

CAPÍTULO 3

DESCRIPCIÓN DEL *CORPUS* Y METODOLOGÍA APLICADA

Presentamos a continuación el *corpus* estudiado, describimos el método propuesto por el lingüista Michael Hoey (1991a) para el análisis de textos no narrativos: tipos de repetición, construcción de matrices de repetición y detallamos el procedimiento efectuado para la confirmación o refutación de las hipótesis planteadas en nuestra investigación.

3. 1. Descripción del *corpus*.

El *corpus* que conforma este trabajo consta de veinte textos auténticos, pertenecientes todos ellos al área de Ciencias Químicas. Diez son introducciones de artículos de investigación extraídos de revistas especializadas en el campo de Química Analítica. Los otros diez consisten en textos obtenidos de revistas educativas dirigidas al profesorado y alumnado de Ciencias Químicas. Todas las revistas seleccionadas se encuentran a disposición de los alumnos, bien en la biblioteca de la Facultad de Ciencias, o bien en el Departamento de Química Analítica.

Las siguientes tablas muestran los títulos de los textos analizados, las revistas a las que pertenecen, y el número de oraciones y palabras que cada texto contiene:

Introducciones de los Artículos de Investigación. (IAI)	Revistas	Número de oraciones	Número de palabras
1. Bioconversion of solid food wastes to ethanol.	The Analyst. (1998), vol. 123: 497-502.	18	437
2. Speciation as an analytical aid in trace element research in infant nutrition.	The Analyst. (1998), vol 123: 821-826	22	606

3. Analysis of carbonaceous aerosols: interlaboratory comparison.	The Analyst. (1998), vol. 123: 851 – 857.	27	664
4. High-precision conductometric detector for the measurement of atmospheric carbon dioxide.	Analytical Chemistry. (1998), vol. 70: 4678-4682.	21	533
5. Refinement of the borohydride reduction method for trace analysis of dissolved and particulate dimethyl sulfoxide in marine water samples.	Analytical Chemistry. (1998), vol. 70: 4864-4867.	10	280
6. Determination of cyanide in whole blood by capillary gas chromatography with cryogenic oven trapping.	Analytical Chemistry. (1998), vol. 70: 4873-4876.	14	311
7. RP-HPLC binding domains of proteins.	Analytical Chemistry. (1998), vol. 70: 5010-5018.	14	494
8. Nanoliter chemistry combined with mass spectrometry for peptide mapping of proteins from single mammalian cell lysates.	Analytical Chemistry. (1998), vol. 70: 5344-5347.	10	362
9. The determination of food colours by HPLC with on-line dialysis for sample preparation.	Food Chemistry. (1992), vol. 43: 137-140.	18	534
10. Analysis of serotonin in whole-blood samples – A novel fully automated method.	LC-GC INT. (1995), vol. 8: 104 -107.	17	334
TOTAL		171	4555

Artículos ‘Académicos Informales’. (AAI)	Revistas	Número de oraciones	Número de palabras
1. Is it real gold?	Journal of Chemical Education. (1999), vol. 76: 198 – 199.	46	1327
2. Why gold and copper are colored but silver is not.	Journal of Chemical Education. (1999), vol. 76: 200	23	549
3. Both nylon and PET fibers burn continuously under atmospheric conditions.	Journal of Chemical Education. (1999), vol. 76: 793.	20	488
4. A chromatographic parable.	Journal of Chemical Education. (2000), vol. 77: 176	29	604
5. High flying polymer.	Education in Chemistry. (1999), vol. 36 / 58: 1	15	391
6. Flash of inspiration wins Nobel prize for chemistry.	Education in Chemistry. (2000), vol. 37 / 1: 8.	30	888
7. Pressure to change solvents.	Education in Chemistry. (2000), vol. 37 / 4: 2-3	49	1128
8. A healthy spread.	Education in Chemistry. (2000), 37 / 6: 2-3	49	1189
9. Apatite for destruction.	Education in Chemistry. (2000), vol. 37 / 66: 1	17	406

10. Hair-raising ideas.	Chemistry in Britain. (2000), vol. 36 / 11: 19	17	497
TOTAL		295	7467

Para la ejemplificación del modelo de análisis efectuado en esta investigación se ha utilizado la introducción de un artículo de revisión denominado *Sol-gel encapsulation methods for biosensors* de la revista *Analytical Chemistry*. (1994), vol. 66: 1120-1127.

3. 2. Modelo de análisis.

El modelo de cohesión textual utilizado en nuestros textos para su análisis es el que el lingüista Michael Hoey propone en su libro *Patterns of Lexis in Text* (1991a) para textos no narrativos¹. Fundamentándose en la construcción de matrices de reiteración léxica, Hoey (1991a) demuestra que la presencia de redes de interrelaciones cohesivas evidencia la organización de la información semántica entre oraciones próximas y no próximas en un texto. Las oraciones se consideran, de este modo, en ‘paquetes’ interrelacionados de mensajes semánticos conectados por repeticiones múltiples y, el léxico, en organizador del significado del texto.

Utilizaremos el texto *Both nylon and PET fibers burn continuously under atmospheric conditions*. (1999)² para ejemplificar el método propuesto por Hoey (1991a) y detallar el procedimiento para la extracción de resúmenes.

3. 2. 1. Tipos de repetición.

Hoey (1991a: 71) distingue entre las unidades léxicas que “do not depend on each other for their meaning” y las unidades textuales que “entirely depend on other items for their interpretation ...their function is to stand in, or substitute, for lexical

¹ Hoey (1991c, 1994b, 1997b, 2000a) también desarrolla un modelo de descripción de textos narrativos.

