

## CAPÍTULO 1

### **EL DISCURSO CIENTÍFICO: COHERENCIA Y COHESIÓN.**

En este capítulo se realiza un breve recorrido por las teorías más representativas y relevantes del análisis del discurso de la ciencia para mostrar el estado de la cuestión. Asimismo, se ofrece una exposición detallada de la teoría de la coherencia y cohesión, incluyendo las aportaciones de diversos autores. Por último, se presenta el modelo de análisis léxico de Hoey (1991a) objeto de nuestro estudio.

#### **1. 1. El discurso científico.**

Un área de gran interés dentro de la Lingüística Aplicada es el estudio de la forma en que los científicos se expresan en inglés (Bloor y Bloor, 1995: 221). El discurso de la ciencia “is one of those concerns which functional linguists find interesting and which an unusually wide range of researchers finds interesting too.” (Martin, 1998: 4).

Por otra parte, no cabe duda de que el inglés se ha convertido a lo largo de los últimos años en la *lingua franca* de la comunicación entre investigadores (Wood, 2001).

“In England and America, when questions of employment or tenure are discussed in academic contexts, the slogan ‘Publish or perish!’ is frequently heard. When similar discussions come up in non-English academic contexts, the slogan changes into ‘publish in English or perish!’.” (Ventola y Mauranen, 1991: 457).

Según Swales (1987: 42) aproximadamente el 50% de los artículos que se publican en el mundo están escritos en inglés. Este avance del inglés se está viendo impulsado aún más por su uso en Internet: “A recent report on America’s National

Public Radio estimated that 85 per cent of the world's homepages used English, and only two per cent were in French.” (Swales, 1998: 3).

En el ámbito internacional, científicos cuya lengua materna no es la inglesa necesitan leer libros y revistas especializadas publicadas en inglés y cada vez son más numerosos los que se ven obligados a publicar sus trabajos de investigación en lengua inglesa:

“Alrededor del 90% de la información recogida en los bancos de datos estadounidenses provienen de aportaciones en inglés, el resto (10%) se reparte entre ruso, alemán, japonés y francés.

Por otro lado, el hecho de que el inglés sea la lengua de las publicaciones hace que sea también la lengua de comunicación entre los científicos”. (Luján García, 2001: 702).

Por otra parte, es bien sabido que no es posible efectuar una línea divisoria clara entre lengua específica y lengua general, puesto que todos los rasgos de cualquier lengua específica están extraídos de la lengua general a la que pertenece:

“one could not...state the “rules” which determine what parts of the grammar or lexicon of English may or may not appear in “scientific English” ...Hence we have more a continuum than a division between LSP and LGP” (Varantola, 1986: 14).

Hutchinson y Waters (1980) opinan que no existe un tipo de comunicación denominada científica o técnica. Para ellos la diferencia entre la lengua general y la lengua de la ciencia se basa en el vocabulario y la frecuencia de uso de determinadas formas gramaticales. Mantienen que “there is little justification for a subject-specific approach to ESP” (Hutchinson y Waters, 1987:166), excepto en lo que concierne a la motivación del alumno. Tarantino (1991: 53), sin embargo, sugiere que expresiones tales como ‘laser-induced light shift’ sólo pueden generarse, utilizarse y ser apreciadas dentro del área de la ciencia. Halliday y Martin (1993:56) sugieren que un texto se reconoce como “scientific English” debido al efecto combinado de “clusters of features” y “it is the combined effect of a number of such related features, and the relations they contract throughout the text as a whole, rather than the obligatory presence of any particular ones, that tells us what is being constructed is the discourse of science.”

## **1. 2. Evolución de la descripción lingüística en el inglés científico.**

Los diversos enfoques de estudio lingüístico sobre la prosa científica, al igual que el concepto de ESP, han ido evolucionado con las teorías lingüísticas y del discurso a lo largo de los años. Conocerlos nos puede servir para orientarnos hacia donde se dirige la investigación del inglés científico.<sup>1</sup> Como es bien sabido, estos estudios se han llevado a cabo desde tres perspectivas: estudios de frecuencia, análisis retórico y análisis del género.

### **1. 2. 1. Análisis de frecuencia o registro.**

Los estudios de frecuencia han sido importantes desde el comienzo de ESP. El concepto de registro se remonta al antropólogo Malinowski en 1923. Malinowski acuñó el término contexto de situación, noción que ejerció una gran influencia en las investigaciones de Firth y más tarde en Halliday. El argumento principal en el que se basaba era que para entender una proposición debemos no sólo tener en cuenta el significado literal de las palabras, sino también todo el conjunto de particularidades sociales en los que la proposición se realiza, ya que “the meaning of any single word is to a very high degree dependent of its context” (Malinowski 1946: 307). Lingüistas como Firth (1957a) y Pike (1967) con sus “restricted languages” referidos a “circumscribed fields of experience or action” y que tienen cada uno “its own grammar and dictionary” (Firth, 1957a: 124, 87, 98 y 105) se fijaron ya en la descripción de la lengua utilizada en áreas tales como la ciencia, la tecnología, el comercio, el deporte o la meteorología: “An area of English for specific purposes (ESP) which could be considered as having its roots in Firth’s concept of ‘restricted languages’ is that of register analysis” (McLaren, 1992: 197).

Halliday, discípulo de Firth, divulgó en los años sesenta el término registro, identificándolo como “a variety according to use, in the sense that each speaker has a range of varieties and chooses between them at different times” (Halliday, McIntosh y Strevens, 1964: 77) y en 1978 precisó el objetivo que debía tener el análisis del registro:

---

<sup>1</sup> Sanz Sainz (2002) realiza un recorrido por las diferentes fases que pueden identificarse en la evolución de ESP o, en español, IFE.

“What a theory of register does is to attempt to uncover the general principles which govern this variation (language variation in context), so that we begin to understand what situational factors determine what linguistic features” (Halliday 1978: 32).

A partir del concepto de registro, la investigación lingüística inglesa de los años sesenta, dentro de los límites de la lingüística oracional, intenta fijar no sólo una metodología de estudio e investigación, sino una clasificación de las distintas variedades que una lengua puede presentar, según la utilidad que se dé a dichas variedades. El concepto de registro queda definido por Halliday, McIntosh y Stevens (1964: 87) en los siguientes términos:

“Language varies as its function varies; it differs in different situations. The name given to a variety of a language distinguished according to its use is register.”

Los registros se diferencian entre sí “by having a number of distinctive words, by using words or phrases in a particular way, and sometimes by special grammatical constructions” (Richards, Platt y Platt 1992: 313). De igual manera se manifiesta Alcaraz (2000: 23) cuando señala que en un registro determinado “sobresale un conjunto de rasgos léxicos y gramaticales que, por repetirse con frecuencia y regularidad, le confiere un cierto sentido de unidad diferenciada”. Los químicos, por ejemplo, usan un registro caracterizado por la reiteración de ciertas formas léxicas y gramaticales y pueden intercambiar entre ellos información y conocimientos especializados, aún cuando no dominen la lengua de su interlocutor, ya que la misma comunidad epistemológica a la que pertenecen ha creado un código que, en determinadas circunstancias, resulta suficiente para poder intercambiar información con otro especialista del mismo campo.

Los trabajos de Herbert (1965) o Ewer y Latorre (1969), que analizaron un *corpus* extenso de textos científicos especializados para establecer las diferencias entre diversos registros, son representativos de este enfoque:

“In order to get a working idea of what this basic language consisted of, a frequency-analysis of the English actually used by scientific writers was required. ....The sample used was comprehensive. In subject, it covered ten main areas of science and a large

number of individual disciplines from anatomy to volcanology. ....In size the sample amounted to well over three million words, almost all of which had been written in the last twenty years” (Ewer y Latorre 1967: 221-222).

Siguiendo a Halliday, McIntosh y Stevens (1964) y Halliday y Martin (1993), los factores situacionales que nos conducen al registro son campo (*field*), tenor (*tenor*) y modo (*mode*), los cuales constituyen la aplicación de las tres macrofunciones en las que Halliday divide el lenguaje: ideacional o experiencial, función conectada con las ideas y hechos; interpersonal, en conexión con las interrelaciones entre emisor y receptor; y textual, ámbito en el cual se organiza el contenido de la información.

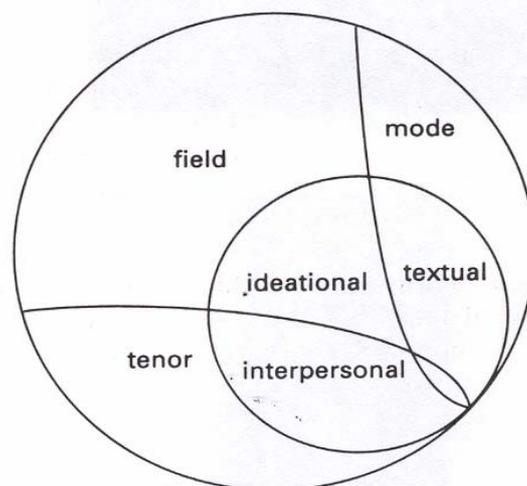


Figura 1. Relación entre las macrofunciones de la lengua y las variables del contexto. (Halliday y Martin, 1993).

Estas tres macrofunciones ejercen su influencia en todos los niveles del lenguaje, desde la semántica a la fonología, pasando por la lexicología y la morfosintaxis, y todos ellos se ven afectados por las características situacionales que impone el registro.

Halliday (1964: 90) define *campo* (field) como la forma en que “registers are classified according to the nature of the whole event of which the language activity forms part”. Gregory y Carroll (1978: 8) redefinen el concepto de una forma más clara: “field of discourse is the consequence of the user’s purposive role, what his language is

‘about’, what experience he is verbalizing, ‘what is going on’ through language”. Más recientemente el mismo Halliday (1993: 32) redefine el concepto como “what is actually taking place”.

*Tenor* es definido por Gregory y Carroll (1978: 8) como “the relationship the user has with his audience”. El *tenor* en palabras de Halliday (1993: 32) “refers to who is taking part, to the nature of the participants, their statuses and roles”.

Halliday (1964: 91) define *modo* como “the medium or mode of the language activity”. Más tarde lo redefine como “the symbolic organisation of the text, the status that it has, and its function in the context, including the channel” (Halliday, 1985: 12).

