JoSTrans
The Journal of Specialised Translation

www.jostrans.org · ISSN: 1740-367X

Chueca Moncayo, F. (2005). The Textual Function of Terminology in Business and Finance Discourse. *The Journal of Specialised Translation, 3*, 40-63.
https://doi.org/10.26034/cm.jostrans.2005.794

The Textual Function of Terminology in Business and Finance Discourse

Fernando J. Chueca Moncayo
University of Valladolid

ABSTRACT
This paper aims to demonstrate the significant role that specialised terms play in the creation of texture and text in Business and Finance discourse. Accordingly, we will explore the suggestion that an additional function could be assigned to terms in addition to those traditionally assigned, i.e. the representative and communicative functions (Cabré, 1999). Thus, we will refer to a textual function of terms. This proposal is applied in a contrastive analysis, the results of which are interpreted using the models of Hoey (1991) and Berber Sardinha (1997), in particular the concept of lexical cohesion.

KEYWORDS
Terminology, Discourse Analysis, Translation Studies, Contrastive Studies, Lexical Cohesion

1. Introduction

Our experience of teaching 'Terminology' has enabled us to explore in greater depth what is usually known as *terminology* or *terminology science*[i]. As a result of this teaching experience, we have noticed the significance of the representative and communicative functions of the terms (Cabré, 1999). Certainly, terms, i.e. the association of a concept and a name, fulfil a function of representation: they reveal the conceptual structure of each subject domain. In other words, they refer to what Desmet & Boutayeb (1994: 311) call *fields of knowledge* as opposed to *fields of experience*. On the one hand, the fields of knowledge are zones of vocabulary associated with scientific and technical knowledge shared by a community of experts. On the other hand, the fields of experience are also zones of vocabulary, but zones associated with the experience and culture of a certain linguistic community.

In addition, terms fulfil a communicative function, since these lexical items make it possible for experts to share the specialised knowledge of each subject field. This communicative function may be expressed in an immediate manner among the experts of a certain subject domain or in a non-immediate manner by means of what it is known as *linguistic mediators*, i.e. translators and interpreters. However, our teaching experience and the study of real communicative instances have permitted us to glimpse the possibility of a third function of terms: a textual function, which facilitate the overall organisation of specialised texts and contribute to the reader's perception of aboutness. This textual function of terms is closely linked to the referential function of language suggested by Bühler or Jakobson, since terms, like words, refer to concepts conventionally established by the specialists in each subject field, thus permitting the identification of meaning networks by means of specialised lexical items (i.e. terms). The idea of a possible additional function of

terms from a textual point of view is not new, and some researchers, such as Cabré (1999), have pointed out this line of research. But, in our opinion, this idea has not attracted enough attention. For this reason, our intention in this paper is to explore the textual role of terms within specialised source and target texts.

2. Theoretical background

Our study has a clear multidisciplinary character, since various disciplines are involved. Thus, it is difficult to establish sharp boundaries as to the range of each one of these disciplines. In any case, we will indicate some basic subjects that support our research:

1. Discourse Analysis.
2. Text Linguistics.
3. Lexicology and Terminology.
4. Corpus-based Studies.
5. Contrastive Studies.
6. Translation Studies.

3. A Proposal for linguistic analysis

We have divided this paragraph into four different sections:

1. A first section dealing with the unit of analysis for this research.
2. A second section that tackles the search for an appropriate category of analysis for the specialised texts in our corpus.
3. A third section that deals with the two basic models chosen to explore the textual function of terms: those suggested by Hoey (1991) and Berber Sardinha (1997).
4. Finally, a fourth section dealing with the corpus supporting this research.

3.1. The unit of analysis

Obviously, if we want to carry out a linguistic analysis of the texts in our corpus, we have to establish the unit of analysis beforehand. In order to do so, a brief discussion about words and terms is required, since it is essential to know what we mean by *word* and *term*.
In this section, we will undertake the study of the lexicon of a language from two different perspectives based on the criterion of specialisation:

1. On the one hand, we can think of a lexicon as composed of lexical items belonging to what it is usually known as *general or common language*.

2. On the other hand, we can think of a lexicon as composed of lexical items belonging to the so-called *special subject languages*.

These two perspectives, in turn, give rise to two different but complementary approaches: the lexicological vs. the terminological approach. In fact, it is quite an oversimplification to associate Lexicology with general or common language on the one hand, and Terminology with special subject languages on the other. The matter is much more complex but, for the purpose of this analysis, this is our methodological decision.

3.1.1. What is a word and what is a term?

As Geeraerts (1994: 2189) cautions, there is a vagueness in the term *word* that calls for a better definition. Jackson (1988: 8) distinguishes four kinds of *word*:

> We have identified orthographic words, words distinguished from each other by their spelling. Secondly, we have identified phonological words, distinguished from each other by their pronunciation. Thirdly, we have identified word-forms, which are grammatical variants. And fourthly, we have identified words as 'items of meaning', the headwords of dictionary entries, which are called lexemes.

The fourth concept is also called *lexical item* in Geeraerts' terminology.
With this classification in mind, we propose an alternative typology of these concepts. Thus, an *orthographic unit* is an item between two blank spaces. A *lexeme* is an item that has meaning, which is the concept that basically lies at the basis of the division of dictionaries into separate entries. And we restrict the use of the term *word* to designate those lexical items (or lexemes) which belong to the so-called general or common language, i.e. the language shared by the vast majority of a linguistic community.