² Cuando algún tipo de repetición no aparezca en este texto, utilizaremos otros textos analizados para ejemplificarlos.

items”. Trataremos en primer lugar los diferentes tipos de repetición establecida mediante las unidades léxicas.

3. 2. 1. 1. Repetición simple: Dos unidades léxicas que son idénticas [*bear – bear*] o dos unidades similares cuya diferencia sea “explicable solely in terms of different choices from a grammatical paradigm” [*bear – bears*] (Hoey, (1991a: 58).

No trataremos como repetición simple la reiteración de unidades gramaticales como determinantes, preposiciones, auxiliares, negativas, coordinadores, subordinadores, marcadores del discurso³ y submodificadores. No quiere decir que estas unidades no puedan contribuir a establecer relaciones entre oraciones. Winter (1974) muestra la contribución que estos elementos gramaticales pueden desempeñar en los modelos de repetición. Pero nuestra decisión es más bien analítica. Nuestro objetivo es la repetición léxica y si ésta contribuye a establecer interrelaciones cohesivas entre las oraciones de un texto. Por lo tanto, tener en cuenta también la repetición de unidades gramaticales nos podría llevar a confundir nuestro objetivo (Hoey, 1991a: 53).

La repetición simple⁴ no está exenta de problemas, ya que el significado de una palabra no necesariamente se mantiene constante cuando se repite. Hoey (1991a: 54), sin embargo, considera que no se requiere una identidad absoluta de significado para establecer una relación de repetición simple entre dos unidades léxicas.

Por ejemplo, en las dos oraciones siguientes, las unidades léxicas **nylon**, **PET** y **fibers**, enmarcadas y unidas por líneas anaranjadas, están relacionadas mediante la repetición simple:

1. We would like to present two series of photographs showing the characteristic burning behaviors of a **nylon** fiber and a polyethyleneterephthalate (**PET**) fiber, in order to help people safely handle these **fibers** in their everyday lives.

2. In many textbooks, especially on textiles, **nylon** and **PET fibers** are classified as flammable but self-extinguishing.

³ Para la función cohesiva de los marcadores del discurso ver Ariel (1998).

⁴ Véase Rimmon-Kenan (1980), McCarthy (1988) en Hoey (1991).

Sin embargo, en las siguientes oraciones del texto *Speciation as an analytical aid in trace element research in infant nutrition* (1998), los elementos léxicos **separation** en las oraciones 2 y 22 no establecen una repetición simple, ya que en el contexto en el que aparecen estas dos unidades léxicas no tiene nada en común la separación o nacimiento del niño con el proceso de separación de las proteínas.⁵ El contexto será básico y decisivo para determinar cualquier tipo de reiteración léxica (Hoey, 1991a: 37, 58-60, 68).

1. After **separation** from the mother, the newborn has to develop its own functions and regulatory systems, including respiration, digestion and immune defenses.

22. We therefore carried out speciation studies to determine the binding form of trace elements in these fluids, combining methods for protein **separation** with methods for trace element determination in the eluted fractions.

3. 2. 1. 2. Repetición compleja: Dos unidades similares que comparten el mismo morfema, o dos unidades idénticas con diferente función gramatical. [*human* (nombre) – *human* (adjetivo), *drug* – *drugging*] (Hoey, 1991a: 55).

Dos unidades léxicas establecen una relación de repetición compleja cuando pueden ser parafraseadas en el contexto del texto en el que aparecen de tal manera que la paráfrasis de una incluya la otra. Hay que tener en cuenta que este criterio no significa una paráfrasis exacta, sino una aproximación en el contexto (Hoey, 1991a: 55). Por ejemplo, los elementos léxicos **burning** y **burned** enmarcados y unidos por una línea verde constituyen una repetición compleja:

⁵ Hoey (1991: 37) ejemplifica cuándo una unidad léxica no se considera repetición debido al contexto. Asimismo, consideramos determinante la función del especialista en el contenido para determinar si una unidad léxica establece relación de repetición.

1. We would like to present two series of photographs showing the characteristic **burning** behaviors of a nylon fiber and a polyethyleneterephthalate (PET) fiber, in order to help people safely handle these fibers in their everyday lives.

10. Once a part of the terminal end was ignited, it **burned** continuously, as shown in the series of photographs in this paper, in contrast to the descriptions in textbooks stating that it "burns slowly but if the sample is removed from the flame it self-extinguishes".

3. 2. 1. 3. Paráfrasis simple: Dos unidades léxicas de la misma categoría gramatical. La paráfrasis simple ocurre “whenever a lexical item may substitute for another without loss or gain in specificity and with no discernible change of meaning [*sedated – tranquilized*]” (Hoey, 1991a: 62); paráfrasis que equivale a la sinonimia de Halliday y Hasan (1976).

La paráfrasis puede ser parcial o mutua. Es mutua si los elementos léxicos que establecen la relación de sinonimia pueden cambiarse el uno por el otro. Es parcial cuando la sustitución se realiza solamente en una dirección.

Por ejemplo, en el texto *Analysis of carbonaceous aerosols: interlaboratory comparison* (1998), las unidades léxicas **used / employing**, enmarcadas y unidas por líneas rojas, establecen una relación de paráfrasis simple mutua, ya que pueden intercambiarse sin que el significado de las mismas se vea alterado. Sin embargo, la relación establecida entre los términos **technique / method** es parcial, ya que **technique** puede reemplazar a **method**, pero no al contrario⁶.

⁶ La labor del informante ha sido determinante a la hora de precisar si estas dos unidades léxicas – **technique/method** – son sinónimas en el texto, ya que en español estos dos términos no son intercambiables

25. Another thermal **technique** based on coulometric detection of CO₂ is being **used** in Europe for occupational monitoring of diesel particulate.