Aunque estas categorías situacionales y contextuales han sido redefinidas en Gregory (1967), Cristal y Davy (1969), Ellis y Ure (1969), Hasan (1973) y Gregory y Carrol (1978), etc., los registros se han definido en términos léxicos-gramaticales y más recientemente en términos semánticos.

Los trabajos de Barber (1962), Cristal y Davy (1969), Ewer y Latorre (1969), Huddleston (1971) y Gustafsson (1975), son representativos de este enfoque de análisis caracterizado por ser estudios cuantitativos de las características lingüísticas de los textos.

Este análisis del registro, de tipo cuantitativo, centrado en el estudio de la oración, que sólo tiene en cuenta las características formales del texto ha sido muy criticado porque, al ser meramente descriptivo, resulta muy limitado desde una perspectiva textual. Además, algunas veces es difícil definir los tipos de texto aún cuando pertenezcan al mismo registro. McCarthy y Carter (1994: 22) nos indican que la lengua de la predicción metereológica, por ejemplo, puede variar dependiendo de la situación: “there are as many weather forecasts as there are situations in which such forecasts are made.” De esta forma, la noción de registro “becomes so slippery as to be impractical for purposes of defining different types of text” (McCarthy y Carter, 1994: 24). Igualmente Bex (1996: 104-105) señala que no es posible decir que “each text can only realise its own register” e indica que se necesita “a higher level of analysis which groups texts into types and demonstrates what they have in common both in terms of their particular linguistic selections and in terms of their shared social purpose”. Este “higher level” es lo que podríamos denominar género. En este mismo sentido parece manifestarse Dudley-Evans (1987: 74) cuando señala:

“If we are to teach the writing of certain very specific texts such as ... the business letter or telex, or a business or technical report, we need a system of analysis that shows how each type of text differs from other types”.

Robinson (1980) considera necesaria una mayor precisión y una menor generalización acerca de las supuestas características de ciertos registros. Asimismo, Widdowson (1983: 28) critica estos estudios tradicionales del registro del léxico y estructuras gramaticales y aboga por un nuevo enfoque que tenga en cuenta los aspectos relacionados con la competencia comunicativa:

“it [se refiere a registro] is an operation on text and does not, as such, reveal how language is used in the discourse process”.

Es cierto que este análisis del registro, de tipo cuantitativo, centrado en el estudio de la oración, ha sido muy criticado porque al ser meramente descriptivo resulta muy limitado desde una perspectiva textual. Sin embargo, hay autores que los consideran válidos e incluso adecuados para las investigaciones de aspectos concretos de la lengua. El mismo Swales (1983: 26) afirma:

“Frequency analyses are descriptions not explanations; they are not discovery-procedures, but they can often indicate which features do call for some sort of exploration... Twenty years later, there are things that are criticized in Barber – certainly his selection of texts- but his modesty, accuracy and intellectual honesty set standards that his successors have not found it always easy to maintain.”

En los últimos años, las opiniones, aunque críticas, muestran una actitud de respeto y reconocimiento ante estos estudios. Bhatia (1993: 17) reconoce la importancia y la utilidad del análisis del registro en el campo de la lingüística, pero señala que en pocas ocasiones se ha establecido una distinción entre variedad (registro) y género. Así, muchos de estos estudios consideran ejemplos representativos del discurso científico, al mismo nivel, tanto un artículo científico de investigación como un informe de laboratorio, sin establecer distinciones entre las características de cada género y, al

mismo tiempo, dando la impresión de que un determinado género es diferente si pertenece a diversas áreas de conocimiento.

“Analyses of varieties or registers on their own reveal very little about the true nature of genres and about the way social purposes are accomplished in and through them in settings in which they are used.” (Bhatia, 1993: 18).

De igual manera se manifiesta Leckie-Tarry (1993: 40) cuando señala que cualquier intento de caracterizar la lengua o la variación dentro de una lengua, “must work through the concepts of register and genre, and ... any characterisation of register/genre .... must specify both contextual features at various levels as well as linguistic features”.

Dentro de la escuela sistémica Reid (1987) o Martin (1992a) consideran también que la teoría del registro es insuficiente para explicar la relación texto / contexto, ya que se centra básicamente en los elementos lingüísticos contextuales de la situación y no del contexto más amplio, cultural o social. Igualmente, Widdowson (1998: 9) sugiere que el medio más apropiado para describir la lengua es el análisis del género que “seeks not simply to reveal what linguistic forms are manifested but how they realize, make real, the conceptual and rhetorical structures, modes of thought and action, which are established as conventional for certain discourse communities”.

Por otro lado, los resultados del análisis del registro mostraron que existían muy pocas diferencias entre la lengua de la ciencia y la lengua general. Como Coffey (1984: 4-5) concluye:

“In short, register cannot be used...because there is no significant way in which the language of science differs from any other kind of language.”

Estos resultados afectaron no sólo a la lengua de la ciencia sino también a todo el conjunto de ESP y su enseñanza: “ESP is *not* a matter of teaching ‘specialised varieties’ of English. The fact that language is used for a specific purpose does *not* imply that it is a special form of the language, different in kind to other forms.” (Hutchinson y Waters, 1987: 18).

Sin embargo, debemos señalar que el análisis del registro ha cobrado más fuerza y más sentido desde la perspectiva del género. Los años 90 fueron testigo de la consolidación de la línea de investigación sobre el discurso científico iniciada por Barber en 1962 en relación a géneros específicos. Así, el análisis de frecuencias en relación con elementos gramaticales o léxicos de textos específicos y el desarrollo de las ideas que subyacen al análisis del género hacen posible establecer una relación entre los datos cuantitativos obtenidos del análisis y los rasgos discursivos del texto.

Actualmente, se están llevando a cabo estudios comparativos del discurso científico a varios niveles: contraste entre géneros, disciplinas, lenguas, etc. y en este área parece concentrarse la futura investigación. Algunos trabajos recientes en este campo establecen contrastes entre artículos de investigación y la divulgación de la ciencia, como el estudio de Gunnarsson (1993) realizado en diversas ramas de la ciencia o el estudio de Myers (1994) en el campo de la genética molecular. Gómez Morón (2001) realiza un análisis comparativo entre la utilización de citas en el artículo de investigación científico y humanístico. Otros estudios analizan las posibles diferencias y similitudes en las prácticas retóricas utilizadas por distintas culturas<sup>2</sup> en la redacción del texto científico en lengua inglesa, como el realizado por Ventola y Mauranen (1991) y Mauranen (1993a, 1993b, 2000), que contrastan el finlandés con el inglés; el efectuado por Taylor y Chen (1991), que exploran las diferencias entre el chino y el inglés; el estudio de Clyne (1987, 1991), entre el alemán y el inglés; o el de Anderson y Gunnarsson (1995), entre varias lenguas incluido el español. En España merecen destacarse las investigaciones realizadas por Moreno (1997) en la redacción del artículo de investigación, el efectuado por White y Andrew (2001) en las introducciones de los artículos de investigación biomédicos y el estudio realizado por Vázquez y del Árbol (2002) sobre las cartas al director y editoriales biomédicas, que analizan las diferencias culturales entre españoles y anglófonos.

La combinación, pues, del análisis del registro con el estudio del género en el análisis textual supone un paso adelante en la investigación del discurso científico, que no se va a limitar a la mera detección de una serie de elementos léxico-gramaticales obtenida a través del análisis del registro, sino que va más allá, pretende examinar cómo

---

<sup>2</sup> Según Kaplan (1966, 1987 y 1988) existe una estrecha relación entre cultura y lengua.

estos contenidos se integran en las convecciones lingüísticas y socio-culturales de los diversos tipos de texto.

### **1. 2. 2. Análisis retórico-funcional.**

La reacción en contra del análisis de registro a principios de 1970 “concentrated on the communicative values of discourse rather than the lexical and grammatical properties of register.” (West, 1997: 36). El análisis retórico-funcional investiga cómo las oraciones se unen dentro de un texto para darle forma: “In discourse one has to work out what concepts, or propositions, are being expressed and what functions, or illocutions, they count as” (Widdowson, 1978: 24). Esto incluye el concepto de coherencia, “the quality of being meaningful and unified” (Cook, 1989: 4), y cohesión, “links between sentences and between clauses” (Cook, 1989: 14).

Se evoluciona a un enfoque cualitativo, en el que lo que importa no es la frecuencia de las estructuras lingüísticas, sino por qué el escritor elige unas determinadas estructuras lingüísticas para dar forma al texto.

“The focus was thus on text (specifically the conceptual paragraph) rather than on the sentence, and on writer’s purpose rather than on form.” (Robinson, 1991: 24).

En la práctica, el análisis retórico-funcional tiende a concentrarse en cómo las oraciones se utilizan para realizar actos de comunicación y su objetivo es generar materiales de enseñanza basados en las funciones. En éstos ya se tienen en cuenta los conceptos de coherencia y cohesión como aspectos esenciales constitutivos del texto.

“One might usefully distinguish two kinds of ability which an English course at this level should aim at developing. The first is the ability to recognise how sentences are used in the performance of acts of communication, the ability to understand the rhetorical functioning of language in use. The second is the ability to recognize and manipulate the formal devices which are used to combine sentences to create continuous passages of prose. We might say that the first has to do with rhetorical coherence of discourse, the second with the grammatical cohesion of text.” (Allen y Widdowson 1974: 2).

La principal crítica a este enfoque es que el análisis seguía siendo fragmentario, ya que identificaba las unidades funcionales de las que el discurso está compuesto a nivel de oración/proposición, pero no indicaba cómo las funciones y oraciones/proposiciones se configuran para dar forma al texto: “We are given little idea of how these functions combine to make longer texts” (Robinson, 1981: 54).

Robinson (1991) cita entre otros los trabajos de Lackstrom, Selinker y Trimble (1973); Selinker y Trimble (1974); Selinker, Trimble y Vroman (1974); Selinker, Todd Trimble y Trimble (1976, 1978); Todd Trimble y Trimble (1982) y Trimble (1985) como representativos de los inicios del análisis retórico-gramatical.

Algunos de estos análisis han dado lugar a falsas generalizaciones, como es el caso de la definición, considerada como una de las estructuras retóricas principales en el discurso científico, y que, como señala Swales (1981a), se utiliza con frecuencia únicamente en los libros de texto de la ciencia, pero raramente en las demás variedades del discurso científico. Otros estudios han revelado un uso más frecuente de la definición en el artículo especializado que en el artículo divulgativo, donde es prácticamente inexistente.