As to the question about what a term is, two possible interpretations of the term *term* can be postulated: first, a term is the association of a concept and a linguistic label; second, a term is just the linguistic label. This applies to words as well. From a Saussurian structuralist point of view, we are in favour of the first interpretation. We cannot understand one without the other. Maybe this is an arguable matter from a cognitive point of view, but it remains quite clear to us from a structuralist point of view. Nevertheless, one should be aware, though, that the linguistic dimension is just one of the points from which terminology can be approached. So, if a term is the association of a concept and a name, does it belong to what Melby (1991: 16) calls *field systems* as opposed to *language systems*? From a structuralist point of view, terms, like words, belong to the system of language, although the pragmatic circumstances of usage of terms and words are evidently different.

That said, we define terms as those lexical items (or lexemes) which belong to what traditionally has been called special subject languages, i.e. that set of codes partially overlapping the general or common language, which can be characterised by means of a bunch of *special* features, such as the subject field, the participants, or the communicative setting (Cabré, 1993: 128). We share Cabré's belief that both the general or common language as well as the special subject languages maintain an intersecting relationship and a relationship of inclusion with regard to the system of language. For us, both kinds of codes are instantiations of the language system.

3.1.2. How can we identify terms?

Much research has been carried out on the intractable question of the identification of specialised lexical items (i.e. terms). From a linguistic viewpoint, we believe that there is no single reliable method for identifying terms in isolation. This forthright statement must not be misunderstood. What we mean by this is that words and terms do not have a life of their own. Rather, they are quite dependent on the context of situation in which they are produced.

But, what makes a lexical item become a word or a term? We agree with Pearson (1998) that it is the communicative setting that makes a lexical item be realised as one or the other. So, there are certain contexts in which terms rather than words are likely to occur. Pearson (1998: 35) distinguishes four different communicative settings:

1. Expert-expert communication (setting 1).
2. Expert to initiated communication[ii] (setting 2).
3. Relative expert to the uninitiated communication (setting 3).
4. Teacher-pupil communication (setting 4).

As Pearson (1998: 39) points out:

> What we are suggesting here is that the terminology used in settings 1, 2 and 4 is likely to be used in a precise way but the terminology used in setting 3 is used in a less rigid manner and more as a part of a general communicative situation. Consequently, we conclude that settings 1, 2 and 4 are reliable sources for potential term candidates, and that when terms are used within these contexts, we can assume that the people using the terms accept the stipulated and agreed meaning associated with these terms.

In this way, there are three important factors which contribute to the identification of terms: (1) the topic of the communication, (2) the knowledge status of the participants, and (3) the way in which those participants use terms. To become a real term, a clear idea about the reference of the term must be shared by the speakers.

Accordingly, it is a good idea for us to refine the definition of words and terms given above. Words and terms are lexical items that can be realised in discourse as either specialised (terms) or non-specialised (words) items depending on the communicative setting in which they are produced[iii]. Anyhow, terms in this sense are the most significant feature of the so-called special subject languages.

3.2. The category of analysis

Now that we have established the basic unit of analysis (i.e. terms), the next step in our proposal is to search for a suitable tool for the linguistic analysis of the source and target texts compiled in our corpus.
Some well-known tools for the linguistic analysis of a text are the cohesive devices suggested by Halliday & Hasan in 1976. When trying to define the term *cohesion*, they state (1976: 4):

> The concept of cohesion is a semantic one; it refers to relations of meaning that exist within the text, and that define it as a text. Cohesion occurs where the INTERPRETATION of some element in the discourse is dependent on that of another. The one PRESUPPOSES the other, in the sense that it cannot be effectively decoded except by recourse to it (Emphasis in original).

Cohesion is the first one of the seven standards of textuality suggested by de Beaugrande & Dressler (1981). For these researchers, cohesion refers to "the ways in which the components of the SURFACE TEXT, i.e. the actual words we hear or use, are mutually connected within a sequence" (Emphasis in original).

The two quotes above represent two different perspectives of understanding this concept and they point directly to the very core question about the debate on the nature of cohesion and its relationship with a close notion: that of coherence. Thus, we observe two approaches to the concept of cohesion:

1. Those for whom the nature of cohesion is a semantic one, although it is expressed at the lexico-grammatical level. This semantic approach makes the use of the term *coherence* unnecessary.
2. Those for whom the notion of cohesion alludes to the set of formal resources at the level of the discourse surface, which reflect the ways in which the components of the textual world, i.e. the configurations of concepts and relations which underlie the surface text, are mutually accessible and relevant (de Beaugrande & Dressler, 1981: 4).

As Hasan (1984: 185 – 186) claims,

> neither the term 'surface' nor 'deep' appears to me to be well-defined or theoretically motivated. If by surface phenomena is meant 'lexico-grammatical categories', then certainly the devices described above [reference, substitution,

> ellipsis, conjunction and lexical cohesion] are surface phenomena. But in accepting this, I do not accept the implicit assumption that therefore they have little or nothing to do with "deep" phenomena, if by deep phenomena is meant semantic. It is not possible to view language as a sign system 'a la Saussure', and to maintain that there is an irreconciliable gulf between 'surface' and 'deep' – between 'form' and 'meaning'.

We mostly agree with Hasan's statement. Accordingly, it is the first approach the one we are adopting in this article. However, cohesion is not only a type of relationship between elements within a text; it can also be used as a tool for the linguistic analysis of a text. Thus, Halliday & Hasan (1976) propose five categories of analysis based on cohesion: reference, substitution, ellipsis, conjunction and lexical cohesion. No doubt all of them are very useful, but since we are mainly dealing with lexical items that belong to open sets in the system of language, i.e. what is usually called *content or lexical words* (basically nouns, verbs, adjectives and adverbs), the focus of our research lies on *lexical cohesion*.