26. Four laboratories **employing** the coulometric **method** also participated in the interlaboratory comparison, giving a total of eleven laboratories

3. 2. 1. 4. Paráfrasis Compleja: Dos elementos de la misma o distinta categoría gramatical. Este tipo de reiteración queda restringida a tres posibilidades:

(1) Antónimos que no comparten el mismo morfema léxico [*hot – cold*]. Así, las unidades léxicas, **slowly** y **quickly**, enlazadas por una línea de color violeta, establecen una relación de paráfrasis compleja por antonimia entre las siguientes oraciones:

11. In nylon thread, as seen in Figure 1, the flame propagates **slowly**.

12. In PET thread, shown in Figure 2, the flame propagates more **quickly** and is accompanied by black smoke.

(2) Dos unidades, una de ellas “is a complex repetition of the another item (for example, **writer** and **writings**) and also a simple paraphrase (or antonym) of a third (for example, **writer** and **author**).” (Hoey, 1991a: 64). En el siguiente ejemplo, las unidades léxicas **caught fire** e **ignition** establecen una relación de paráfrasis compleja. El elemento intermedio es **ignited** que determina una sinonimia con **caught fire** y a su vez una repetición compleja con **ignition**.

6. So we very carefully performed experiments to see what would happen when fibers **caught fire**.

8. Thread samples about 50 cm long were hung up just in front of a focused camera and then **ignited** at the bottom end with a tiny flame from a cigarette lighter.

14. When the **ignition** flame was removed, the threads continued to burn.

Un ejemplo de antónimo lo encontramos en las siguientes oraciones del texto *A healthy spread*. Los términos **risk** y **beneficial** establecen una paráfrasis compleja, ya que la unidad intermedia es **benefits**, que constituye una antonimia con **risk** y a su vez una repetición compleja con **beneficial**.

13. To reduce the **risk** of heart disease, people therefore need to lower both their total cholesterol levels and their LDL-cholesterol levels in the plasma.

19. As people become more conscious of the **benefits** of cutting down excess intake of fats, especially saturated fats, spreads that contain <80 per cent fat are gaining in popularity.

49. Although these products can be **beneficial**, Donnelly says that she hopes ‘people do not begin to rely on them because they are not miracle cures and there are many other factors involved in heart disease, which these products do not address.

(3) Por último, estableceremos una relación de paráfrasis compleja aún cuando la unidad léxica que sirve de enlace no esté presente en el texto, siempre que esta unidad “is capable of paraphrasing exactly in that context one of the items and of repeating the other”. (Hoey, 1991: 66). Así, en las siguientes oraciones del texto *High-precision conductometric detector for the measurement of atmospheric carbon dioxide* (1998), los elementos léxicos **described** y **reports** forman una repetición de paráfrasis compleja siendo su

unidad intermedia de enlace el término **descriptions**, no presente en el texto, que puede sustituir a **reports** y a su vez constituye una repetición compleja con **described**.

13. The new technique **described** here for measurement of CO₂ mixing ratios is based on the increase in conductivity that occurs when deionized water makes contact with air by use of microporous hollow fiber membranes.

15. There are previous **reports** of conductometric techniques for measuring gas-phase CO₂. (**descriptions** sería la unidad léxica intermedia, no presente en el texto, que establece una sinonimia con **reports**).

Igualmente, las unidades léxicas **higher** y **low**, extraídas de las siguientes oraciones del texto *Determination of cyanide in whole blood by capillary gas chromatography with cryogenic oven trapping* (1998), están relacionadas mediante una paráfrasis compleja. La unidad intermedia, no presente en el texto, es **high**, que forma una relación de antonimia con **low** y a su vez una de repetición compleja con **higher**.

7. In most of these reports, conventional packed columns, which give relatively **low** sensitivity and poor separation, were used.

14. This means that 10-100 times **higher** sensitivity can be obtained by this method as compared with that of the previous methods. (**high** sería la unidad intermedia, no presente en el texto, que establece una antonimia con **low**).

3. 2. 1. 5. Hiponimia y co-referencia. Sólo si tienen el mismo referente en común y si el hipónimo aparece primero [*bear – animals*] (Hoey, 1991a: 69). Por ejemplo, en las siguientes oraciones, **nylon** y **PET** establecen una relación de hiponimia con **fibers**. El término hiperónimo es **fibers**, relacionado con **nylon** y **PET** mediante una línea turquesa.

5. On the basis of these pyrolysis data, we were doubtful about the flammable but self-extinguishing classification for **nylon** and **PET**.

6. So we very carefully performed experiments to see what would happen when **fibers** caught fire.

También incluiremos en este tipo de relación de hiponimia aquellos casos en que el hiperónimo aparezca antes, siempre que las dos unidades “are interpreted as having identical referents” (Hoey, 1991: 70). Por ejemplo, en las siguientes oraciones del texto *The determination of food colours by HPLC with on-line dialysis for sample preparation* (1992), se considera que **technique** y **dialysis** establecen una relación de repetición aún cuando el hiperónimo aparezca antes, ya que el contexto nos indica que ambas unidades comparten el mismo referente⁷.

17. The power of the **technique** is further extended by allowing enrichment of the determinand in the dialysate on small trace enrichment cartridge prior to elution to the analytical HPLC column.

18. The combination of **dialysis** and trace enrichment then leads to a complete sample preparation systems for microconstituents of foods, which is marketed under the acronym ASTED (automated sample treatment through enrichment of dialysates).

3. 2. 1. 6. Otras formas de repetición.

