

Document downloaded from:

<http://hdl.handle.net/10251/35312>

This paper must be cited as:

Soler Monreal, C.; Carbonell Olivares, MS.; Gil Salom, ML. (2011). A contrastive study of the rhetorical organisation of English and Spanish PhD thesis introductions. *English for Specific Purposes*. 30(1):4-17. doi:10.1016/j.esp.2010.04.005.



The final publication is available at

<http://dx.doi.org/10.1016/j.esp.2010.04.005>

Copyright Elsevier

## A contrastive study of the rhetorical organisation of English and Spanish PhD thesis introductions

CARMEN SOLER-MONREAL  
MARÍA CARBONELL-OLIVARES

LUZ GIL-SALOM  
*Departamento de Lingüística Aplicada*  
*Escuela Técnica Superior de Informática Aplicada*  
*Universidad Politécnica de Valencia*

*Camino de Vera s/n*

*46022- Valencia*

*Spain*

*Fax: 34 963877539*

[csoler@idm.upv.es](mailto:csoler@idm.upv.es)  
[mscarbon@idm.upv.es](mailto:mscarbon@idm.upv.es)  
[mlgil@idm.upv.es](mailto:mlgil@idm.upv.es)

### Abstract

This paper presents an analysis of the introductory sections of a corpus of 20 doctoral theses on computing written in Spanish and in English. Our aim was to ascertain whether the theses, produced within the same scientific-technological area but by authors from different cultural and linguistic backgrounds, employed the same rhetorical strategies to introduce the work presented. The analysis follows the Swalesian approach and is based on a move/step/sub-step model proposed for PhD introductions in Spanish (Carbonell-Olivares, Gil-Salom, & Soler-Monreal, 2009). The Spanish academic conventions appear to be that move 1 (M1-*Establishing the Territory*) and move 3 (M3-*Occupying the Niche*) are obligatory moves in PhD thesis introductions in Spanish, while move 2 (M2-*Establishing the Niche*) is optional. The structure of English thesis introductions reveals that they conform more closely to the M1-M2-M3 arrangement. Moreover, combinations of moves and patterns, cyclicity and embedding make their organisation more complex. The step analysis suggests that introductions in both languages rely mainly on the presentation of background information and the work carried out. However, the English introductions tend to stress the writer's own work, its originality and its contribution to the field of study. They also present more embedding and overlapping of steps and sub-steps than the Spanish texts.

**Keywords:** Contrastive rhetoric; Intercultural rhetoric; Genre analysis; Doctoral thesis; Introduction; Academic writing; Computing

### 1. Introduction

Contrastive rhetoric (CR) started as linguistic text analysis which aimed to identify problems in essays written by English as a Second Language (ESL) students in university classes due to the interference caused by cultural and linguistic conventions of the writer's first language (Connor, 1996; Kaplan, 1966). The approach was both theoretically-based and pedagogically-oriented. More recently, CR has been re-framed as intercultural rhetoric (Connor, 2004) and refocused on writing for specific purposes

(Connor, 2008, p. 303). The analysis of specific purpose genres, such as research articles, research reports, grant proposals, texts for professional purposes and theses, is a growing area of research and pedagogical endeavour. In addition the field has also begun to include the analysis of the social situation of writing (Connor, 2008, p. 3) and has benefited from a variety of approaches, particularly those from discourse-based, socio-cognitive and ethnographic fields.

New directions in CR focus on the processes that lead to the final written products and describe the complexities of the cultural, social, situational and contextual factors affecting writing (Connor, 2004, p. 292; Connor, 2008, p. 304). Although much research has been carried out to compare texts written in English by non-native and native writers, recent studies also compare different varieties of a language (e.g. Ädel, 2008; Pak & Acevedo, 2008) and different languages (e.g. Árvay & Tankó, 2004; Burgess, 2002; Lee, 2000; Loukianenko Wolfe, 2008; Martín-Martín, 2003; Moreno, 1997; Taylor & Chen, 1991; Suárez & Moreno, 2008; Wang, 2008). In these cases, the corpora analysed are either translations or comparable corpora. The assumption is that there is some similarity between cross-linguistic aspects but that different sociocultural and socio-rhetorical parameters directly influence the way arguments and ideas are organised and expressed. Some studies focus on the dominant discursive and cultural features of the texts investigated, particularly those dealing with eastern and northern European languages (e.g. Ahmad, 1997; Duszak, 1997; Gnutzmann & Oldenburg, 1991; Melander, Swales, & Frederickson, 1997; Taylor & Chen, 1991). Others follow genre-oriented approaches that highlight the discursal and rhetorical patterns of the texts under comparison and consider the role of the writer in the discourse community and the expectations of that community (e.g. Burgess, 2002; Feng, 2008; Lee, 2000; Loukianenko Wolfe, 2008; Martín-Martín, 2003; Martín-Martín, & Burgess, 2004; Moreno, 1997, 1998, 2004; Suárez & Moreno, 2008; Wang, 2008; Yakhontova, 2002).