But what is lexical cohesion? In general terms, we can define this phenomenon as the relationship between two or more elements that belong to open sets of the system of language. In fact, just as with the other types of cohesion, lexical cohesion may also be understood as (1) a device for the creation of texture and text, a device which provides a hint for the appropriate interpretation of the text's message by the addressee; and (2) a category for the linguistic analysis of texts, which is our main interest here.
Lexical cohesion, then, is achieved by means of the selection of vocabulary items. Halliday & Hasan (1976: 274) distinguish two types of lexical cohesion: (1) Reiteration, which is defined as

> A form of lexical cohesion which involves the repetition of a lexical item, at the one end of the scale; the use of a general word to refer back to a lexical item, at the other end of the scale; and a number of things in between – the use of a synonym, near-synonym, or superordinate (Halliday & Hasan, 1976: 278).

(2) collocation, which occurs "between any pair of lexical items that stand to each other in some recognizable lexicosemantic (word meaning) relation" (Halliday & Hasan, 1976: 285) or between any pairs whose connection "depends not so much on any systematic semantic relationship as on their tendency to share the same lexical environment" (Halliday & Hasan, 1976: 286).

Accordingly, when talking about lexical cohesion, we seem to be handling three different groups of criteria to establish connections between lexical items:

1. Morphological criteria: there is a link between items by means of a simple repetition of those items, whether the repetition is partial (i.e. inflected forms) or total.

2. Semantic criteria: a semantic connection between lexical items can be recognised by means of the meaning relations of those lexical items, that is, synonymy, hyponymy, antonymy, meronymy or any other semantic relationship.
3. Syntagmatic criteria: the tendency of two lexical items to share the same linguistic environment can also be taken as a criterion to identify a link between two vocabulary items.

Given the polysemy of the term *collocation* in recent discussion, the use of this term by Halliday & Hasan must not be confused with the phraseological use of the term to refer to those "ready-made memorized combinations in written and spoken language" (Cowie, 1998: 1). Therefore, we must distinguish two different senses of the term collocation:

1. A first sense by which a collocation alludes to the tendency of two items to occur in the very same linguistic context or co-text. We can call this *collocation in a cohesive or Firthian sense*.
2. A second sense of the term by which a collocation refers to those fixed or semi-fixed expression that constitute the 'phrasicon' of a language (Gläser, 1994/1995: 45).

In this paper, we will use the first sense of the term 'collocation'. The application of the subcategory of collocation to the linguistic analysis of a text has traditionally posed serious threats to the validity of the results. Not surprisingly, some researchers such as Hasan (1984) have eliminated the subcategory of collocation from their studies. Sense relations, on the other hand, have been studied profusely and they constitute an objective basis for analysing a text. The same applies to morphological criteria (Berber Sardinha, 1997).
Therefore we have also eliminated the subcategory of collocation from our investigation, as we will illustrate. However, this does not imply a lack of interest in collocation, which is worth researching in greater detail.

3.3. Basic models of analysis

So far, we have established the unit of analysis (i.e. terms or specialised lexical items) and the tool for carrying out our task (i.e. the category of lexical cohesion, which is based both on semantic as well as morphological criteria). Therefore, the next step is to establish a method by which we can explore the textual role of terms in specialised source and target texts. In other words, we have answered the *what* and *with what* questions. It is time to answer the *how* question.

3.3.1. Hoey's model

The significance of Hoey's model (1991) is outstanding, judging from the high number of references to this work in other researchers' studies. In Hoey's model, lexical cohesion is the basic tool to accomplish Hoey's aim: the study of how lexical cohesive relations are combined in order to contribute to the organisation of the text. The motivation underlying this article is not very different from that of Hoey's, although the subject of study is not the same in both investigations. Unlike Hoey, we are *only* interested in specialised lexical items.

Hoey's intention is to describe lexical patterns within the text by means of the basic notion of repetition, which includes not only morphological repetition but also other concepts such as hyponymy, synonymy or antonymy. The starting point of Hoey's model is the comprehension of a text as a unit of interconnecting packages of information. Since the concept of *interconnecting packages of information* is quite an abstract concept, Hoey matches this notion with the traditional concept of *sentence*. In fact, a text is not a sequence of sentences in Hoey's model. In accordance with Halliday & Hasan (1976: 2), Hoey believes that a text is realised by, or encoded in, sentences.

Hoey changes the traditional notions of *tie* and *chain* for the concept of *link*. Thus, a potential lexical item establishes a cohesive relationship not only with a preceding lexical item to which it is tied, but also with all the preceding or following items in the text, with which it has any kind of cohesive relationship. The result of this connectedness is a net of cohesive links rather than a cohesive chain.

The integrative approach to lexis in Hoey's model is worth emphasising. This textual integration of lexis calls for the existence of lexical networks throughout a text that proves the cohesive and, therefore, the organisational function of lexis in texts.

3.3.2. Berber Sardinha's model

The second model used in this paper is the so-called *Link Set Median Procedure* or LSM suggested by Berber Sardinha (1997). Berber Sardinha's model is based on the work by Hoey (1991), which we have just introduced.