Hasta ahora nos hemos concentrado en aquellos enlaces léxicos que permiten a un escritor 'decir algo de nuevo'. Hay, sin embargo, otros tipos de enlaces en el texto que tienen la misma función. Son los que pertenecen a los pronombres y sustitutos, estudiados por Quirk, Greenbaum, Leech y Svartvik. (1972, 1985); Winter (1974, 1979); Halliday y Hasan (1976) y Gutwinski (1976). Estas unidades textuales se diferencian de las unidades léxicas en que dependen de otros elementos para su

⁷ Es importante señalar que la labor del informante en estos casos es decisiva, debido a que la comprensión del contenido es determinante a la hora de precisar los referentes comunes.

interpretación y no tienen significado por ellas mismas. Su función es sustituir las unidades léxicas.

- A) **Pronombres personales.** De los pronombres personales sólo consideraremos que establecen una relación de repetición **he, she, it** y **they**. **I, you** y **we** serán considerados como repetición si aparecen en citas. Si no es así, éstos se refieren a algo externo al texto (exófora). En nuestro análisis se les da un trato especial. Se les considera como si fueran unidades léxicas y, por lo tanto, en la segunda y sucesivas apariciones son analizadas como enlaces léxicos. Sin embargo, son anotadas como casos dudosos y, como veremos, esto afectará a la interpretación de nuestros resultados. (Hoey, 1991a: 71)
- B) **Adjetivos y pronombres demostrativos.** Se considera que los pronombres **this, that, these** y **those** establecen enlaces de repetición sólo si funcionan como pronombres. Cuando funcionan como adjetivos, no se contabilizan como repetición, a no ser que los sustantivos a los que modifican no estén en repetición léxica o paráfrasis. La razón es puramente práctica. Si el sustantivo ya realiza la función del demostrativo mediante su repetición léxica o paráfrasis, establecer una unión con el demostrativo sería contar doble. (Hoey, 1991a: 72).
- C) **(an)other, the other, (the) same.** Se les considera igual que a los modificadores demostrativos. Si acompañan a una unidad léxica que establece una relación de repetición con otra unidad léxica anterior, son ignorados para evitar contar doble. Si la unidad léxica a la que acompañan no establece tal unión, entonces se considera que forman una relación de repetición. (Hoey, 1991a: 74).
- D) **Elipsis.** Cuando partes del discurso o de la oración deban ser suplidas por el lector para la correcta interpretación del mensaje. [*a work of art – the work*] (Hoey, 1991a: 74, 87).

3. 2. 2. Criterios para establecer enlaces entre oraciones.

- A) Tomaremos como punto de partida para nuestro análisis la oración, definida ésta, desde un punto de vista pragmático, como aquella parte del texto que empieza por mayúscula y termina en punto. Por esta razón, se presentan todos los textos con las oraciones numeradas para facilitar su reconocimiento. Cuando establecemos una relación de repetición, consideramos que ésta es siempre anafórica.
- B) Una unidad léxica forma una relación de repetición con cada una de sus apariciones previas, no sólo con su inmediata predecesora. Así, la unidad léxica **nylon** en la oración 4 repite no sólo **nylon** en la oración 3, sino también **nylon** en las oraciones 1 y 2, creando una red de enlaces como se muestra en las siguientes oraciones.

1. We would like to present two series of photographs showing the characteristic burning behaviors of a **nylon** fiber and a polyethyleneterephthalate (PET) fiber, in order to help people safely handle these fibers in their everyday lives.

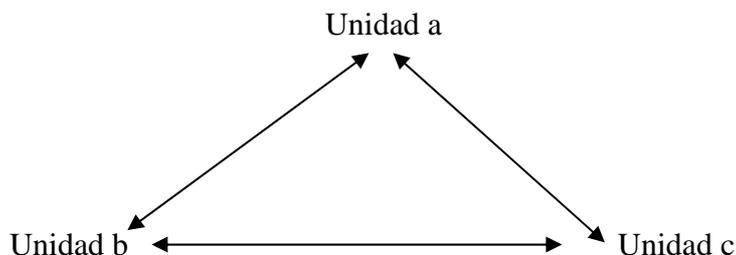
2. In many textbooks, especially on textiles, **nylon** and PET fibers are classified as flammable but self-extinguishing.

3. In other references, we have read that **nylon** and PET give off combustible gases when they are heated above their decomposition temperatures.

4. According to references, **nylon** gives propylene (8.8% in volume of total detected gases evolved), cyclopentanone (32.2%), hexamethylenediamine and other methylene amines (22.5%), and others (3), and PET gives ethylene (8.3% in volume of total detected gases evolved), acetaldehyde (10.9%), benzoic acid (37.5%), and other phenyl compounds.

- C) Si una unidad léxica *a* establece un enlace con las unidades léxicas *b* y *c*, las unidades léxicas *b* y *c* establecerán igualmente una relación de

repetición. La secuencia es irrelevante. Esta relación se puede representar por medio del siguiente triángulo (Hoey, 1991a: 65):



Así, en las siguientes oraciones, la unidad léxica **PET** está relacionada con **polymer** y **fibers**. Por medio del triángulo, estableceremos un enlace entre **polymer** y **fibers**, indicado por medio de una línea quebrada negra.

12. In **PET** thread, shown in Figure 2, the flame propagates more quickly and is accompanied by black smoke.

15. During the course of the burning, pictures were taken of the small spherical fire balls composed of a molten **polymer**.

20. If you want to do this type of experiment, you should wear a glove made of non-flammable **fibers** so your hand will not be burned.