In the context of genre analysis, special attention has been paid to the organisational patterns of introductory sections of English research articles (RAs) and to PhD theses. Swales's (1990) *Create a Research Space* (CARS) model for RA introductions has been validated by a number of descriptions of RA introductions written in English (e.g. Bhatia, 1997; Nwogu, 1990; Paltridge, 1994). However, other studies have pointed to the necessity of considering cyclicity (Crookes, 1986), embedding (Samraj, 2002) and new steps in the CARS model to effectively describe the

rhetorical organisation of the texts analysed (Anthony, 1999). The Swalesian framework of analysis has also been used as a reference in studies of RAs from different language groups (Árvey & Tankó, 2004; Burgess, 2002; Lee, 2000; Yakhontova, 2002). As regards English and Spanish, Burgess (2002) and Martín-Martín (2003) have investigated RA introductions and abstracts, respectively.

A number of studies on PhD theses written in English have described their overall organisation (e.g. Paltridge, 2002; Thompson, 2001), as well as specific features, such as metatextual references (Bunton, 1999), stance (Charles, 2003), modal verbs and citation practices (Thompson, 2001, 2005). Other studies have followed the Swalesian approach to analyse particular sections or chapters (e.g. Bunton, 2002, 2005; Kwan, 2006; Ridley, 2000). As for PhD thesis introductions, Bunton (2002) posited a model that showed a greater number of steps than Swales's. According to Swales (2004), this is because of the different nature and extent of the PhD thesis and the RA. Cross-cultural studies on PhD theses (e.g. Cooley & Lewkowicz, 1997; LoCastro, 2008) have investigated the contexts of both the situations and cultures of doctoral research work, comparing writings subject to different traditions and notions of what constitutes an acceptable thesis in different countries.

However, we have not found any genre-based studies drawn from a comparable corpora of PhD theses written in English and in Spanish. To our knowledge our study is the first to compare PhD theses in these two languages. Like much other work in genre analysis, this paper focuses on the introduction section and follows the Swalesian approach. In it we examine the rhetorical structure of Spanish and English PhD thesis introductions in the field of computing from a comparative point of view.

Computing is a relatively recent field of knowledge which was initiated in the Anglo-American scientific community and then exported to the rest of the world. Computing curricula in Spanish universities are founded on this tradition. In addition, the internationalisation of scholarship and the dominance of English as the language of science, propitiate the use of standardised Anglo-American patterns. Although it seems natural that text production will share certain features that go beyond linguistic and ethnic frontiers, cultural differences in communication strategies are also to be expected. Our purpose, therefore, is to identify the similarities and differences in the strategies adopted in both languages. From a pedagogical standpoint studies of this kind may help students to decide what rhetorical patterns to choose to communicate

effectively in their disciplinary field of study and the language in which they are writing.

## 2. Method

A corpus of 10 PhD Spanish theses and 10 PhD theses in English on computing was established. As the objective of the study was to conduct a detailed analysis of the texts, this was deemed an adequate number to work with. The English texts were randomly selected from the Internet. Each thesis had to satisfy three criteria: access is free; it was written within the field of computing and was submitted to an Anglophone institution. The texts found were defended between 1991 and 2005 at American, Australian, British and Canadian universities. It is not always clear whether the authors were native English writers or not, however they are assumed to have produced texts of the same standard as those of native English writers since their theses were supervised and/or assessed by English-speaking academics. The Spanish corpus was selected in the order in which it appeared in the online library, ProQuest Information and Learning, where theses and dissertations defended at the UPV (Universidad Politécnica de Valencia) are published. The theses were submitted between 2000 and 2003. Both corpora belong to the sub-fields of computation, computer engineering, and programming languages. The difference in the submission years of the theses in the two corpora is not expected to have a relevant impact on the results. More important aspects, such as the field of knowledge, discourse community and audience are indeed comparable, making the corpora suitable for our research purposes. As Moreno (2008, p. 29) explains, comparable corpora are equivalent to the extent that the text exemplars contained in them may be considered similar in all relevant contextual factors. In order to assess the comparability of the corpora used in the study, informants among the members of the PhD faculty from the School of Computer Engineering (UPV) were consulted. They corroborated that the aspects mentioned above, as well as the proportion of applied vs. theoretical theses and the submission years, are roughly comparable.