Estudios posteriores dentro del análisis retórico-gramatical (Swales, 1981a; Tarone, Dwyer, Gillette y Icke, 1981; Trimble, 1985 y Malcom, 1987) han mostrado más atención a géneros específicos que a un registro general de la ciencia. Esta progresión gradual en la descripción lingüística se mueve, por tanto, de general a específico. A esta evolución se refería Widdowson (1979: 54), que establece una secuencia progresiva: “attention on text, textualization, and discourse”.

Las diferencias entre el trabajo de Barber (1962) y el de Tarone et al. (1981)<sup>3</sup>, elegidos respectivamente como primer y último episodio de (Swales, 1985) “ponen de manifiesto no sólo el diferente modo de entender la relación con el texto sino que nos sirven de guía hacia donde se orienta la investigación del discurso científico” (Serrano, 1999: 15).

---

<sup>3</sup> En 1998 Tarone, Dwyer, Gillette y Icke revisan y actualizan esta investigación llegando a la conclusión de que en los artículos de astrofísica la voz activa es más utilizada que la voz pasiva y que la primera persona plural “we” se utiliza en puntos estratégicos relacionados con la estructura retórica de estos artículos y dependiendo del objetivo comunicativo del autor.

### 1. 2. 3. Análisis del género

Como hemos señalado anteriormente, un avance significativo dentro del análisis del discurso lo constituye el análisis del género. Frente al análisis del registro, el énfasis se sitúa en “schematic structure of texts and on attempts to identify the functional/rhetorical roles of specific linguistic phenomena” (Bloor, 1994: 16). Este nuevo enfoque intenta aproximarse al texto en su totalidad y considera que es algo más que una colección de unidades o funciones aisladas.

Los teóricos del género, a diferencia de los teóricos del registro, que analizan los rasgos léxico-gramaticales del texto, examinan, con una metodología más integrada, cómo éstos “forman parte de las convecciones lingüísticas y socio-culturales de los textos utilizados por los profesionales y los académicos en los distintos acontecimientos comunicativos de su vida diaria, en la empresa, la universidad, el hospital, el foro, etc.” (Alcaraz Varó, 2000: 134). De esta forma, el análisis genérico constituye un modo de abarcar todos los niveles del contexto, así como también la estructura lingüística, lo cual proporciona una percepción del texto como producto y proceso.

“Genres are both ‘products’ and ‘processes’-‘systems’ and ‘performances’. Each time a text is produced so as to realize and construct a situation-type it becomes the model for another text and another situation-type. As a model, it functions like a static, finished product or a system according to which new texts can be constructed. Once the constructing begins it becomes again a dynamic process, a ‘performance’ which will inevitably change the model with which it begins. (Threadgold, 1989: 102).

Los géneros, pues, no se consideran entidades permanentes, sino “goals structuring generic activity need to be viewed as multiple rather than unified, varying with time, place, and participants.” (Prior, 1995: 55). Como Berkenkotter y Huckin (1993: 481) señalan: “Genres must do more than encapsulate intersubjective perceptions of recurring situations”, los géneros reflejan los cambios socioculturales que afectan a la comunidad discursiva: “Genres change over time in response to their users’ sociocognitive needs”. (Berkenkotter y Huckin, 1995: 4).

Esta característica dinámica de los géneros ha sido confirmada por las diversas investigaciones realizadas sobre la evolución del artículo de investigación. Huckin

(1987) analiza 350 artículos de investigación y encuentra cambios significativos en la forma en que los científicos organizan la información. Específicamente, cómo los resultados de la investigación se avanzan y se presentan en títulos, *abstracts* e introducción y cómo las secciones de métodos y procedimientos pasan a un segundo plano. Atkinson (1999) nos muestra cómo los cambios sufridos dentro del paradigma científico entre 1675 y 1975 han traído consigo cambios graduales en la estructura retórica. Valle (1997) describe la evolución del artículo científico desde la perspectiva de la comunidad discursiva; en particular, la evolución en el tema de los textos y su estructura retórica general, la audiencia a la que van dirigidos los textos y la presencia o ausencia de una ‘motivación’ explícita en el texto, es decir, el objetivo.

Así, los géneros se transforman en respuesta a las necesidades de la comunidad discursiva y los miembros que la forman. La naturaleza del género, por lo tanto, es estática y dinámica. El género constituye un modelo a seguir, ya que impone normas a la hora de formular proposiciones del mismo tipo. Al mismo tiempo, evoluciona, ya que cualquier nueva proposición inevitablemente modifica el modelo añadiendo nuevos elementos, si bien preservando los más esenciales. A este respecto Swales (2001: 12) afirma:

“The applied genre enthusiasts have played their cards well by de-emphasizing the earlier ‘systems of rules’ aspects of genre-based instruction and by coming round to stress that genres are better seen as flexible and negotiable ‘strategic resources’”.

Es cierto que, desde el punto de vista estrictamente lingüístico, género y registro han sido y son objeto de una considerable polémica sobre si deben considerarse como dos conceptos diferentes o como dos denominaciones del mismo fenómeno, que Downing (1996: 11) señala como “socially motivated variation across discourse types”. Hay autores como Ghadessy (1993) que utilizan indistintamente el término de género o registro, mientras que otros como Goatley (1994) los funden en un solo vocablo “registro-género”. El mismo Swales (1990a: 40) señala que “the relationship between genre and the longer established concept of register is not always clear”.

Incluso en la investigación realizada sobre la lengua inglesa, el concepto de género y análisis del género comprende un área bastante amplia y con frecuencia

confusa. El mismo término no tiene un significado unificado y las dos escuelas principales sobre el análisis del género, representadas por Halliday y Swales, lo conciben de forma diferente. Sin embargo, aunque difieran en sus formulaciones teóricas, “they do not appear to be in fundamental conflict and there is much to be learnt from both.” (Hoey, 2001: 8).

Bhatia (1997: 181) nos ofrece la siguiente definición de género que conjuga las diferentes tendencias actuales dentro del análisis del género.

“Genre analysis is the study of situated linguistic behaviour in institutionalized academic or professional settings, whichever way one may look at it, whether in terms of typifications of rhetorical action, as in Miller (1984), and Berkenkotter & Huckin (1995); regularities of staged, goal oriented social processes, as in Martin, Christie, & Rothery (1987) and Martin (1993); or consistency of communicative purposes, as in Swales (1990) and Bhatia (1993)”.

### **1. 2. 3. 1. El género en la escuela sistémica.**

En el seno de la escuela sistémica el género se basa en el hecho de que todo comportamiento lingüístico viene determinado por factores socio-culturales y tiene un propósito comunicativo. Los lingüistas sistémicos (Halliday, 1978; Ventola, 1984; Halliday y Hasan, 1985; Martin, 1985; Martin, 1992a, 1992b; Eggins, 1994; Ventola, 1995a; Christie y Martin, 1997; Eggins y Martin, 1997 y Martin, 2001) proponen el registro como realización del género, estando ambos inmersos en el concepto ‘contexto’. En la siguiente figura se muestra la relación entre lengua, registro y género:

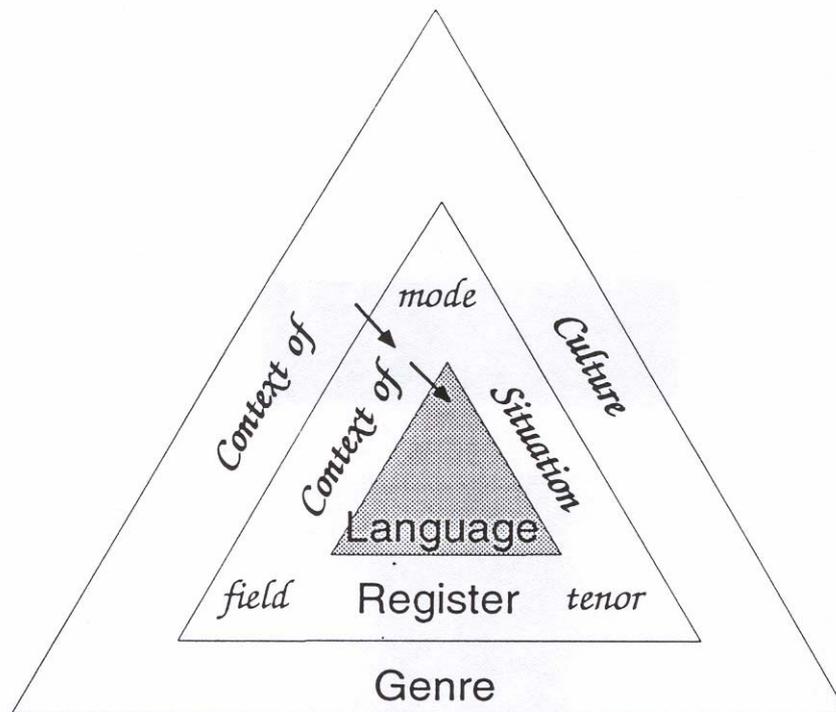


Figura 2. Lengua, género y registro. (Eggins, 1994: 34).

Consideran que el género estaría determinado por el contexto socio-cultural o de orden superior y el registro por el contexto situacional o de orden inferior. Por una parte, el análisis del género permite establecer para cada tipo de texto unos patrones formales compuestos por funciones retóricas que varían según el propósito comunicativo de aquél. Por otra parte, el concepto de registro (Halliday, 1978) y su división en categorías de campo, tenor y modo permite otro nivel de análisis distinto al del género y complementario según la sistémica australiana, ya que estudia las realizaciones concretas de cada género. Martin (1986: 17) resume de este modo el modelo:

“This model is being developed within the framework of systemic functional linguistics.... Its main peculiarity is that unlike Halliday (1978) or Halliday and Hasan (1985), it includes a fourth variable beyond the more traditional field, mode and tenor; and that unlike Gregory and Carrol (1978) it makes this fourth variable, referred to as

genre, dominate the other three. Field, mode and tenor are then referred to as register, which is treated as a semiotic system realized through language; while genre in turn is treated as a semiotic system realized through register and language”.