D) Consideraremos que una relación es dudosa de acuerdo con los criterios aplicados a nuestro análisis cuando dos unidades léxicas no tengan el mismo referente, pero la similitud en el contexto permita establecer un enlace mucho más débil (Hoey, 1991a: 62). De esta forma, en las siguientes oraciones, **change** en la oración 1 se refiere a los cambios en la composición traza de la atmósfera y **change** en la oración 5 se refiere al cambio climático. Establecemos una relación de repetición simple dudosa, señalada con línea quebrada, porque,

aunque los referentes sean distintos, el cambio climático se produce precisamente por los cambios en la composición traza de la atmósfera

1. The recent increase in atmospheric CO₂ mixing ratio is one of the most significant **changes** in the trace gas composition of the atmosphere.

5. In addition, flux measurements are necessary to improve the global circulation models that predict future CO₂ concentrations and climate **change**.

E) También consideraremos dudosas las relaciones establecidas por los pronombres de primera y segunda persona cuando hacen referencia a algo no perteneciente al texto. Por ejemplo, en las siguientes oraciones **we** presenta una referencia exofórica.

6. So **we** very carefully performed experiments to see what would happen when fibers caught fire.

7. **We** selected typical sewing threads for sewing machine (supplied by Teijin Co., Ltd., and Asahi Chemicals Co., Ltd.) for testing.

F) Una unidad léxica en una oración sólo puede establecer un enlace con otra unidad léxica en otra oración. El orden de prioridad en caso de que una unidad léxica esté relacionada con dos elementos léxicos es el siguiente:

- repetición simple
- repetición compleja
- paráfrasis simple mutua
- paráfrasis simple parcial
- paráfrasis compleja de antonimia
- otros tipos de paráfrasis
- sustitución
- co-referencia
- elipsis

(Hoey, 1991a: 83)

Así, en las siguientes oraciones, la unidad léxica **PHA** establece una repetición simple con **PHA** y una relación de hiponimia con **polymer**. Por el criterio de prioridad elegimos el enlace de repetición simple.

8. The backbone structure of **PHA** meant that it could be a useful thermoplastic (softens on heating) for forming into films and fibres.

9. Also at temperatures of ca 180-200° C, **PHA** converts with very little mass loss to water and a different **polymer**, i.e. the rigid high – strength polybenzoxazole (PBO, 2), which decomposes only at very high temperatures (ca 600° C).

Igualmente, en las siguientes oraciones, el elemento léxico **synthetic** está relacionado con **synthetic** por medio de una repetición simple y con **natural** mediante antonimia. La unidad léxica **dyes** establece una repetición simple con **dyes** y una paráfrasis simple mutua con **colours**. Por el criterio de prioridad elegiremos la repetición simple para ambas unidades.

3. The current trend is, however, away from the use of such **synthetic** **dyes** despite the extensive toxicological screening which they have undergone.

4. The lists of permitted **synthetic** **dyes** are progressively being reduced and a number of food processors are relying on the use of **natural** **colours** to impart the desired colour to their products.

3. 2. 3. Elaboración de las matrices de repetición.

Una vez realizado el análisis detallado y minucioso de las relaciones que establecen las unidades léxicas de cada oración con las unidades léxicas del resto de oraciones pertenecientes al texto, éstas se presentan de dos formas:

- A) Matriz con las unidades léxicas pormenorizadas junto con el tipo de repetición que establecen. A continuación se presenta, como ejemplo, una parte de la matriz de repeticiones léxicas – oraciones 1, 2, 3, 4 y 5 – del texto *Both nylon and PET fibers burn continuously under atmospheric conditions* (1999), junto con el análisis de repeticiones de la oración 1 como muestra de cómo se ha efectuado la recogida de datos. Las oraciones son las siguientes:

1. **We** would like to present two series of photographs showing the characteristic burning behaviors of a **nylon fiber** and a polyethyleneterephthalate (**PET**) fiber, in order to help people safely handle these **fibers** in their everyday lives.
2. In many textbooks, especially on textiles, **nylon** and **PET fibers** are classified as flammable but self-extinguishing.
3. In other references, **we** have read that **nylon** and **PET** give off combustible gases when they are heated above their decomposition temperatures.
4. According to references, **nylon** gives propylene (8.8% in volume of total detected gases evolved), cyclopentanone (32.2%), hexamethylenediamine and other methylene amines (22.5%), and others (3), and **PET** gives ethylene (8.3% in volume of total detected gases evolved), acetaldehyde (10.9%), benzoic acid (37.5%), and other phenyl compounds.
5. On the basis of these pyrolysis data, **we** were doubtful about the flammable but self-extinguishing classification for **nylon** and **PET**.

En las celdas se presentan en color negro las unidades léxicas, por orden de aparición, que establecen repetición con las oraciones posteriores y en color rojo los elementos léxicos de la siguiente oración. Así, en la celda primera nos encontramos en color negro las unidades **nylon**, **PET** y **fibers** de la oración número 1 que establecen repetición simple con **nylon**, **PET** y **fibers**, en color rojo, de la oración número 2. Además, se especifica el tipo de repetición que forman para su contabilización. Las abreviaturas en negrita utilizadas en la matriz de este texto son las siguientes:

- | | |
|---------------------------------------|----------------------------------|
| rs. = repetición simple | rc. = repetición compleja |
| psm. = paráfrasis simple mutua | hip. = hiponimia |
| e. = elipsis | d. = deixis |
| + = dudosa (referencia exofórica) | |

1				
2	rs. nylon – nylon rs. PET –PET rs. fibers – fibers	2		
3	rs. we – we+ rs. nylon – nylon rs. PET –PET e. fiber – 0	hip. textbooks – references rs. nylon – nylon rs. PET –PET e. fibers – 0 psm. flammable – combustible	3	
4	rs. nylon – nylon rs. PET –PET e. fiber – 0	hip. textbooks – references rs. nylon – nylon rs. PET –PET e. fibers – 0	rs. references – references rs. nylon – nylon rs. PET –PET psm. give off – give rs. gases – gases	4
5	rs. we – we+ rs. nylon – nylon rs. PET –PET e. fiber – 0	rs. nylon – nylon rs. PET –PET e. fibers – 0 rc. classified – classification rs. flammable – flammable rs. self-extinguishing – self-extinguishing	rs. we – we+ rs. nylon – nylon rs. PET –PET psm. combustible – flammable	d. gives... compounds – these rs. nylon – nylon rs. PET –PET

Este tipo de matriz nos ofrece una valiosa información de cómo el texto está estructurado mediante la cohesión, de cómo las oraciones próximas y lejanas se interrelacionan mediante la repetición del léxico. Además, nos permite la contabilización de los diferentes tipos de reiteración para una posible diferenciación entre los dos tipos de texto que se van a analizar en nuestra investigación.