There are different procedures that have been used to research text segmentation such as those by Hearst (1993) or Okumura & Honda (1994) but we have chosen Berber Sardinha's model because it is at least in part a continuation of Hoey's work, which best fits the motivation underlying this study. Furthermore, this motivation coincides with that in Hoey and Berber Sardinha's models, i.e. to explore the cohesive role of lexis in the organisation of texts so as to validate or refute the textual function of terminology. The only difference between our research and

that carried out by Hoey and Berber Sardinha is the specialised reference of the lexical items examined in this paper. To be more precise, Berber Sardinha's focus is on reports, which is a type of specialised text. However, to our knowledge, he does not express explicitly the specialised character of the basic unit of analysis, although we assume that a great majority of those lexical units are terms, given the specialisation of the texts studied. In any case, Berber Sardinha's model is a bridge between Hoey's model and our own model of analysis.
Segmentation

> here means using the computer to divide a written text into acceptable parts (or segments). An acceptable segment is one which conforms to a certain standard, and in the case of the present work, it means matching the section divisions placed by the authors of the text (Berber Sardinha, 2000: 213).

So, Berber Sardinha's aim is to search for segments, or, to be more precise, lexical segments, which can be defined as "contiguous portions of written text consisting of at least two sentences held together by lexical cohesive links" (Berber Sardinha, 2000: 214).

In spite of applying Hoey´s model because it is amenable to computer treatment, Berber Sardinha´s starting point is just the opposite to that of Hoey. If Hoey's central point is text integration, Berber Sardinha's emphasis is on text segmentation. In applying Hoey's model to text segmentation, Berber Sardinha intends

> to show that integration and segmentation are two manifestations of the same phenomenon: 'all texts are about difference and sameness' (Hoey, personal communication); difference surfaces as segments, while sameness is made evident by the existence of meaningful multiple repetition (bonding)" (Berber Sardinha, 2000: 213).

In summary, we think of both models as complementary. In this way, the results achieved in this research, whether positive or not, can be considered to be more reliable, since they will have been obtained by the application of two opposite but complementary methods.

3.4. A corpus-based study

Up to this point, we have researched the *what*, *with what* and *how* questions. Finally, before showing the results of this investigation, it is our intention to answer the *where* question. In order to confirm our working hypothesis (that is, the cohesive role of terms in specialised texts and, therefore, the textual function of this kind of lexical items), we need a corpus, i.e. "a collection of pieces of language that are selected according to explicit linguistic criteria in order to be used as a sample of language" (Sinclair, 1996). Thus, we have compiled a corpus of source and target texts in Business and Finance discourse. Each text represents the so-

called *Group Chairman's Statement* of the corresponding annual reports of British and Spanish banking institutions. The terms used in the Group Chairman's Statement (both in English and Spanish) are in some way more generic than those used, for instance, in the balance sheet, the profit and loss account or the notes to the accounts, which are three of the documents included in the annual report (in English and Spanish). The advantage of analysing the Group Chairman's Statement is that it constitutes a semantic unit by itself and, whether terms are generic or specific, the fact is that a great number of the lexical items in this type of text are realised as terms under the communicative setting theory presented above.

Consequently our corpus is composed of two subcorpora:

1. A comparable subcorpus[iv], which is made up of two sets of texts from the same text type: the first one contains original English texts, the second one original Spanish texts.
2. A translation subcorpus[v], which is formed by a set of source texts in Spanish and their translations into English[vi].

In addition, our corpus can be classified[vii] as follows:

3. Sample corpus: it is made up of subtexts according to principles concerning size and location of the sample within the full text.
4. Synchronic corpus: it contains texts produced within a restricted period of time.
5. Terminological corpus: it includes texts originated within a special subject field, i.e. that of Business and Finance.
6. Bilingual corpus: it is made up of texts produced in two languages (English and Spanish), selected according to identical criteria.
7. Written corpus: it is made up entirely of written texts.
8. Professional corpus: it consists of translations carried out by professional translators.
9. Published corpus: it consists of translations that have been published and are widely available.

Now that we have established the four basic steps of our proposal for the analysis of the source and target texts in our corpus, we have to show the results we have achieved after the application of this method of analysis.

4. Results

This section deals with the contrastive data achieved after applying the method of analysis we have commented on in the previous section. Thus, we have divided this section into two subsections: (a) the first subsection (4.1) presents what we can designate as *preliminary data*, by which we mean statistical data such as the number of sentences of each text, the type/token ratio, lexical density and so on; (b) the second subsection (4.2) puts forward the results that are specific to this research, those data

that will confirm or refute the existence of a possible textual function of terminology.

4.1. Preliminary data

This section focuses on established observation. The aim of this section is to present some statistical results which can make easier to understand the statements put forward in section 4.2. below. Therefore, we are suggesting empirical evidence rather than digging into the causes for contrastive differences.

4.1.1. Number of sentences

There is not a fixed number of sentences for texts in each subcorpus. We observe a variable number, ranging from 89 sentences in one of the texts originally written in English to 11 sentences in one of the texts originally written in Spanish. As for translations, all texts in the translation subcorpus have a similar number of sentences to those source texts in Spanish from which they are derived.

4.1.2. Sentence length

In this research, we mean by *sentence length* the quantity of orthographic units which each sentence contains. Sentence length thus defined shows interesting results: on the one hand, those sentences originally written in Spanish are considerably longer than those originally written in English. The former have almost fourteen more orthographic units on average than the latter. It could also be confirmed that the Spanish preference for hypotactic structures is the main reason for the greater sentence length.

On the other hand, there is only a slight difference regarding in sentence length between sentences in the Spanish source texts and their corresponding translations into English: nevertheless, the latter texts have fewer orthographic units than the former texts (See Annex I below).

4.1.3. Type-token ratio

According to Baker (1995: 236),

> a high type-token ratio, for instance, may be interpreted as a consequence of the process of lexical simplification which has been reported as taking place in a variety of mediated communicative activities, including translation.