### **1. 2. 3. 2. El género según Swales.**

La figura representativa de la segunda escuela es Swales. Tanto Swales (1990a) como Couture (1986) formulan que ambos conceptos, género y registro, no tienen el mismo radio de acción. Mientras el género impone restricciones en el ámbito de la estructura del discurso, el registro lo hace en el ámbito lingüístico gramatical y léxico:

“Unlike register, genre can only be realized in completed texts or texts that can be projected as complete, for a genre does more than specify kinds of codes extant in a group of related texts; it specifies conditions for beginning, continuing and ending a text”. (Couture, 1986: 82).

Swales (1990a: 58) ofrece la siguiente definición de género, que pone de manifiesto el énfasis en el objetivo comunicativo y las convenciones que comparte la comunidad discursiva que lo utiliza con respecto a contenido, estructura, estilo y audiencia:

“A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constraints choice of content and style. Communicative purpose is both a privileged criterion and one that operates to keep the scope of a genre as here conceived narrowly focused on comparable rhetorical action. In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience. If all high expectations are realized, the exemplar will be viewed as prototypical by the parent discourse community”.

Swales (1981b, 1990a) analiza varios géneros dentro del inglés científico-técnico, especialmente el artículo de investigación, y otros relacionados como el

resumen científico, la tesis, los trabajos académicos y las comunicaciones orales. El autor presenta una propuesta de análisis, especialmente desarrollada para la introducción del artículo de investigación, cuya característica principal es la división del texto en fases o movimientos, *moves*. Cada una de estas fases contiene una función específica y sistemática que evidencia la organización del texto (Johnson, 1993: 203). Nwogu (1991: 114) define el término *move* de la siguiente manera:

“a text segment made up of a bundle of linguistic features (lexical meanings, propositional meanings, illocutory forces, etc.) which give the segment a uniform orientation and signal the content of discourse in it. Each “move” is taken to embody a number of “Constituent Elements” or submoves which combine to constitute information in the move”.

### **1. 2. 3. 3. El artículo de investigación.**

Uno de los géneros más estudiados ha sido el artículo de investigación, puesto que constituye la principal vía con la que cuenta el científico para trasladar a los foros profesionales, por medio de una revista especializada, “las investigaciones y descubrimientos más recientes publicados en ningún libro” Burgos (1996: 19),

Las diversas secciones del artículo de investigación científico se han analizado por medio de movimientos o *moves*, siguiendo a Swales (1981b, 1990a). Por ejemplo, la introducción ha sido investigada por Crookes (1986), Anthony (1999) y Samraj (2002). Peng (1987), Hopkins y Dudley-Evans (1988), Dudley-Evans (1994) y Peacock (2002) realizan el análisis de la sección de discusión, mientras que Thompson (1993) y Brett, (1994) analizan la sección de resultados y Salager-Meyer (1990, 1992), Carrió Pastor (1997a), Verdejo Segura (1997)<sup>4</sup> y Gea Valor (1998) investigan el *abstract*.

Swales (1990a) establece para los artículos de investigación científicos la macroestructura *IMRD* (*introduction, methods, research y discusión*), para diferenciarlos de otros géneros académicos como los artículos de revisión, la tesis, libros de texto, etc. Este esquema ha quedado validado por los diversos trabajos realizados y ha ampliado su aplicación a otras especialidades como el efectuado por

---

<sup>4</sup> Verdejo Segura (1999) aporta unas sugerentes aplicaciones metodológicas sobre el estudio del *abstract* en el aula.

Lewin y Fine (1996) en el área de psicología y sociología, el realizado por Holmes (1997) sobre historia, política y sociología, el de Nwogu (1997) sobre las ciencias médicas o el de Posteguillo (1999) sobre los artículos de investigación en informática.

En consonancia con lo anteriormente expuesto, hemos tenido en cuenta dicha macroestructura en la selección de las introducciones que van a ser analizadas en esta tesis doctoral y todas ellas pertenecen a artículos de investigación que contienen *IMRD*.

### **1. 3. Tipos de textos seleccionados para su análisis.**

Antes de adentrarnos en el concepto de cohesión y coherencia, creemos oportuno exponer las características de los dos tipos de texto seleccionados para su análisis en esta investigación, teniendo en cuenta lo anteriormente mencionado sobre registro y género:

1. Textos correspondientes a la introducción de artículos de investigación, escritos por investigadores y dirigidos a colegas y especialistas del mismo campo científico.
2. Textos ‘académicos informales’, escritos por científicos y profesores, dirigidos a una audiencia más general – tanto a alumnos como a personal docente universitario – que tiene conocimientos de la ciencia y con un objetivo pedagógico.

Utilizamos el término ‘académico informal’ para diferenciarlo de la divulgación de la ciencia, proceso por el cual un escritor hace llegar a un público no especializado y amplio el saber producido por especialistas en una disciplina científica y que ha sido investigado por diversos lingüistas como Myers (1990), Fahnestock (1993), Kelley (1993), Myers, (1994) y Fuller (1998).

Creemos que la implicación del análisis del género para la presente investigación es de gran importancia, ya que las unidades léxicas y la repetición de las mismas no se consideran como un mero recurso estilístico sino como la representación de la estructura cognitiva subyacente del texto, dotándolo de unidad y diferenciándolo de otros textos que no pertenezcan al mismo género.

El mismo Hoey (1991a: 190) sugiere que puede haber una correlación entre género y cohesión: “There may also be a link with genre; possibly bonding occurs to a greater or lesser extent in texts of a certain type or of certain level of formality”. Aunque Hoey no desarrolla esta idea en su trabajo de 1991a, sí señala que sería necesario investigar si la cantidad de conexiones establecidas por medio de repeticiones léxicas, o su ausencia, podrían estar relacionada con el tipo de texto.

El artículo de investigación y el texto ‘académico informal’ pueden considerarse dos géneros distintos. El objetivo comunicativo y la comunidad discursiva (Swales, 1990a) a la que pertenecen ambos tipos de texto son distintos aunque estén relacionados. El artículo científico pertenece a una comunidad discursiva cerrada en el sentido que sólo aquél que domine el área de la Química Analítica es capaz de comprenderlos en su totalidad, mientras que el escrito ‘académico informal’ pertenece a una comunidad discursiva más amplia que incluye a profesores de cualquier especialidad de la Ciencia Química y a alumnos con un conocimiento general de Química.

El objetivo comunicativo es claramente distinto. Como afirma Swales, el objetivo de la introducción es “to create a research space for the writer-researcher” (Swales, 1990b: 204) y pertenece al artículo de investigación cuyo propósito es trasladar a los foros científicos nacionales e internacionales los resultados de investigaciones puntuales. El objetivo del escrito ‘académico informal’ es pedagógico e ilustrativo de sencillos experimentos o narraciones para su posible utilización en clase o de lectura para alumnos.

Al hacer esta distinción, nos basamos más en el aspecto sociocultural que en el puramente lingüístico. Siguiendo la tradición de Miller (1984), Swales (1990a), Mauranen (1993a), (1993b), Freedman (1994), Dudley-Evans (1995) y Swales (2001), enfatizamos el objetivo comunicativo y la comunidad discursiva. Creemos que sería interesante investigar si “there may be some rough and ready correlation between the level at which bonds are formed and the genres of which the texts are instances (Hoey, 1991a: 92).

Por otro lado, las variables de campo, tenor y modo pertenecientes al registro nos ofrecen una visión distinta y complementaria a la hora de establecer las diferencias

entre estos dos tipos de texto. En 1976, Halliday y Hasan (1976: 23) predecían que la interpretación de los patrones de los enlaces léxicos dependía del registro:

“The concept of cohesion can therefore be usefully supplemented by that of register, since the two together effectively define a text. A text is a passage of discourse which is coherent in these two regards: it is coherent with respect to the context of situation, and therefore consistent in register; and it is coherent with respect to itself, and therefore cohesive”.

Como reitera Halliday (1994: 339), para que un texto sea coherente “it must deploy the resources of cohesion in ways that are motivated by the register of which it is an instance”.

Las variables del contexto situacional de estos dos tipos de texto varían según la siguiente tabla:

<b>Tipo de texto</b>	<b>Campo</b>	<b>Tenor</b>	<b>Modo</b>
Introducción del artículo de investigación.	Química analítica.	De especialista a especialista.	<input type="checkbox"/> Escrito. <input type="checkbox"/> Perteneciente a revistas especializadas. <input type="checkbox"/> Información visual: Representaciones gráficas con diagramas tablas, etc.
Escrito ‘académico informal’.	Química general.	De profesor universitario a profesor universitario y alumnos.	<input type="checkbox"/> Escrito. <input type="checkbox"/> Perteneciente a revistas educativas. <input type="checkbox"/> Información visual: Fotografías.

Tabla 1. Variables del contexto situacional.

Si para establecer la coherencia o textura de un texto, como señala Martin (1992b: 418), hay que tener en cuenta “consistency of register (field, mode and tenor) and genre (text structure)”, sería razonable investigar si los recursos cohesivos y las relaciones léxicas varían o, por el contrario, no existen diferencias significativas en estos dos tipos de texto, pertenecientes a dos géneros distintos y con variables contextuales diferentes.

#### **1. 4. Cohesión y coherencia**

La producción de un texto no consiste simplemente en la combinación o secuencia de oraciones; es el reflejo del proceso mental de un escritor manifestado por la elección de palabras, estructuras gramaticales, contenido y organización. En un intento de descubrir este proceso, los estudios del discurso escrito han explorado qué determina la cohesión textual o coherencia (Connor, 1984; Weisberg, 1984; Hölker, 1989; Lundquist, 1989a; Lundquist 1989b; Villemoes, 1989; Myers, 1991; Kuo, 1995 y Tanskanen, 1995;), qué clase de elementos caracteriza un buen escrito (Witte y Faigley, 1981; Yang, 1989 y Parsons, 1990 y 1991), o cómo varía el uso de recursos cohesivos según la cultura a la que pertenezca el escritor (Johns, 1984; Fernández Polo, 1995 y Connor, 1996). Entre estos estudios, aquéllos que se han realizado sobre la cohesión textual han sido principalmente cuantitativos basados en el análisis de frecuencias de los recursos cohesivos.