B) Matriz que contabiliza el número de enlaces que cada oración establece con las anteriores y posteriores. Así, de la matriz anterior se extrae la siguiente información que nos indica la densidad de enlaces entre oraciones. Los

números entre paréntesis indican el número de enlaces teniendo en cuenta los enlaces dudosos.

	1			
2	3	2		
3	3 (4)	5	3	
4	3	4	5	4
5	3 (4)	6	3(4)	3

3. 2. 4. Número de enlaces y conexiones entre oraciones.

Consideraremos que existe una conexión entre oraciones cuando señalemos, como mínimo, tres enlaces entre las mismas: “any two sentences are connected as packages of information if they share at least three points of reference” (Hoey, 1991a: 36). La razón es meramente práctica. Si reducimos este umbral, entonces todas las oraciones van a aparecer relacionadas con las demás, lo cual no nos va a aportar datos relevantes sobre la cohesión. Phillips (1985) también llegó a la misma conclusión al tomar tres enlaces como mínimo para llegar a resultados significativos.

Existen circunstancias en que tendremos que tomar más de tres enlaces para establecer una conexión: “if a high proportion of sentences are linked by three or more repetitions, it is necessary to increase the cut-off point” (Hoey 1991a: 91). ¿Qué proporción se considera la adecuada? Hoey (1991a) no aporta un criterio fijo: dependiendo del tipo de texto y de la cantidad de enlaces, se tendrá en cuenta tres o más. El criterio que hemos adoptado en nuestra investigación ha sido establecer el tanto por ciento de los pares de oraciones que establecen como mínimo tres enlaces, que no debe ser superior a aproximadamente el 30%; si no es así, incrementaremos el número de enlaces hasta conseguir esta proporción. Este 30% aproximado constituye el umbral que hemos establecido *a priori*, ya que se podría considerar, dentro del ámbito estadístico, como el que puede ofrecer datos más fiables, según el Dr. Espinosa Hidalgo. Para ello confeccionamos la siguiente tabla que muestra los porcentajes de la cantidad de enlaces:

Número de enlaces	Número de celdas	Tanto por ciento
0 (sin tener en cuenta los enlaces dudosos)	72	37'8%
0 (teniendo en cuenta los enlaces dudosos)	67	35'2%
A partir de 3 (teniendo en cuenta los enlaces dudosos)	31	16'3%
Total de celdas	190	

En este texto en cuestión tomaremos tres enlaces para establecer una conexión, aún cuando represente el 16'3% del total de celdas, ya que, como hemos señalado anteriormente, dos enlaces no son significativos para establecer una conexión. En otros textos tendremos que incrementar el número de enlaces cuando el porcentaje se incremente más allá de aproximadamente el 30%. En cada texto analizado indicaremos el número de enlaces que hemos tomado como mínimo para establecer conexiones entre oraciones.

De la matriz numérica extraemos la siguiente tabla que muestra las conexiones que cada oración presenta con el resto de las oraciones del texto. El primer dato numérico entre primer paréntesis indica el número de conexiones – en este texto a partir de tres enlaces – que la oración establece con las anteriores; el segundo dato numérico entre primer paréntesis indica el número de conexiones que establece con las posteriores. De esta forma, la oración 2 establece 1 conexión con una oración anterior y 5 con posteriores. Las rayas que aparecen en las oraciones 1 y 20 indican que no es posible establecer conexiones con oraciones anteriores a la oración primera del texto, ni tampoco establecerlas con las oraciones posteriores a la última oración del texto. Los datos numéricos entre corchetes y en rojo contabilizan el total de conexiones de cada oración con todas las oraciones del texto. Los datos numéricos entre segundos paréntesis y corchetes contabilizan las conexiones teniendo en cuenta los enlaces dudosos.

1. (-,8) [8]	2. (1,5) [6]	3. (2,3) [5]
4. (3,2) [5]	5. (4,1) [5]	6. (0,0) [0]
7. (0,0) [0]	8. (0,3) [3]	9. (1,0) [1]
10. (4,3) [7]	11. (1,1) [2]	12. (3,0) [3]
13. (3,0) [3]	14. (2,0) [2]	15. (1,3) [4] (1,4) [5]
16. (0,0) [0]	17. (1,0) [1]	18. (1,0) [1] (1,1) [2]
19. (1,0) [1]	20. (1,-) [1] (3,-) [3]	

Este tipo de tabla ofrece una información fundamental para nuestra investigación, ya que muestra qué oraciones presentan un mayor número de conexiones – oraciones centrales – y cuáles establecen 0 conexiones – oraciones marginales –. Además, indica qué oración / es establecen un mayor número de conexiones con las posteriores – oración principal – y cuáles establecen un mayor número de conexiones con las anteriores – oración concluyente –. En resumen, se obtiene una representación gráfica de las hipótesis que esta tesis doctoral se propone verificar.

3. 3. Procedimiento para la verificación de las hipótesis.