Both the texts originally written in English and the translations from Spanish show a higher type-token ratio than those texts originally written in Spanish, which implies a tendency of the former towards lexical simplification and, consequently, they are easier to read. Blum-Kulka &

Levenstone (1983: 119) define "lexical simplification" as "the process and/or result of making do with *less* words" (Emphasis in original). This proves to be true for the present research as data contained in Chart 2 below (See Annex II) demonstrate. This fact, in turn, could be a possible explanation for the higher type-token ratio of the texts originally written in English and the translations from Spanish into English.

4.1.4. Lexical density

Baker (1995: 237) defines *lexical density* as "the percentage of lexical as opposed to grammatical items in a given text or corpus of texts". Our research reveals a high lexical density in the source texts in English as well as the translations from Spanish, which means that these kinds of texts are more informative and less predictable than those texts originally written in Spanish, according to the postulates[viii] by Baker (1995).

4.1.5. Number of terms

After identifying and delimiting the candidates for term status in our corpus and checking that they can be found in specialised lexicographic as well as terminological reference works, we observe that those texts originally written in English have a higher percentage of terms than the Spanish texts, which is consistent with what we said in the previous subsection with respect to lexical density. In addition, the translations into English also have a higher percentage of terms than the corresponding source texts in Spanish, so, in this sense, translations come close to the tendency of texts originally written in English (See Annex II below).

4.1.6. Grammatical category of the terms

The statistical results show a clear tendency towards a nominal terminology, i.e. nouns are the most frequent grammatical category of the terms in each subcorpus of texts. Those verbs and adjectives with a terminological value are not common, and adverbs with this terminological value are very rarely used. The findings in this type of discourse (Business and Finance discourse) are consistent with those findings in other types of discourse, in which nouns are mainly used with a terminological value (see some examples in Annex III below).

4.2. Results specific to the study

In the previous subsection, we have examined statistical data that are particularly valuable for this research from a contrastive point of view. In this subsection, we will investigate those data that can, in our opinion, confirm or refute the working hypothesis of this study, that is, the cohesive role of terms and their textual function in specialised source and

target texts. The corresponding lexical cohesion matrices for each text in both subcorpora yield the following data:

4.2.1. Repetition, hyponymy, synonymy and antonymy

The data obtained in this research confirm those data achieved by Hoey (1991) and Berber Sardinha (1997): the notion of repetition is an essential feature when researching the cohesive lexical links among sentences in texts, particularly specialised texts.

All texts in both subcorpora exhibit a prevalence of morphological repetition, let it be partial or total. However, repetition is considerably higher in those texts originally written in English, which is not surprising, since this is a known rhetorical feature of texts in this language. Thus, the writers of the texts originally written in English use the term "company" to refer to the name of the banking institution most of the times. Conversely, Spanish writers tend to use a set of designations in order to make reference to the institution's name, such as "entidad", "institución" or "organización".
The source texts in Spanish and their corresponding translations into English reveal other types of cohesive relationships, such as hyponymy, synonymy and antonymy. Interestingly enough, translations into English have a higher rate of cohesive links based on these sense relations (hyponymy, synonymy, antonymy) than the English source texts.
Revealing data on this topic are displayed in table form in Annex IV below.

4.2.2. Variation in the lexical cohesive density

Lexical cohesive density refers to the number of bonds among sentences in the source and target texts, established by the number of links among the various lexical items in the texts. A particular number of links forms one bond (see below). An example of this can be found in Annex V. Lexical cohesive density is significantly lower in those texts originally written in English than those source texts in Spanish and their translations into English.

A feasible explanation for this fact can be found in the length of sentences in each component of subcorpora. In fact, we should talk about covariance rather than explanation. As we have seen in subsection 4.1.2, the sentences in the source texts in Spanish and their translations into English are longer on average than those in the texts originally written in English. Likewise, we have said earlier in this paper that the orthographic sentence[ix] is taken as the basic unit for the application of Hoey's and Berber Sardinha's models of analysis. For this reason, the following axiom can be laid down: the longer the sentence, the higher the rate of lexical links between sentences and, consequently, the higher the rate of bonds.

4.2.3. Cut-off point

As Berber Sardinha (2000: 213) points out, clusters of lexically cohesive items are arranged in a net-like rather than in a string-like fashion. As a result, the lexical links among these occurrences can be represented in a net. A significant number of lexical links among sentences is what Hoey (1991) calls a *cut-off point*. As this author cautions (1991: 91), "the cut-off point of a text is related somewhat indirectly and uncertainly to the relative length and lexical density of the sentences of the text in question". In addition, we dare to say that it is related to the text type as well, although this statement needs further investigation. However that may be, the cut-off point is subjected to those parameters and is, therefore, variable in nature.

Hoey's cut-off point for his research is three links. Thus, three links form a bond. Since the cut-off is variable in nature and is dependent on factors such as sentence length and lexical density, and bearing in mind that we are only interested in specialised lexical items (which is another restrictive factor) the cut-off point in our research is two lexical links. If we intend to explore the cohesive role of terms and their contribution to the textual organisation, we must compute those significant cohesive links above the average. A cut-off point for one link is irrelevant, because that means a 100% of the lexical links. A cut-off point for three links implies 25% of all the links, which is a low cut-off point. A cut-off point for two lexical links implies over 50% of links, which can be considered to be an appropriate cut-off point for this text type.

4.2.4. Number and distribution of links

The number and distribution of lexical links among sentences is highly significant in both subcorpora (comparable and translation subcorpora), since a net of links can be determined, not only between contiguous sentences but also between distant sentences in the text.