Aunque cada tipo de recurso cohesivo – siguiendo la teoría de cohesión de Halliday y Hasan (1976) –, ha sido analizado en diversas lenguas y tipos de texto, la **cohesión léxica**, que puede revelar información crucial sobre la elección de palabras por parte del escritor en el proceso de confección de un escrito, no ha sido suficientemente analizada con el objetivo de explorar su función organizativa dentro de un texto escrito. Como afirma Carter (1998: 85), “analysis of the kinds of textual relations and patterns realized by vocabulary remains a major lexical research goal”.

#### **1. 5. Diferencia entre cohesión y coherencia.**

El término cohesión indica el medio por el cual los textos están lingüísticamente conectados. Se puede definir como “the use of explicit linguistic devices to signal relations between sentences and parts of texts.” (Connor, 1996: 83) o “the means

available in the surface forms of the text to signal relationships that exist between sentence or clausal units in a text”. (Grabe y Kaplan, 1996: 56).

Halliday y Hasan (1976) consideran que la cohesión es una propiedad lingüística del texto que contribuye a la coherencia del mismo. Al proponer el concepto de cohesión como un factor determinante de lo que normalmente se entiende por coherencia, señalan:

“If a speaker of English hears or reads a passage of the language which is more than one sentence in length, he can normally decide without difficulty whether it forms a unified whole or is just a collection of unrelated sentences. This book is about what makes the difference between the two.

The word TEXT is used in linguistics to refer to any passage, spoken or written, of whatever length, that does form a unified whole. We know, as a general rule, whether any specimen of our own language constitutes a TEXT or not.

This suggests that there are objective factors involved – there must be certain features which are characteristic of texts and not found otherwise; and so there are. We shall attempt to identify these, in order to establish what are the properties of texts in English, and what it is that distinguishes a text from a disconnected sequence of sentences.” (Halliday y Hasan, 1976: 1).

En esta misma línea parece manifestarse Hoey (1994a: 44) cuando señala que los textos incorrectos “can be shown to arise from ‘faulty’ or missing signaling”.

Mientras que la cohesión se manifiesta en la conexión entre los elementos lingüísticos de un texto, existe al mismo tiempo una continuidad de sentido a un nivel más profundo. Este nivel tiene que ver con el contexto, la experiencia y el conocimiento que todo lector aporta al texto desde el mundo exterior. De esta forma, podemos señalar que la cohesión es un hecho objetivo y medible, mientras que la coherencia es subjetiva y es el resultado de la interacción entre el texto y el saber enciclopédico que supuestamente comparten emisor y receptor (Swales, 1990b: 189; Hoey, 1991a: 11; Hunston, 1993: 57; Eggins, 1994: 87; Connor, 1996: 84; Stubbs, 1996: 7 y Downing, 2001: 57).

La cohesión estaría íntimamente ligada con el *modo* del discurso, mientras que la coherencia lo estaría con el *campo* y el *tenor* del mismo. Es función de la cohesión “to

guide the reader towards the coherence intended by the producer.” (Thompson, 2001: 59).

### 1. 6. Diferencia entre textura y texto

Halliday y Hasan (1976: 2) definen textura como la propiedad de “being a text”; esto es lo que distingue un texto de un no-texto y deriva su textura del hecho de que funciona como una unidad con respecto a su entorno. El texto queda definido como una unidad semántica (Halliday y Hasan, 1976: 293), una unidad de sentido, oponiéndola a una unidad gramatical. Esta distinción es de gran importancia, ya que nos conduce a la pregunta de cómo esta unidad semántica se estructura.

Halliday y Hasan (1976: 9) consideran que los textos no están constituidos por oraciones; más bien se realizan o codifican por medio de éstas. Los textos y las oraciones son unidades lingüísticas diferentes – el texto es una unidad semántica, la oración una unidad gramatical. A diferencia de los textos, las unidades gramaticales consiguen la unidad por medio de la estructura gramatical. Ya que el texto no es una unidad estructural, su unidad proviene no de la estructura gramatical sino de la cohesión que lo constituye como un edificio semántico (Halliday y Hasan, 1976: 27).

### 1. 7. Concepto de vínculo (*tie*)

Un concepto importante introducido por Halliday y Hasan es el de vínculo (*tie*) o “a single instance of cohesion” (Halliday y Hasan, 1976:3). Lo cohesivo no es la presencia o no de una clase particular de elementos sino “the relationship between a cohesive item and the item it presupposed in a text” (Martin, 2001: 37); por lo tanto, la cohesión se manifiesta cuando la interpretación de un elemento del discurso depende de la del otro. En el siguiente ejemplo bien conocido de Halliday y Hasan (1976):

“Wash and core six cooking apples. Put them into a fireproof dish”

*them* en la segunda oración se refiere a *six cooking apples* y sólo podemos interpretar *them* si establecemos la relación con *six cooking apples*.

A este tipo de relación entre unidades Halliday y Hasan la denomina *tie*. Este concepto de vínculo hace posible el análisis de las propiedades cohesivas de un texto “and give a systematic account of its patterns of texture.” (Halliday y Hasan, 1976:4).

Es importante señalar que, aunque la relación se efectúe mediante el sistema léxico-gramatical de la lengua, la naturaleza del vínculo es semántica; los dos términos que establecen un vínculo están conectados por medio de una relación de significado. Tales relaciones semánticas constituyen la base para establecer la cohesión entre los mensajes de un texto (Hasan, 1985: 73).

Halliday y Hasan (1976) describen con detalle los diferentes tipos de vínculos cohesivos de la lengua inglesa – referencia, sustitución, elipsis, conjunción y cohesión léxica –, siendo su función esencial contribuir a la coherencia o en términos de Halliday y Hasan textura.

### **1. 8. Definición de cohesión léxica.**

La definición de cohesión léxica dada por Halliday y Hasan en 1976 está aún vigente. La definen como “selecting the same item twice, or selecting two that are closely related” (Halliday y Hasan, 1976:12). McCarthy (1991: 65) afirma que la cohesión léxica consiste en “exact repetition of words and the role played by certain basic semantic relations between words in creating *textuality*<sup>5</sup>, that property of text which distinguishes it from a random sequence of unconnected sentences”. Este concepto de cohesión léxica constituye el punto de partida de nuestro trabajo de investigación, ya que equipara cohesión léxica con repetición, “most cohesion serves to repeat” (Hoey, 1995: 73), e implica que es posible estudiar la cohesión léxica por medio del estudio de la repetición.

El léxico, a diferencia de las otras relaciones, es el único tipo de cohesión “that regularly forms multiple relationships ...lexical cohesion becomes the dominant mode of creating texture” (Hoey, 1991a: 10). Igualmente, Greenbaum y Quirk (1990: 434) argumentan que “lexical choice” es más frecuente a la hora de establecer “the internal unity of texts” que las unidades puramente sintácticas o gramaticales. Por lo tanto, al

---

<sup>5</sup> Ventola (1995b: 110) también utiliza el término ‘textualidad’ para referirse a como “discourse hangs together.”

estudiar la repetición como un hecho discursivo investigaremos “the way lexis patterns in text” (Hoey, 1991a: 74).

Los dos tipos de cohesión léxica, según Halliday y Hasan (1976: 279-280), son reiteración y colocación. El criterio de reiteración cubre la repetición de:

- a) una unidad léxica idéntica
- b) una unidad léxica relacionada por sinonimia o casi-sinonimia,
- c) un hiperónimo,
- d) una unidad léxica general.

Por ejemplo, en las siguientes oraciones:

“There’s a boy climbing that tree. The boy’s going to fall if he doesn’t take care”

El vínculo (*tie*) que se establece entre las dos unidades léxicas ‘boy’ se clasifica como reiteración por repetición.

Si reemplazamos la segunda oración, obtenemos diversos tipos de repetición:

2ª oración: “The lad’s going to fall if he doesn’t take care”

El vínculo se establece entre ‘boy’ y ‘lad’ por sinonimia.

2ª oración: “The child’s going to fall if he doesn’t take care”

Se establece una relación de hiperonimia, ya que ‘child’ es un término hiperónimo de ‘boy’.

2ª oración: “The idiot’s going to fall if he doesn’t take care”

El vínculo se establece mediante la unidad léxica general ‘idiot’.

El otro tipo de cohesión léxica es la colocación, que Halliday y Hasan (1976: 284) la definen como “the association of lexical items that regularly co-occur”. Incluye:

- unidades léxicas opuestas, por ejemplo ‘boy’ – ‘girl’;
- antónimos, por ejemplo, ‘like’ – ‘hate’;
- unidades léxicas que pertenecen a una misma serie ordenada, por ejemplo ‘Tuesday’ – ‘Thursday’;
- unidades léxicas que no pertenecen a una serie ordenada; por ejemplo ‘road’ – ‘rail’.

Hoey (1991a: 7) critica la denominación de colocación dada por Halliday y Hasan (1976), ya que “collocation has long been the name given to the relationship a lexical item has with items that appear with greater than random probability in its (textual) context (for example, Firth 1957a; Sinclair 1966).” Sin embargo, sí reconoce el intento de Halliday y Hasan de enfatizar el papel que el léxico realiza a la hora de estructurar y organizar el texto.

### **1. 9. Cadenas cohesivas**

Una de las aportaciones más importantes desarrolladas a partir del análisis de la cohesión en los textos propuestos por Halliday y Hasan (1976) la constituye el trabajo de Hasan (1984, 1985) que, tras investigar los enlaces cohesivos en la escritura infantil, observa que no existe una correlación entre el uso de recursos cohesivos y la coherencia. Concluye que la cohesión contribuye a la coherencia en tanto en cuanto los enlaces cohesivos forman cadenas y define el concepto de cadena cohesiva como un conjunto de unidades que se relacionan unas con otras por medio de relaciones semánticas (Hasan, 1985: 84).

Basándose en los tipos de relaciones semánticas, es posible distinguir entre dos tipos de cadenas: de identidad y de similitud. Las primeras son aquéllas que comparten los mismos referentes. Las segundas se establecen cuando distintas entidades están implicadas en procesos o descripciones paralelas. Por ejemplo, en el siguiente extracto de un cuento analizado por Hasan – extraído de Hoey (1991a: 15) –:

1. Once upon a time there was a little girl
2. and she went out for a walk
3. and she saw a lovely little teddy bear
4. and so she took it home
5. and when she got it home she washed it
6. and she had the teddy bear for many many weeks and years.