Mientras que el criterio para considerar una oración marginal es fijo, no ocurre igual con las oraciones centrales: “just as the minimum number of links criterial for bonding varies from text to text, so the number of bonds criterial for centrality likewise varies” (Hoey, 1991a: 113-114). Sugiere que el número de conexiones mínimas para establecer las oraciones centrales podría fijarse a partir del número de conexiones que ninguna oración establece. Cada texto presentará un umbral distinto e incluso puede presentar dos umbrales, como le ocurre a Hoey (1991a) en su texto analizado y a nosotros en los nuestros.

Este criterio nos presentó las siguientes dificultades con respecto a la definición de umbral y al informante. En relación con la definición de umbral, no todos los textos presentan este umbral de forma tan clara. Así, este texto, según muestra la siguiente tabla, no presenta un número de conexiones que ninguna oración establezca. Decidimos incrementar a 1 oración para fijar el umbral. Como podemos observar, el número de conexiones que 1 oración establece es 4 y 6. El umbral sería, por lo tanto, 4 y 6 conexiones, ya que, si tomamos 7 u 8 conexiones, el resumen contendría sólo una o dos oraciones.

Número de conexiones	Número de oraciones
0	3
1	5
2	2
3	3
4	1
5	3
6	1
7	1
8	1

Presentamos al informante los 2 resúmenes teniendo en cuenta estos 2 números mínimos de conexiones – 4 y 6 – y, cuando recabamos su información sobre la validez de los mismos, observamos que 5 es el número mínimo a tener en cuenta para establecer las oraciones centrales.

Igualmente, la introducción del artículo de investigación *Anayisis of carbonaceous aerosols: interlaboratory comparison* no presenta un número de conexiones que ninguna oración establezca o que establezca 1 oración, como muestra la siguiente tabla:

Número de conexiones	Número de oraciones
0	1
1	1
2	6
3	2
4	2
5	5
6	4
7	3
8	2
9	1

Decidimos incrementar el número a 2 oraciones para fijar el umbral. Como podemos observar, el número de conexiones que 2 oraciones establecen es 3 y 4. El umbral sería, por lo tanto, 3 y 4 conexiones, ya que, si tomamos 8 conexiones, el resumen contendría sólo dos oraciones.

Presentamos al informante los resúmenes confeccionados con las oraciones que establecen a partir de 3 y 4 conexiones y, cuando recabamos su información sobre la validez de dichos resúmenes, observamos que, si elevamos el número de conexiones a 6, el resumen contiene la información fundamental del texto original.

En otros textos, aunque el umbral era claro, el número de conexiones fijado mediante dicho umbral no nos resultó válido. Así, la introducción del artículo de investigación *Speciation as an analytical aid in trace element research in infant nutrition* presenta la siguiente relación entre el número de conexiones y oraciones:

Número de conexiones	Número de oraciones
0	3
1	3
2	1
3	2
4	2
5	0
6	2
7	1
8	0
9	2
10	2
11	2
12	0
13	2

Este texto presenta varios umbrales: 5 y 8 conexiones. Presentamos al informante los resúmenes teniendo en cuenta estos dos números mínimos de conexiones y, cuando recabamos su información sobre la validez de los mismos, observamos que 10 es el número mínimo a tener en cuenta para establecer las oraciones centrales y, por lo tanto, para ofrecer un resumen que contenga la información fundamental del texto original.

En cuanto al informante, el criterio para la definición de umbral propuesto por Hoey (1991a) presentó el grave problema de que el informante, al tener que efectuar la lectura de varios resúmenes diferentes de un mismo texto, perdía la noción de lo que se supone constituye la información esencial. Él mismo nos comentaba la confusión que le producía la introducción o eliminación de una o varias oraciones a la hora de juzgar la

marginalidad o esencialidad de las mismas en los textos analizados. Por esta razón, modificamos el criterio de definición de umbral propuesto por Hoey (1991a) para fijar las oraciones centrales, ya que no nos proporcionaba resultados fiables.

A esto habría que añadir que el informante tiene que efectuar la lectura de cuatro resúmenes de un mismo texto original:

- a) El resumen confeccionado con las oraciones centrales, cuyo número de oraciones, como acabamos de señalar no es fijo, sino que va a oscilar dependiendo de si contiene o no la información esencial.
- b) El resumen confeccionado con la eliminación de las oraciones que establecen 0 conexiones.
- c) El resumen confeccionado con las oraciones que conexionan con la oración principal.
- d) El resumen confeccionado con las oraciones que conexionan con la oración concluyente.

Pudimos observar con preocupación que el criterio del informante sobre esencialidad y/o marginalidad fluctuaba entre texto y texto y, lo que es aún peor, que dicho criterio se veía mediatizado e influenciado al presentarle los resúmenes ya elaborados.

Por ello, optamos por la siguiente solución para fijar las oraciones centrales: el Dr. Espinosa Hidalgo lee el texto original y señala las que, desde el punto de vista de un científico, considera como oraciones esenciales. Con esta información objetiva se analizan las características propias de las oraciones del texto con respecto a la capacidad de las mismas para generar conexiones entre sí. De esta forma, si una oración seleccionada por el informante como esencial no establece suficientes conexiones y existe otra u otras en el texto que parafrasean a la seleccionada por el especialista y presentan un mayor número de conexiones, ésta o éstas serán las elegidas. Por ejemplo, las siguientes oraciones de la introducción del artículo de investigación *Anayisis of carbonaceous aerosols: interlaboratory comparison* presentan una información similar:

3. Some ~~of these~~ aerosols [asphalt fumes, oil mists, cigarette and wood smokes, carbon black, and diesel exhaust] are known or suspect human carcinogens (e.g., cigarette smoke and diesel exhaust, respectively) and have been linked to other adverse health effects (e.g., asthma, heart disease)
4. Exposure to diesel exhaust is of particular concern because it has been classified a probable human carcinogen and diesel equipment use is widespread in (e.g., trucking, transit, mining, railroads, agriculture).