4.2.5. Segmenting texts

Where applicable, we performed the LSM procedure proposed by Berber Sardinha (1997). The LSM is a procedure which was devised in order to identify lexical segments in large computerised corpora. He defines (2000: 214) "lexical segment" as "contiguous portions of written text consisting of at least two sentences held together by lexical cohesive links". The basic tasks of this procedure involve (1) assessing the similarity between contiguous sentences, and (2) assessing the dissimilarity between contiguous sentences. As Berber Sardinha points out (2000: 226), in order to evaluate the performance of segmentation, two measures commonly found in the Information Retrieval literature are used: recall

and precision. In this research, recall scores refer to the total boundaries correctly inserted by LSM divided by the number of reference segments, i.e. existing orthographic section boundaries determined by the author of the text. Precision, in turn, is the number of section headings divided by the total number inserted by LSM. Thus, a perfect score on recall indicates that the procedure has identified all of the reference segments in the texts, while a perfect score on precision shows that the procedure has only inserted segment boundaries that matched reference segments.

For this particular research, the LSM segmentation procedure achieved over 59% recall and 48% precision, which is a noteworthy performance.
On the lines of what we have said so far, the application of the above-mentioned proposal of analysis confirms our basic working hypothesis, i.e. terminology as a set of specialised vocabulary items plays a significant role as a cohesive element in the specialised texts we have researched in this investigation (Business and Finance discourse). Consequently, terminology makes an important contribution to the creation of texture[x] and, therefore, to the whole organisation of the text.

We believe that all we have said so far verify our working hypothesis,

The data gathered in this paper verify the results achieved by Hoey and Berber Sardinha. Therefore, the formal repetition of lexical items, whether partial or total, constitutes the major resource for the creation of cohesion in the language samples studied here, although we should not forget other types of resources such as hyponymy, synonymy or antonymy. Accordingly, there is a remarkable trend in texts originally written in English towards repetition. Despite the fact that repetition is also the most important resource of cohesion in texts originally written in Spanish and their corresponding translations into English, these kinds of texts show a relatively high percentage of usage of other types of meaning relationships such as synonymy, which is another rhetorical difference between English and Spanish texts: writers in Spanish tend to favour lexical variation.

5. Conclusion

In spite of the relative success of the work done here, the investigation poses a number of problems. Some are related to the decision to choose the orthographic sentence as the starting point for the analysis of specialised texts. We decided to opt for the orthographic sentence because of the apparent objectivity of this unit, although, as we could have noticed, the level of performance of our method of analysis varies according to the length of sentences, since the counting of bonds is based on the links between sentences.

Nevertheless, we have accomplished the aim stated at the beginning of this study. Thus, we have explored the cohesive role of specialised terms in Business and Finance discourse and the consequences this has for a third function of terminology: a textual function. It is worth noting here that we have concentrated on just one dimension of Terminology, the linguistic dimension, which does not mean it is the most important one or it is the only way of looking at terminology: it is just one of the three ways of approaching Terminology (linguistic, cognitive and communicative; Sager: 1990).

Certainly, the contrastive work done so far in this paper has significant implications for two subject studies:

1. Technical writing in a second language.
2. Translation Studies.

The linguistic analysis of the comparable and translation subcorpora in this investigation has revealed some contrastive differences and similarities as to the linguistic behaviour of terms in specialised texts, whether they were originally written in the L_1 or translated into the L_2. It is therefore worth bearing in mind all those differences and similarities when writing a technical text or a translation if we want our readers to grasp the intended message of the text.

ANNEX I

Subcorpora	Text number	Number of sentences	Average sentence length
Texts originally written in English	Text 1	53	22,2
	Text 2	41	20,1
	Text 3	89	18,9
	Text 4	42	21,8
	Text 5	21	25,6
Total average sentence length			**21,7**
Texts originally written in Spanish	Text 1	64	34,2
	Text 2	59	33
	Text 3	24	29,3
	Text 4	39	37
	Text 5	11	44,

ANNEX II

Subcorpora	Text number	Number of orthographic units	Number of terms
Texts originally written in English	Text 1	1174	210
	Text 2	860	147
	Text 3	1677	254
	Text 4	916	202
	Text 5	538	120
Total amount of orthographic units and terms		**5165**	**933**
Orthographic units and terms on average per text		**1033**	**186,6**
Percentage of terms			**18,06**
Texts originally written in Spanish	Text 1	2188	351
	Text 2	1946	359
	Text 3	704	121
	Text 4	1440	222
	Text 5	490	86
Total amount of orthographic units and terms		**6768**	**1139**
Orthographic units and terms on average per text		**1353,6**	**227,8**
Percentage of terms			**16,82**
Translated texts into English	Text 1	1644	305
	Text 2	1789	360
	Text 3	628	122
	Text 4	1325	226
	Text 5	446	87
Total amount of orthographic units and terms		**5832**	**1100**
Orthographic units and terms on average per text		**1166,4**	**220**
Percentage of terms			**18,86**

Chart 2: Number of terms

ANNEX III

Term candidate	**Term**	**Sentence number**
Aim	Aim	1
Successful company	Company	1
Measured	Measured	1
Success	---	1
Value	Value	1
Aim	Aim	2
Customers	Customers	2
Financial services	Financial services	2
Generation	---	3
Measure	Measure	3
Corporate management performance	Corporate management performance	3
Challenging objective	Objective	4
Company	Company	4
Current performance	Performance	5
Companies	Companies	6
Value creation	Value	6
Customer	Customer	6

