticias de la prensa británica por medio de la cohesión léxica". *Odisea*, 5: 45-60.
- DÍAZ GALÁN, Ana y María del Carmen FUMERO PÉREZ. 1997. "Making reading easier". En Bolaños, Lydia Esther, Alexander Cárdenes y M. García. (eds.). 1997. *Actas de las I Jornadas Canarias de Lenguas Aplicadas a la Ciencia y la Tecnología*. Las Palmas de Gran Canaria: Universidad de las Palmas de Gran Canaria: 102-107.
- 2003. "The Problem-Solution Pattern: A Tool for the Teaching of English?". *BELLS* 13. Versión electrónica <http://www.publicacions.ub.es/revistes/bells13/>
- DIK, Simon. 1997a. *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Segunda edición revisada por Kees Hengeveld. Berlín: Mouton de Gruyter.
- 1997b. *The Theory of Functional Grammar. Part II: Complex and Derived Constructions*.
- FRANCIS, Gill. 1986. *Anaphoric Nouns*. Discourse Analysis Monographs No 11. Birmingham: English Language Research, University of Birmingham.
- 1994. "Labelling discourse: an aspect of nominal-group lexical cohesion". En Coulthard, Malcolm. (ed.). *Advances in Written Text Analysis*. Londres: Routledge: 83-101.
- HALLIDAY, Michael. A. K. 1994. *An Introduction to Functional Grammar*. Londres: Arnold.
- HALLIDAY, Michael A. K. y Ruqaiya HASAN. 1976. *Cohesion in English*. Londres: Longman.
- 1989. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford: Oxford U. P.
- HASAN, Ruqaiya. 1984. "Coherence and Cohesive Harmony". En Flood, James. (ed.). *Understanding Reading Comprehension*. Delaware: Delaware International Reading Association: 181-219.
- 1989. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective. Part B*. Oxford: Oxford U. P.
- HENGEVELD, Kees. 1997 "Cohesion in Functional Grammar". En Conolly, John H., Roel Vismans, Christopher Butler y Richard Gatward. (eds.). 1997. *Discourse and Pragmatics in Functional Grammar*. Berlín: Mouton de Gruyter: 1-16.
- 2004. "The Architecture of a Functional Discourse Grammar". En Mackenzie, Lachlan y María Ángeles Gómez González. (eds.). *A New Architecture for Functional Grammar*. Berlín: Mouton de Gruyter: 1-21.
- HOEY, Michael. 1983. *On the Surface of Discourse*. Londres: Allen and Unwin.

- 1991. *Patterns of Lexis in Text*. Oxford: Oxford U. P.
- 2001. *Textual Interaction. An Introduction to Written Discourse Analysis*. Londres: Routledge.
- KEIZER, M. Evelien. 1991. "Referring Expressions in Functional Grammar: how to define reference and referring expressions". *Working Papers in Functional Grammar*, 43.
- LEECH, Geoffrey. 1969. *A Linguistic Guide to English Poetry*. Londres: Longman.
- MAES, Alfons. 1987. "The Pragmatic Value of Cataphoric Relations." En Nuyts, Jan y Georges De Schutter. (eds.). *Getting one's Words into Line. On Word Order and Functional Grammar*. Dordrecht: Foris: 131-146.
- MARTIN, James R. 1992. *English Text: System and Structure*. Amsterdam: John Benjamins.
- 2001. "Cohesion and Texture". En Schiffrin, Deborah S, Deborah Tannen y Heidi Hamilton. (eds.). *The Handbook of Discourse Analysis*. Oxford: Blackwell: 35-53.
- RENOUF, Antoinette y Alex COLLIER. 1995. *A system of Automatic abridgement*. Research and Development Unit for English Studies: University of Liverpool.
- SARACENI, Mario. 2002. "Review of *Patterns of Text: In Honour of Michael Hoey*". *Linguist List* 13.230. Descargado en Mayo de 2004 desde <http://www.linguistlist.org/issues/13/13-230.html>
- SCOTT, Mike y Geoff THOMPSON. (eds.). 2001. *Patterns of Text. In Honour of Michael Hoey*. Amsterdam: John Benjamins.
- VAN DIJK, Teun. 1980. *News as Discourse*. Hillsdale: Lawrence Erlbaum.
- 1986. "News schemata". En Cooper, Charles R. y Sidney Greenbaum. (eds.). *Studying Writing. Linguistic Approaches*. Londres: Sage: 155-185.
- WINTER, E. O. 1979. "Replacement as a Fundamental Function of the Sentence in Context". *Forum Linguisticum*, 4 (2): 95-133.

Apéndice

Relación y codificación de los textos integrantes del corpus

Las treinta noticias integrantes del corpus se han codificado con una letra, *a*, *b* o *c*, según el grupo socio-económico del diario británico en el que fueron publicadas el 12 de marzo de 1992 (*a*: *up-*, *b*: *mid-* o *c*: *down market*) y con un número que hace referencia a su contenido. De esta manera, por ejemplo, los textos 1a, 1b y 1c corresponderían a la misma noticia publicada en tres tipos de diario diferente. Enumeramos a continuación sus títulos y el diario en que aparecieron y, para una consulta más completa, remitimos al lector interesado al volumen Díaz Galán 2004a.

Algunas consideraciones sobre la propuesta de análisis de ...

- 1a:** "West threatens to bomb Iraqi nuclear plants for defying UN" *The Times*
- 1b:** "Countdown to war as UN runs out of patience" *Daily Mail*
- 1c:** "The Saddam factor" *Daily Star*
- 2a:** "Slap in the face for a Tango ad" *Daily Telegraph*
- 2b:** "TV ad scrapped after school slapping craze" *Daily Express*
- 2c:** "Tango-ing! Telly Chiefs slap ban on Mr. Orange" *The Sun*
- 3a:** "McDonald=s woman in hoax bomb call" *The Independent*
- 3b:** "Burger wars girl in hoax bomb call" *Today*
- 3c:** "Doing a bomb in burgers" *Daily Star*
- 4a:** "MPs admonish bank over BCCI" *The Times*
- 4b:** "MPs admonish bank over massive BCCI fraud" *Daily Express*
- 4c:** "Major BCCI rap" *Daily Mirror*
- 5a:** "Carey condemned as 'ayatollah' over gay book" *The Guardian*
- 5b:** "Gays attack 'ayatollah' archbishop" *Daily Express*
- 5c:** "Gay prayer book fury" *Daily Star*
- 6a:** "Nurse faces new charges of attempted murder" *Daily Telegraph*
- 6b:** "Nurse accused of hospital killings faces new charges" *Daily Mail*
- 6c:** "Nurse on new charges" *Daily Mirror*
- 7a:** "Country cottage was drugs center" *The Guardian*
- 7b:** "Cops turn dustmen to crack LSD gang" *Today*
- 7c:** "60m LSD factory at Lawless Cottage" *Daily Mirror*
- 8a:** "Woman sinks pool opponents" *The Guardian*
- 8b:** "Susan wins the right to cue up for pool room" *Today*
- 8c:** "Cue ace Sue wins battle to turn pro" *The Sun*
- 9a:** "TV rivals fire first salvoes in battle for election" *The Guardian*
- 9b:** "Battle of TV Experts" *Daily Express*
- 9c:** "Major ducks live TV head-to-head" *Daily Mirror*
- 10a:** "Bush backing for 'superb' Major" *The Guardian*
- 10b:** "A pat on the back from Bush" *Daily Express*
- 10c:** "Major is a superb PM says Bush" *Daily Star*