La oración 3 establece 7 conexiones, mientras que la 4 establece 6. Elegiremos la 3 como esencial, puesto que es la que establece más número de conexiones.

Estas oraciones esenciales serán denominadas como pertenecientes al 'cuerpo del texto', ya que son las que portan la información fundamental del mismo y van a ser las que se utilicen para fijar el número mínimo de conexiones a tener en cuenta para establecer las oraciones centrales, número que se obtiene a partir del número mínimo de conexiones que establecen las oraciones seleccionadas por el informante como pertenecientes al 'cuerpo del texto'. Éstas, en principio, deberán aparecer en todos los resúmenes independientemente de si se elaboran mediante la eliminación de oraciones marginales, la combinación de oraciones centrales, a partir de la oración principal o a partir de la oración concluyente. Cuando los resúmenes confeccionados mediante las conexiones que establece la oración principal o concluyente no contengan estas oraciones pertenecientes al 'cuerpo del texto', el informante realizará la lectura de los mismos y verificará si pueden considerarse válidos aún cuando no contengan las oraciones pertenecientes al 'cuerpo del texto'.

Siguiendo con el ejemplo del texto que aquí presentamos, las oraciones pertenecientes al 'cuerpo del texto' son las siguientes: 2, 3, 4, y 10.⁸

2. In many textbooks, especially on textiles, nylon and PET fibers are classified as flammable but self-extinguishing.
3. In other references, we have read that nylon and PET give off combustible gases when they are heated above their decomposition

⁸ La cohesión del texto original se suprime, porque consideramos las oraciones como 'paquetes' independientes de información. Por ello, las referencias de pronombres y artículos determinados aparecen en los resúmenes entre corchetes. Igualmente, se eliminan las marcas de coherencia marcando en el texto resultante una línea que las cruza.

temperatures. **4.** According to references, nylon gives propylene (8.8% in volume of total detected gases evolved), cyclopentanone (32.2%), hexamethylenediamine and other methylene amines (22.5%), and others (3), and PET gives ethylene (8.3% in volume of total detected gases evolved), acetaldehyde (10.9%), benzoic acid (37.5%), and other phenyl compounds.

10. ~~Once~~ a part of the terminal end [of thread samples about 50cm long] was ignited, [with a tiny flame from a cigarette lighter] [and] it burned continuously, as shown in the series of photographs in this paper, in contrast to the descriptions in textbooks stating that it "burns slowly but if the sample is removed from the flame it self-extinguishes".

La tabla de las oraciones con el número de conexiones en la sección 3. 2. 4., muestra que el número mínimo de conexiones establecidas es 5, por lo tanto éste constituirá el umbral a tener en cuenta y tomaremos todas las oraciones que posean como mínimo este número de conexiones. Las oraciones 1, 2, 3, 4, 5 y 10 serán, pues, las oraciones centrales y las que ofrezcan el resumen.

1. We would like to present two series of photographs showing the characteristic burning behaviors of a nylon fiber and a polyethyleneterephthalate (PET) fiber, in order to help people safely handle these fibers in their everyday lives.

2. In many textbooks, especially on textiles, nylon and PET fibers are classified as flammable but self-extinguishing. **3.** In other references, we have read that nylon and PET give off combustible gases when they are heated above their decomposition temperatures. **4.** According to references, nylon gives propylene (8.8% in volume of total detected gases evolved), cyclopentanone (32.2%), hexamethylenediamine and other methylene amines (22.5%), and others (3), and PET gives ethylene (8.3% in volume of total detected gases evolved), acetaldehyde (10.9%), benzoic acid (37.5%), and other phenyl compounds.

5. On the basis of these pyrolysis data, we were doubtful about the flammable but self-extinguishing classification for nylon and PET.

10. ~~Once~~ a part of the terminal end [of thread samples about 50cm long] was ignited, [with a tiny flame from a cigarette lighter] [and] it burned continuously, as shown in the series of photographs in this paper, in contrast to the descriptions in textbooks stating that it "burns slowly but if the sample is removed from the flame it self-extinguishes".

(15. During the course of the burning, pictures were taken of the small spherical fire balls composed of a molten polymer.)

La oración 15, entre paréntesis, establece 4 conexiones y, si tenemos en cuenta el enlace dudoso con la oración 20 (véase matriz de repetición de unidades léxicas en el anexo, sección 2. 3. 1.) presenta 5 conexiones (dato numérico recogido en el segundo paréntesis de la tabla de la sección 3. 2. 4.). Esta oración 15, como podemos observar, no es necesaria para el resumen, ya que no pertenece al ‘cuerpo del texto’ Por lo tanto, aquellas dudosas no necesarias para el resumen no las contabilizaremos a la hora de realizar el cómputo definitivo de oraciones que constituyen un resumen. En este texto, con 20 oraciones, 6 oraciones son las centrales, representando el 30% del total de oraciones del texto original.

El procedimiento anteriormente mencionado nos conduce, pues, a la obtención de los siguientes datos:

- e) Número de enlaces y conexiones tomadas en cada texto y posible diferenciación entre los dos tipos de texto analizados.
- f) Porcentaje de oraciones esenciales y marginales con respecto al total de oraciones del texto original y su posible comparación con los dos tipos de texto analizados.
- g) Verificación de la hipótesis de la eliminación de las oraciones que establecen 0 conexiones.
- h) Confirmación de la oración principal y concluyente.