Chart 3: Term candidates, terms and excluded items

ANNEX IV

Subcorpora	Text number	Repetition	Hyponymy	Synonymy	Antonymy
Texts originally written in English	Text 1	92,02	5,14	1,95	0,88
	Text 2	85,17	8,27	4,13	2,41
	Text 3	88,33	8,06	2,57	1,02
	Text 4	84,24	5,25	7,99	2,51
	Text 5	94,47	1,84	1,84	1,84
Percentage on average		**88,84**	**5,71**	**3,69**	**1,73**
Texts originally written in Spanish	Text 1	94,85	3,13	1,25	0,76
	Text 2	93,71	3,22	1,84	1,21
	Text 3	75,96	17,82	2,32	3,87
	Text 4	82,12	9,14	5,19	3,53
	Text 5	65,38	24,30	5,12	5,12
Percentage on average		**82,40**	**11,52**	**5,26**	**2,89**
Translated texts into English	Text 1	93,29	3,72	2,15	0,81
	Text 2	93,17	2,78	3,41	0,62
	Text 3	58,51	12,59	19,25	9,62
	Text 4	85,62	9,71	3,03	1,61
	Text 5	55,40	33,78	4,05	6,75
Percentage on average		**77,13**	**12,51**	**6,37**	**3,88**

Chart 4: Repetition, hyponymy, synonymy and antonymy

ANNEX V

Term	Sentence number	Term	Sentence number	Type of link
Aim	1	- Aim - Objective	2 4	- Total repetition (TR) - Synonymy
Company	1	- Company - Companies	4 6	- TR - Partial repetition (PR)
Financial services industry	1	Financial services	2	PR
Measured	1	Measure	3	PR
Value	1	Value	6	TR
Aim	2	- Aim - Objective	1 4	- TR - Synonymy
Customers	2	Customer	6	PR
Financial services	2	Financial services industry	1	PR
Measure	3	Measured	1	PR
Corporate management performance	3	Performance	5	PR
Objective	4	- Aim - Aim	1 2	- Synonymy - Synonymy
Company	4	- Company - Companies	1 6	- TR - PR
Performance	5	Corporate management performance	3	PR
Companies	6	- Company - Company	1 4	- TR - TR
Value	6	Value	1	TR
Customer	6	Customers	2	PR

Chart 5: Example of lexical cohesive density identification.

Sentence number	**(1)**				
(2)	2	**(2)**			
(3)	1	0	**(3)**		
(4)	2	1	0	**(4)**	
(5)	0	0	1	0	**(5)**
(6)	2	1	0	1	0
	(1)	**(2)**	**(3)**	**(4)**	**(5)**

Chart 6: Number of links for terms in chart 2 (Lexical cohesive matrix).

REFERENCES

Baker, Mona (1995). "Corpora in Translation Studies: An Overview and Some Suggestions for Future Research." *Target* 7 (2)., 223 – 243.

Baker, Mona (1996). "Corpus-based Translation Studies: The Challenges that Lie Ahead." Somers, Harold (1996). *Terminology, LSP and Translation*. Amsterdam / Philadelphia: John Benjamins, 175 – 186.

Bell, Roger (1991). *Translation and Translating.* London: Longman.

Berber Sardinha, Anthony P. (1997). *Automatic Identification of Segments in Written Text*. Unpublished PhD Thesis, English Dept., University of Liverpool, http://www.theses.f2s.com/phd.htm (consulted 20.05.2001)

Berber Sardinha, Anthony P. (2000). "Lexical Segments in Text." Scott, Mike & Thompson, Geoff. (Eds) (2000). *Patterns of Text. In Honour of Michael Hoey*. Amsterdam / Philadelphia: John Benjamins, 213 – 238.

Blum-Kulka, Shoshana & Levenstone, Eddie A. (1983). "Universals of Lexical Simplification." Claus Faerch & Kasper, Gabriele (Eds) (1983). *Strategies in IL Communiccation*. London: Longman. 119 – 139.

Bravo Gozalo, José María & Fernández Nistal, Purificación (1998). "La lingüística del corpus, las nuevas tecnologías de la información y los Estudios de Traducción en la década de 1990." Fernández Nistal, Purificación & Bravo Gozalo, José María (Eds) (1998). *La traducción: orientaciones lingüísticas y culturales*. Valladolid: Universidad de Valladolid, 205 – 257.

Brown, Gillian & Yule, George (1983). *Discourse Analysis.* Cambridge: Cambridge University Press.

Cabré Castellví, María Teresa (1993). *La Terminología: Teoría, Metodología, Aplicaciones.* Barcelona: Editorial Antártida / Empúries.

Cabré Castellví, María Teresa (1999). *La terminología: representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos.* Barcelona: Institut Universitari de Lingüística Aplicada.

Connor, Ulla (1996). *Contrastive Rhetoric.* Cambridge: Cambridge University Press.

Cowie, Anthony P. (1998). "Introduction." Cowie, Anthony P. (Ed) (1998). *Phraseology. Theory, Analysis and Applications.* Oxford: Clarendon Press, 1 – 20.

De Beaugrande, Robert & Dressler, Wolfgang U. (1981). *Introduction to Text Linguistics.* London: Longman.

Desmet, Isabel & Boutayeb, Sammy (1994). "Terms and Words: Propositions for Terminology." *Terminology*, 1 (2), 303 – 325.

Galinski, Christian & Budin, Gerhard (1989). "Terminology, Knowledge Theory and Language Industries." *L'Actualité Terminologique / Terminology Update*, 22 (3), 3 – 4.

Geeraerts, Dirk (1994). "Lexicology." Asher, R.E. & Simpson, M.Y.(Eds) (1994). *The Encyclopaedia of Language and Linguistics.* Oxford: Pergamon, vol. 4, 2189 - 2192.

Gentzler, Edwin (1993). *Contemporary Translation Theories.* London / New York: Routledge.