Aparecen tres cadenas de identidad:

- a) 1. girl 2. she 3. she 4. she 5. she, she 6. she
- b) 3. teddy bear 4. it 5. it, it 6. teddy bear
- c) 4. home 5. home

y también dos cadenas de similitud:

- a) 2. went out 5. got...home
- b) 4. took 6. had

Hasan (1985: 83) observa que, aunque este texto contiene un número de cadenas cohesivas, lo mismo ocurre con otro texto completamente incoherente. El texto incoherente está compuesto en realidad por una serie de oraciones aisladas, cuyas unidades léxicas presentan un alto grado de repetición:

“A cat is sitting on a fence. A fence is often made of wood. Carpenters work with wood. Wood planks can be bought from a lumber store”.

y concluye que la formación de cadenas no es un buen indicador del porqué los textos se diferencian en la coherencia. La razón es que, cuando se analiza la formación de cadenas, no se está teniendo en cuenta el texto completo sino simplemente palabras aisladas. Lo que se necesita es un método que permita captar cómo las cadenas se relacionan unas con otras para formar el mensaje. A esto Hasan (1985) lo denomina ‘interacción de cadenas’.

Como hemos visto, la presencia de una o más cadenas aisladas no garantiza la coherencia del texto. Para Hasan (1984 y 1985) el factor clave estriba en que las cadenas interactúen unas con otras. A continuación se refleja la interacción de las cadenas del texto anterior en Hoey (1991a: 16):

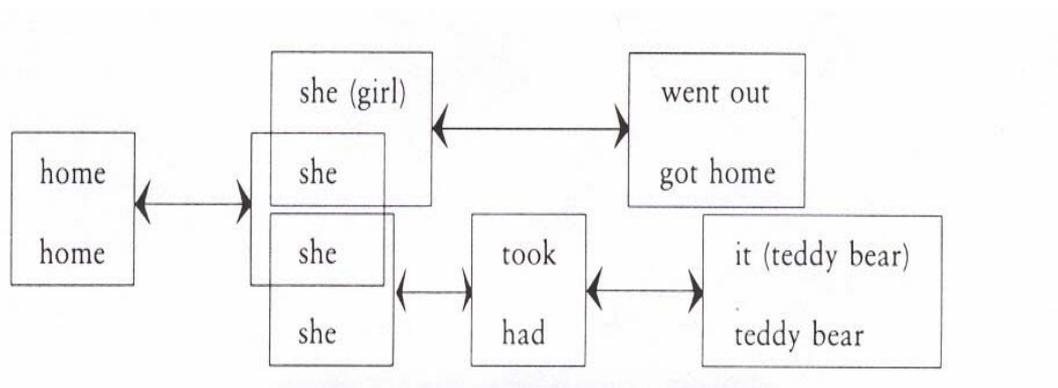


Figura 3. Interacción de cadenas.

Esta interacción entre cadenas es crucial para la metodología propuesta por Hoey (1991a), ya que la combinación de los recursos cohesivos aparece como el factor determinante para poder entender la cohesión textual, y no la presencia aislada de los mismos. Como afirma Hoey (1994b: 4):

“If we treat cohesive ties in isolation,..... we will miss three of the most important features of cohesion, namely that it is predominantly a **lexical phenomenon**, that connections may be between **non-adjacent** sentences and that such connections **cluster**”.

## 1. 10. Repetición

La función pragmática de la reiteración en el discurso se ha visto estimulada por el análisis funcional de la repetición de Winter (1974, 1977, 1979, 1994). Su interés está, más que en la clasificación de los recursos cohesivos, en la función común que éstos desempeñan en el discurso; esta función es la repetición e incluye dentro de ella la elipsis, la sustitución y la repetición léxica (Winter, 1979: 102). El argumento principal estriba en que las relaciones semánticas que unen los componentes de un texto, sean

cláusulas, oraciones o incluso grupos de oraciones, están señaladas lingüísticamente en el texto; no son inferidas simplemente por el lector.

Una forma de explorar la relación de los mensajes dentro de un texto es por medio de la repetición y el reemplazo (replacement)<sup>6</sup>, originalmente propuesto por Winter (1974, 1979). Siguiendo a Dixon (1965) y a Quirk, Greenbaum, Leech, y Svartvik (1972), Winter sugiere que una de las formas principales en la que muchos tipos de texto no-narrativos se desarrollan es por medio de la repetición constante y sistemática de elementos semánticos, que proporciona un sentido de continuidad y de reemplazo constante – la introducción de elementos nuevos en el mensaje – que genera el movimiento hacia delante del texto.

Winter establece dos categorías básicas: relaciones de comparación (*matching relations*), cuando dos cláusulas o partes del texto que se repiten están comparando y secuencias lógicas (*logical sequencing*), cuando aparecen ordenadas lógicamente y/o temporalmente.

Hoey (1991a: 19) presenta un ejemplo de Winter (1979) extraído del *New Statesman* en el que podemos observar qué elementos constituyen la repetición y el reemplazo.

Pressures built up on all sides: **his** father, a ‘moderately successful plumbing contractor’ (said *Time*) **demanded** performance. **His** mother, who left her husband in Florida and moved to Austin to be near her son, **demanded** love. Whitman could provide neither.

Las unidades léxicas en negrita, al repetirse, hacen posible que la atención se dirija a lo que se reemplaza. Así Winter (1979: 103) señala que **his mother** reemplaza a **his father**, un reemplazo que se realiza no sólo por sus posiciones sintácticas de sujeto sino también por la repetición del posesivo **his**. En segundo lugar, **love** reemplaza a **performance**, siendo la repetición del verbo **demanded** el que crea las condiciones

---

<sup>6</sup> Reemplazo no se refiere a la simple sustitución de una palabra por otra, sino a la introducción de una información nueva dentro de un contexto que ya ha sido previamente presentado.

necesarias para el reemplazo. La representación de los elementos constantes y variables sería la siguiente:

	His	father	demanded	performance
	His	mother	demanded	love
REPETICION	his		demanded	
REEMPLAZO		*****		*****
CONSTANTE	His	parent	demanded	response by him
VARIABLE		which sex		what kind of response

La principal aportación de las investigaciones de Winter – de gran influencia en su discípulo Hoey – es haber demostrado que existe una clase de unidades léxicas cuya función principal es organizar el texto así como la interrelación entre gramática, léxico y texto. Hoey (1991a: 20) resume los puntos clave:

- “1. If cohesion is to be interpreted correctly, it must be interpreted in the context of the sentences where it occurs.
2. We are more likely to arrive at a satisfactory account of how cohesion works if we concentrate on the way repetition clusters in pairs of sentences.
- 3 It is the common repeating function of much cohesion that is important, not the classificatory differences between types of cohesion.
- 4 There is informational value to repetition, in that it provides a framework for interpreting what is changed.
- 5 Relations between sentences established by repetition need not be adjacent and may be multiple”.

### 1. 11. Redes de repetición

Phillips (1985) investiga la organización léxica de un *corpus* consistente en 8 libros de texto (5 pertenecientes al área de la ciencia, 2 literarios y 1 de divulgación de la ciencia). Su objetivo no es analizar la cohesión, sino investigar la función que el léxico desempeña en la organización del texto. Tomando como unidad de análisis el capítulo, su punto de partida es la colocación y su capacidad para organizar textos tan extensos como un libro.

Con la ayuda de una metodología estadística compleja descubre que las colocaciones entre unidades léxicas frecuentemente se interrelacionan, es decir, que si una unidad léxica *a* coloca con *b*, *c* y *d*, existe una probabilidad considerable de que *b* coloque con *c* y *d* y así sucesivamente. Esto permite a Phillips (1985: 125) representar las colocaciones entre unidades léxicas como sistemas o redes de interconexiones, como muestra la siguiente figura:

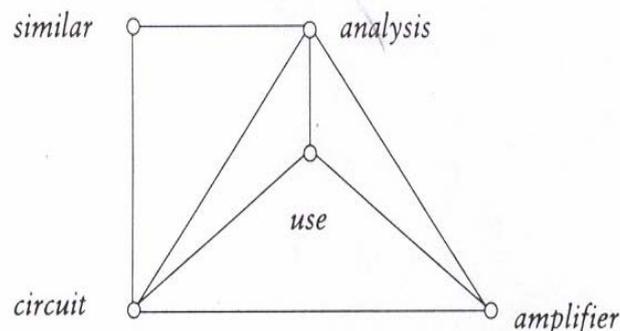


Figura 4. Red de interconexiones entre unidades léxicas.

Después de realizar las redes de interconexiones de varios capítulos, Phillips (1985) compara las redes colocacionales resultantes con las de otros capítulos del mismo libro y observa que dichas redes no son constantes, sino que varían en cada capítulo. En algunos, estas redes no aparecen, en otros sí están presentes, pero con una distribución de unidades léxicas diferente. Esto permite a Phillips demostrar que los agrupamientos léxicos aparecen de forma irregular y no necesariamente entre capítulos contiguos. La siguiente figura representa la relación entre 2 capítulos que, aunque

presentan una red colocacional de unidades léxicas distinta, comparten 3 unidades léxicas iguales (Phillips, 1985: 160):

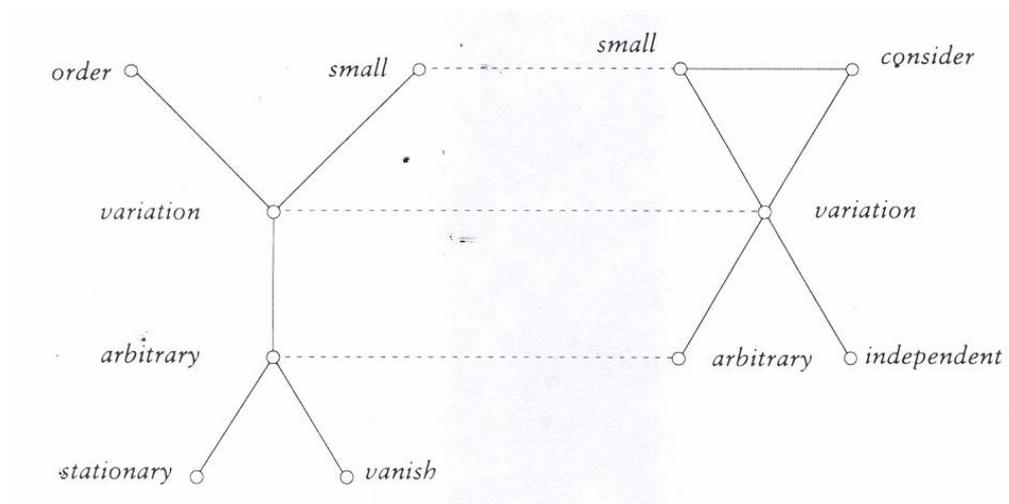


Figura 5. Interrelación de las redes colocacionales entre capítulos.

Dado que las redes colocacionales léxicas varían entre capítulos, la presencia de redes muy similares puede interpretarse como la existencia de una relación estrecha de contenido entre los capítulos.

Phillips (1985) realiza la interrelación entre capítulos según la cantidad de unidades léxicas similares que las redes colocacionales comparten. Ciertamente, es muy probable que la mayoría de los capítulos de un mismo libro presenten una o dos unidades relacionadas: “Indeed it would be surprising if this were not the case with science text viewed from a lexical standpoint.” (Phillips, 1985: 162). En cuanto al número mínimo de unidades léxicas similares que se considera necesario para establecer que los capítulos están relacionados, el criterio que Phillips (1985: 162-163) adopta es que deben existir al menos tres unidades léxicas similares:

“It was necessary to develop a notion of significant linkage in order to reveal whatever patterning might be latent in an otherwise undifferentiated mass of interrelationships .... There appeared to exist .... an interesting discontinuity in the data which proved useful as the basis for an empirical criterion. This continuity occurred between the number of times chapters displayed one .... link and the frequency of two or more such linkages.

In CMEC [uno de los libros de texto analizados], for example, there is a total of twenty-three single linkages between chapters but only eight double linkages. Beyond that, only nine times are chapters linked by more than two ties. The figures for the other science texts are comparable. A conservative criterion would, then, be to require that chapters must be connected by a minimum of three .... links in order to qualify as significantly related.

Este criterio es determinante para la metodología de Hoey (1991a), y es el que adopta en su análisis de textos. Sólo considera que las oraciones de un texto están relacionadas entre sí cuando existen por lo menos tres unidades léxicas que establecen enlace mediante la repetición. La razón es meramente práctica. Como afirma Hoey (1991a):

“if less than three repetitions are treated as establishing a significant connection, then virtually every sentence will be connected to virtually every other sentence, and that is not going to tell us anything interesting about the function of cohesion except that it is pervasive”.

Aplicando este criterio, Phillips (1985: 166) presenta la organización de los capítulos de los libros analizados. Un ejemplo de la organización de los capítulos de un libro basada en la reiteración léxica sería:

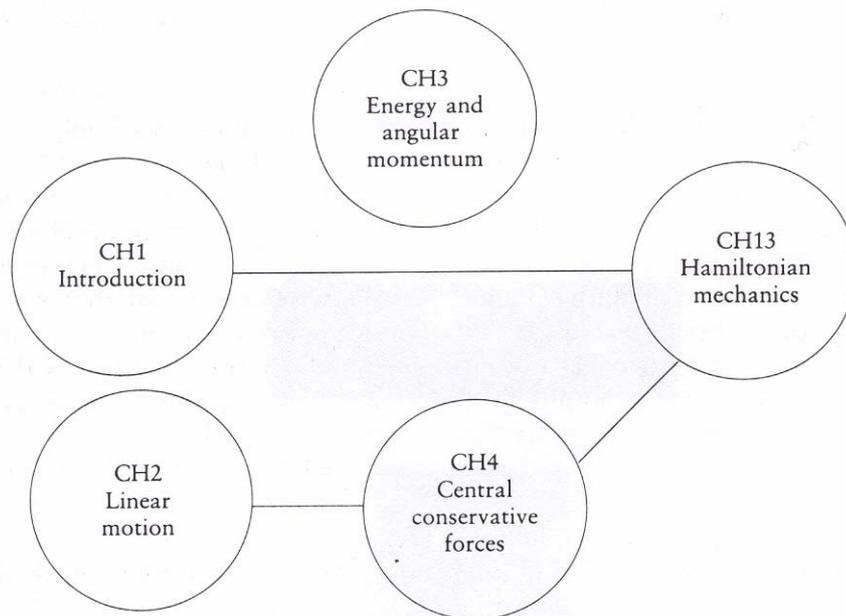


Figura 6. Relación de capítulos basada en la repetición léxica.

Phillips (1985: 165-166) considera que existen tres tipos de organización:

1. Secuencial: cuando existe una sucesión de capítulos relacionados, el primero de los cuales contiene enlaces sólo con los posteriores y el último sólo con los anteriores. En la figura podemos observar que los capítulos 2, 4 y 13 presentan una organización secuencial
2. Sinóptica: cuando dos capítulos no consecutivos están relacionados. Este tipo de organización queda representada por los capítulos 1 y 13
3. Aislada: Cuando un capítulo no establece ninguna relación con otro. El capítulo 3 es un ejemplo de este tipo de organización.

Hemos detallado la investigación de Phillips (1985) porque va a tener una gran repercusión en la metodología propuesta por Hoey (1991a) y, por tanto, en nuestra investigación y ello por varias razones:

1. El estudio del léxico permite observar el modelo de organización de un texto.
2. La repetición sistemática del léxico permite la identificación de la apertura de un tema y su cierre.
3. Las relaciones léxicas indican que los textos académicos contienen agrupaciones de repeticiones entre capítulos distantes.
4. Los capítulos que presentan un contenido común están relacionados por medio de la repetición léxica.
5. La cohesión léxica realiza una función esencial en la organización de un texto.
6. La cohesión léxica sí contribuye a la coherencia.

### **1. 12. El modelo de análisis léxico de Michael Hoey**

La función determinante del léxico en la organización textual ha sido y sigue siendo una de las principales líneas de investigación de Hoey (1985, 1988a, 1988b, 1991a, 1991b, 1991c, 1993, 1994a, 1994b, 1995, 1996a, 1996b, 1997a, 1997b, 2000a, 2001). Específicamente, el trabajo de Hoey (1991a) sobre los patrones léxicos de los textos no narrativos es la base de estudio que presenta esta tesis doctoral. Se fundamenta en la noción de que la cohesión léxica forma agrupamientos entre oraciones.

Metodológicamente, su trabajo presenta un modelo innovador de análisis en la investigación de la cohesión léxica entre oraciones. Teóricamente, su profundo análisis de la forma en que la cohesión léxica opera en los textos enfatiza la importancia de la misma sobre los otros tipos de cohesión.

La propuesta de Hoey (1991a) se encamina a armonizar tres investigaciones realizadas sobre la cohesión y expuestas anteriormente. La primera se apoya en el trabajo de Hasan (1985) y se concentra en cómo las cadenas léxicas se interrelacionan. La segunda, se apoya en el trabajo de Winter (1979) y, en particular, en la presunción de que la función fundamental de la cohesión es repetir. Finalmente, siguiendo a Phillips (1985), identifica relaciones cohesivas léxicas que aparecen a larga distancia entre oraciones. El resto del capítulo se dedicará a explicar con gran detalle la propuesta de Hoey (1991a), que es la que se propone verificar esta tesis doctoral.

### 1. 12. 1. La cohesión léxica y la organización del texto.

Hoey (1991a) enfatiza la importancia de la cohesión léxica haciendo notar que, incluso en los textos analizados por Halliday y Hasan (1976), ésta constituye el tipo dominante de cohesión; casi un 50% de los enlaces son léxicos<sup>7</sup>. Además, considera que la cohesión léxica es el único tipo de cohesión que puede establecer múltiples relaciones y que éstas, siguiendo a Phillips (1985), se realizan en forma de red. Por ejemplo, la unidad léxica **reader** (Hoey, 1991a: 246-247) establece un enlace con cada una de las apariciones de **reader** en las oraciones 1, 7, 10, 12, 14 y 16 y cada una de las apariciones de **reader** están interconectadas. Su representación gráfica sería como se muestra en la siguiente figura:

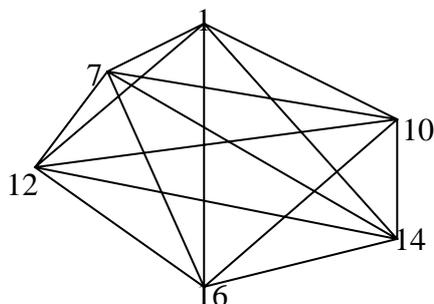


Figura 7. Interconexión de una unidad léxica. (adaptado de Hoey, 1991a: 81).

La relación entre cohesión léxica y organización del texto constituye la esencia del trabajo de Hoey (1991a) que, siguiendo la línea de Winter (1979), considera que la función común de los diversos tipos de recursos cohesivos es la repetición, función que organiza el texto mediante la creación de redes que se distribuyen a través del texto uniéndose mensajes aislados.

En opinión de Hoey (1991a) el texto, que es relativamente un objeto nuevo para la investigación lingüística, tiene dos aproximaciones. Primera, como oración, y aquí incluye a lingüistas como Pike (1967), Longacre (1968), van Dick (1972), Rumelhart (1975), Pike y Pike (1977), Thorndyke (1977) y Stein y Glenn (1979). Segunda, como un diálogo “in terms of the questions a reader might ask a text at any moment” (Hoey,

---

<sup>7</sup> No se tiene en cuenta las conjunciones, que Hoey (1991a) considera que realizan una función diferente en la organización del texto.

1991a: 30) e incluye las investigaciones realizadas por Winter (1974), Gray (1977a y 1977b), Widdowson (1979), Winter (1977), Edmonson (1981), Tadros (1981 y 1985) y Sinclair (1988).

Según Hoey, estos dos enfoques han fracasado a la hora de proporcionar los medios necesarios para que los lingüistas puedan tratar de forma apropiada la complejidad de un texto y esto se debe principalmente a que ambos están basados en principios estructurales.

Un enfoque estructural no permite la descripción válida de un texto, lo cual se debe principalmente a la característica de predecir ciertas secuencias y de suponer que ciertas secuencias son imposibles. Como Hoey (1991a: 204) señala: “It would be a daring person indeed who risked declaring which combinations of elements could not occur together in a text and which had to”.

Se necesita una nueva metáfora si lo que se pretende es tener en cuenta los modelos complejos que el léxico crea dentro del texto. Hoey (1991a: 31) propone una colección de textos ya que “such a comparison would build on the premise that texts are made up of interrelated but separate packages of information – sentences – just as a collection of texts might be”. El tipo de texto que establece relaciones con otros textos sería el artículo académico con sus notas de pie y su bibliografía.