Gläser, Rosemarie (1994 / 1995). "Relations between Phraseology and Terminology with Special Reference to English." *ALFA*, 7 / 8, 41 –60.

Halliday, M.A.K. & Hasan, Ruqaiya (1976). *Cohesion in English.* London: Longman.

Hasan, Ruqaiya (1984). "Coherence and Cohesive Harmony." Flood, James (Ed) (1984). *Understanding Reading Comprehension: Cognition, Language and Structure of Prose.* Newark: International Reading Association, 181 – 219.

Hatim, Basil & Mason, Ian (1989) *Discourse and the Translator.* London: Longman.

Hearst, Marti A. (1993). "TextTiling: A Quantitative Approach to Discourse Segmentation." Project Sequoia Technical Report 24/93. Berkeley: University of California.

Hoey, Michael (1991). *Patterns of Lexis in Text.* Oxford: Oxford University Press.

Jackson, Howard (1988). *Words and their Meaning.* New York: Longman.

L'Homme, Marie-Claude (2000). "Understanding Specialized Lexical Combinations." *Terminology*, 6 (1), 89 – 110.

Laviosa, Sara (1997). "How Comparable Can 'Comparable Corpora' Be?" *Meta*, 9, (2), 289 – 319.

Laviosa, Sara (1998). "The Corpus-based Approach: A New Paradigm in Translation Studies." *Meta*, 43 (4), 474 – 479.

Melby, Alan (1991). "Des causes et des effects de l'asymétrie partielle des réseaux sémantiques liés aux langues naturelles." *Cahiers de lexicologie*, 58 (1), 6 – 43.

Okumura, Manabu & Honda, Takeo (1994). "Word Sense Disambiguation and Text Segmentation Based on Lexical Cohesion." *Proceedings of the Fifteenth International Conference on Computational Linguistics*, 755 – 761.

Pearson, Jennifer (1998). *Terms in Context.* Amsterdam / Philadelphia: John Benjamins.

Phillips, Martin (1985). *Aspects of Text Structure.* Amsterdam: North-Holland.

Roberts, Roda (1996). "Text and Translation." Fernández Nistal, Purificación & Bravo Gozalo, José María (Eds) (1996). *A Spectrum of Translation Studies.* Valladolid: S.A.E., 37 – 53.

Sager, Juan Carlos (1990). *A Practical Course in Terminology Processing.* Amsterdam / Philadelphia: John Benjamins.

Sager, Juan Carlos; Dungworth, David & McDonald, Peter F. (1980). *English Special Languages*. Wiesbaden: Brandstetter.

Sinclair, John (1996). *EAGLES, Preliminary Recommendations on Corpus Typology*, http://www.ilc.pi.cnr.it/EAGLES96/corpustyp/node5.html (Consulted 30.05.2001).

Vázquez Ayora, Gerardo (1977). *Introducción a la traductología: curso básico de traducción.* Washington, D.C.: Georgetown University Press.

Fernando J. Chueca Moncayo is currently a Reader in English and Translation at the Faculty of Translation and Interpreting (University of Valladolid). He received his PhD for a dissertation on Terminology in 2002: *The Cohesive Role of Specialised Terms in Business and Finance Discourse*. His main research areas are terminology, special subject languages and translation studies.
E-mail: chuecafj@lia.uva.es

[i] The designation *terminology science* has been coined by Galinsky & Budin (1989: 3).
[ii] According to Pearson (1998: 37), "frequently, experts working within a subject domain are called upon to communicate with others in their field who, while they have some knowledge of the field, do not have the same level of expertise. They may be students of a particular discipline, as in the case of advanced students in third level institutions. They may be people working within the same area but with a different training background, e.g. engineers and technicians, medical specialists and general practitioners."
[iii] According to what has been said so far, we think it is more appropriate to say that terms are *used* in special subject languages rather than that they *belong to* special subject languages. For practical purposes the difference is slight, but we want our postulates to remain clear from the very outset.
[iv] We use here the terminology proposed by Bravo Gozalo & Fernández Nistal (1998: 225): "Comparable corpora are composed of texts originally written in two or more languages, which show a set of similarities in spite of not being translations" (We have translated this quote from Spanish into English).
[v] Once again, we are using Bravo Gozalo & Fernández Nistal's terminology*:* "This term [translation corpus] is usually used for naming those corpora that are composed of a set of source texts and the subsequent translated texts" (We have translated this quote from Spanish into English).
[vi] Up to now, we have not used translations from English into Spanish, because they are not available for the text type we are investigating. For this reason, we have supplemented our translation corpus with a "comparable corpus"
[vii] This classification is based on the typology suggested by Laviosa (1997).
[viii] As Baker (1995: 237) suggests, "lexical words are less predictable than grammatical words because they belong to open-ended categories: there are thousands and thousands of them".
[ix] According to Hoey (1991: 215), "a sentence is defined solely in terms of a commencing capital letter and a concluding full-stop".
[x] Halliday & Hasan (1976:2) point out that "the concept of TEXTURE is entirely appropriate to express the property of 'being a text'. A text has texture, and this is what distinguishes it from something that is not a text. It derives this texture from the fact that it functions as a unity with respect to its environment" (Emphasis in original).

Biographical note

Mr. Fernando J. Chueca Moncayo is currently a Reader in English and Translation at the Faculty of Translation and Interpreting (University of Valladolid). He received his PhD for a dissertation on Terminology in 2002: *The Cohesive Role of Specialised Terms in Business and Finance Discourse*. His main research areas are: Terminology, Special Subject Languages and Translation Studies.
E-mail: chuecafj@lia.uva.es