El artículo académico nos ofrece una metáfora más apropiada para analizar las interrelaciones entre oraciones, básicamente porque cada artículo que cita a otro está en cierta manera repitiéndolo, repetición que es el eje central del enfoque de Hoey (1991a). Es más, es posible distinguir los artículos que son centrales de aquellos que son marginales; los primeros son los más citados, los segundos los menos citados.

Si los términos de la metáfora se trasladan a sus equivalentes textuales, es posible reemplazar los artículos académicos por oraciones y la colección de textos por el texto. Esto nos va a permitir distinguir entre oraciones centrales y marginales. Las oraciones centrales, como los artículos académicos centrales, “make a number of connections with other sentences” y “are germane to the development of the theme(s) of a text” (Hoey 1991a: 43), mientras que las oraciones marginales “contribute less to the development of its theme” y “show fewer signs of connection with the rest of the text.”

### **1. 12. 2. Relación entre coherencia y cohesión léxica.**

Frente a lingüistas como Widdowson (1978), de Beaugrande y Dressler (1981) y Brown y Yule (1983), que consideran coherencia y cohesión como dos aspectos separados, para Hoey (1991a) – en la línea de Winter (1979), Hasan (1985) y Phillips (1985) – ambos aspectos están relacionados.

Los resultados del análisis<sup>8</sup> de Hoey (1991a: 133) indican que los pares de oraciones conectadas por un número significativo de unidades léxicas resultan coherentes, lo cual no se aplica a los pares de oraciones elegidas al azar, de las que tan sólo menos del 50% han mostrado coherencia (Hoey, 1991a: 192). En palabras de Hoey (1991a: 126), “the co-occurrence of the requisite number of repetitions is sufficient to compel a reading of the pairs as intelligible”.

### **1. 12. 3. La oración.**

Hoey (1991a: 33) argumenta que las oraciones son “miniature packages of information”. Su status es gramatical y textual: “In so far as cohesion occurs across clause boundaries, it reveals the sentence to be a textual category; in so far as there are restrictions on the ways one may repeat within a sentence, the sentence is shown to be a grammatical category.” (Hoey, 1991a: 216).

Si tomamos la oración como una unidad completa, la pregunta siguiente es cómo la cohesión contribuye a crear relaciones entre oraciones. Hoey (1991a : 20) se basa en el trabajo de Winter (1979), que subraya el hecho de que la repetición “provides a framework for interpreting what is changed” y, por lo tanto, tiene un valor informativo.

La función de la repetición, al mostrar las relaciones entre pares de oraciones, es más evidente y efectiva si se tienen en cuenta grupos de repetición. Éstos crean relaciones entre oraciones distantes una de otra. Este enfoque proporciona no sólo el marco para una metodología que observe cómo la cohesión está relacionada con la organización del texto, sino que también pone de manifiesto la función de la cohesión al permitir percibir las relaciones entre oraciones.

---

<sup>8</sup> Análisis llevado a cabo en el texto *Masters of Political Thought*.

#### 1. 12. 4. Conceptos de enlace (*link*) y conexión (*bond*).

El sistema de análisis propuesto por Hoey (1991a) para captar las relaciones entre oraciones se basa en dos nociones clave: enlace (*link*)<sup>9</sup> y conexión (*bond*)<sup>10</sup>. El enlace aparece siempre que hay repetición de una unidad en dos oraciones, ya sea contigua o no.

El término enlace (*link*) se prefiere al término tradicional de vínculo (*tie*) utilizado por Halliday y Hasan (1976) porque “*tie* seems to imply directionality” (Hoey, 1991a: 52), mientras que enlace (*link*) indica multidireccionalidad y así permite la creación de redes entre las unidades léxicas:



Figura 8. Representación gráfica de la diferencia entre *tie* y *link*. (Hoey, 1991b: 389).

Además, Halliday y Hasan (1976) incluyen en el concepto de vínculo (*tie*) ciertos tipos de recursos cohesivos que el término enlace (*link*) no considera como tales; en particular, conectores y colocaciones.

La mayoría de los recursos cohesivos que se cuentan como enlaces (*links*) son léxicos: repetición simple y compleja, paráfrasis simple y compleja, hiperónimos e hipónimos. La repetición no léxica incluye co-referencia, sustitución y elipsis.<sup>11</sup>

<sup>9</sup> Traducimos el concepto de *link* como enlace para diferenciarlo del concepto de *tie* de Halliday y Hasan, que hemos traducido por vínculo.

<sup>10</sup> El término *bond* será traducido por conexión a lo largo de este trabajo para diferenciarlo de enlace (*link*) y vínculo (*tie*)

<sup>11</sup> Los tipos de enlace se explicarán y ejemplificarán en el siguiente capítulo donde se presenta la metodología a seguir en esta investigación.

Hoey (1991a) considera que existe una conexión (*bond*) entre oraciones siempre que exista un número suficiente de enlaces entre dos oraciones. La diferencia entre enlace (*link*) y conexión (*bond*) la establece Hoey (1991a: 91 en los siguientes términos:

“lexical items form links, and sentences sharing three or more links form bonds”.

### 1. 12. 5. Número de enlaces

Siguiendo a Phillips (1985), el criterio que Hoey (1991a: 190) adopta para establecer una conexión entre oraciones es la existencia de un mínimo de tres enlaces “because of the greater likelihood of two repetitions occurring in a pair of sentences by chance”. Ahora bien, el número de enlaces que constituye una conexión está relacionado con el tipo de texto y con el promedio de enlaces del texto: “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).

Los pares de oraciones relacionadas por conexiones tienen la característica esencial de que comparten un contenido común, están relacionadas semánticamente e incluso son coherentes. Por ejemplo, en las dos oraciones siguientes las unidades léxicas relacionadas aparecen enmarcadas y unidas por líneas:

[1] What is attempted in the following volume is to present to the reader a series of actual excerpts from the writings of the greatest political theorists of the past; selected and arranged so as to show the mutual coherence of various parts of an author’s thought and his historical relation to his predecessors or successors; and accompanied by introductory notes and intervening comments designed to assist the understanding of the meaning and importance of the doctrine quoted. [2] What, then, is the advantage which we may hope to derive from a study of the political writers of the past?

(adaptación de Hoey, 1991a: 129).

La oración primera “specifies what the writer is offering the reader”, mientras que la segunda “raises the issue of what the reader might gain from the offer.” (Hoey,

1991a: 129). Hay que destacar que las dos oraciones del ejemplo están separadas por 16 oraciones en el texto.

### 1. 12. 6. Matrices de repetición

La representación gráfica de los enlaces se realiza en forma de red (véase figura 5). Sin embargo, esta representación no es válida para reflejar con detalle las interconexiones entre oraciones. Para esto Hoey (1991a) utiliza lo que denomina como *matrices de repetición*.

Una matriz de repetición recoge todos los enlaces entre una oración y las restantes del texto. Hoey (1991a) distingue entre las matrices que detallan las unidades léxicas que establecen enlaces y las que recogen el número de enlaces entre oraciones. En nuestra tesis doctoral ambas matrices se presentan.

La matriz de repetición se construye dibujando una serie de columnas y filas, una para cada oración del texto. Las columnas se enumeran empezando con la primera oración del texto, mientras que las filas se enumeran a partir de la segunda oración del texto. Las celdas resultantes se completan con las unidades que establecen enlaces entre los pares de oraciones. La siguiente figura reproduce la matriz con el número de enlaces de las cinco primeras oraciones del texto *Master of Political Thought*, analizado por Hoey (1991a):

	1			
2	6	2		
3	2	1	3	
4	5	1	2	4
5	1	0	1	0

Figura 9. Matriz con el número de enlaces entre oraciones. (Adaptación de Hoey, 1991a: 90).

Analizar una matriz de repetición nos puede revelar aspectos importantes sobre la cohesión léxica de un texto, principalmente cuáles son aquellas oraciones que comparten un mayor número de conexiones, y a su vez puede indicarnos la densidad de conexiones a través del texto.

En la figura anterior podemos observar un área densa de enlaces entre las oraciones 1 y 2 y entre las oraciones 1 y 4 del texto, en comparación con los restantes pares de oraciones. Esto nos permite la identificación de las oraciones que comparten más conexiones en el texto, lo cual a su vez nos lleva a la clasificación de oraciones con respecto al grado de conexión. La primera clasificación que se lleva a cabo es la de *oraciones centrales y marginales*:

- Oraciones centrales: Hoey (1991a: 113) las define como “sentences that have an unusually high level of bonding from the remainder” y considera que éstas son esenciales dentro del texto.
- Oraciones marginales: Consideramos oraciones marginales “the sentences that form no bonds at all with other sentences” (Hoey 1991a:105). Es necesario señalar que el concepto de marginalidad utilizado en esta clasificación no implica que estas oraciones carezcan de importancia. Las oraciones marginales pueden considerarse como oraciones metatextuales, cuya función es determinar la intención del escritor o aclarar el contenido proposicional, por lo que el papel que desempeñan a la hora de la comprensión del texto es de gran importancia. Se consideran marginales en tanto en cuanto que éstas “do not directly contribute to that main theme” (Hoey 1991a: 105).

Es posible establecer una segunda clasificación entre *oraciones principales* o introductorias del tema y *oraciones concluyentes* o cierre del tema:

- Oraciones principales o introductorias: Hoey (1991a: 119) propone la siguiente hipótesis: “those with a high second co-ordinate are *topic-opening*”.

- Oraciones concluyentes o cierre de tema : “those with a high first co-ordinate are *topic-closing*.” (Hoey 1991a: 119).

Para identificar estos dos tipos de oraciones se calcula el número de conexiones que cada oración establece con sus predecesoras y posteriores. Así, de la matriz anterior podemos extraer la siguiente tabla:

<b>Oración</b>	<b>Antes</b>	<b>Después</b>
1	0	2
2	1	0
3	0	0
4	1	0

Tabla 2. Conexiones de oraciones con anteriores y posteriores

Según esta tabla, la oración 1 es la que está conectada con más oraciones posteriores del texto y, por lo tanto, se le considera introductoria del tema. Las oraciones 2 y 4 establecen conexiones con las oraciones anteriores y se les considera concluyentes.