A three-level model presented in a previous study (Carbonell-Olivares, Gil-Salom, & Soler-Monreal, 2009, p. 157) was applied to both the English and the Spanish texts to describe their rhetorical organisation. The highest level is that of the Move, and three categories were identified: *Establishing the Territory*, *Establishing the Niche* and

*Occupying the Niche*. The other levels concern steps and sub-steps. The model had been developed from the analysis of a corpus of 21 PhD thesis introductions in computing, written in Spanish. Although it was based on Bunton's (2002) move-step model for PhD thesis introductions in English, it included new steps and posited sub-steps. In M1, two new steps were added: *Explaining the institutional/research group context* and *Summarising previous background research*. Another step, *Defining terms*, proposed by Bunton as an independent step occurring both in M1 and M3, was found to include classifications as part of extended definitions. Five sub-steps were identified in step 2: *Making topic generalisations and giving background information: Indicating a problem or need, Indicating limitations, Giving examples, Defining terms and Giving or anticipating solutions*. In M3, the new step *Field of research*, was postulated to specify the field to which the study pertains. Bunton's step, *Findings or Results (Announcing or predicting principal findings)*, was widened to include the product of research, the model proposed, contributions and solutions to the problems or aspects investigated. Three sub-steps were included in step 2: *Work carried out/Announcing research: Work done, Work or aspects out of scope and Previous requirements*. Step 8, *Thesis structure*, was also subdivided into several substeps: *Overall thesis structure, Chapter structure, Chapter contents and Chapter goal*.

To ensure its reliability, the analysis was carried out in several phases. First, each of the three researchers independently identified and coded every segment of text. Then, individual codings were discussed in pairs by the researchers and agreement was reached where the codings differed. Finally, a consensus about the codings was reached by all three researchers.

### **3. Results and discussion**

When formal aspects are considered, some differences arise between the two corpora. The introductions present a great deal of variability as regards their length. The Spanish ones range from 3 to 18 pages (average 9.1). Half of the introductions have sections and the longest ones also have subsections. One introduction presents sub-subsections. The introductions in the English corpus tend to be longer and present more subdivisions than the Spanish ones. They range from 2 to 33 pages (average 12.1). Most

of them (8 out of 10) are divided into sections, and longer ones contain subsections and sub-subsections.

### 3.1 *Move analysis*

The three moves established in the CARS model are found in both groups of introductions. However, the move structures set out in [Tables 1 and 2](#) show that not all the introductions conform to the archetypal M1-M2-M3 (*Establishing the Territory-Establishing the Niche-Occupying the Niche*) structure. There is variation with respect to the model as far as the presence, the sequence, the cyclicity and the embedding of moves are concerned. By cyclicity we mean a recurrence of moves that makes up "cycling configurations" (Swales 1990, p. 158). Moves may be considered to be embedded when rhetorical aims typical of a move are found within another move. Such phenomena are common to both corpora and can be explained by the type of audience and the length of introductions. The graduate student believes s/he should recursively show the supervisor and the committee members what s/he has read and done, and explain the terminology and the purpose of the work carried out at different stages throughout the presentation of her/ his work. The length of the introduction also leads to cyclical sequences and embedding of moves. The theses with comparatively shorter introductions do not have repeated moves. On the other hand, long introductions, with several hypotheses as starting points, always present moves embedded within other moves. The differences of the move structures between both corpora are discussed in the following subsections.

#### 3.1.1 *Presence of moves*

The first striking difference between both corpora is that the English PhD thesis introductions have a more complex organisation (they contain a total of 145 moves vs. 50 in the Spanish corpus). English introductions contain 3 to 26 moves while the Spanish ones contain from 2 to 9 moves. As shown in [Tables 1 and 2](#), M1 and M3 are present in every thesis in each corpus. As for M2, although the move is present in all the English introductions, it is absent in two introductions in the Spanish corpus. This leads us to infer that the only obligatory moves in Spanish thesis introductions are M1, which presents the field of the thesis, and M3, which provides a description of the research carried out. A sociological explanation for this finding might be that Spanish graduate

students compete less for research space than for favourable assessment of the work they have done. Swales (2004, pp. 243–245) points out that in a corpus of non-English texts that can be considered to be equivalent to English ones, claiming knowledge and good performance in a specific field, seems to have a higher priority than establishing that there exists a gap in previous research that needs filling. Taking this into consideration, it seems that Spanish graduates write their introductions mainly for the immediate audience of the thesis, i.e. the supervisor and the committee.

In contrast, the presence of M2 in all the English introductions, along with M1 and M3, conforms to the CARS pattern closely. Following Yakhontova's considerations (2002, p. 229), we think that the writers' strategy of establishing the niche responds to the need to vie for attention and promote the research in the Anglo- American market, making their discourse both persuasive and self-promotional. The strategy of establishing the niche facilitates both the presentation of the research as novel and the claim for originality.

### 3.1.2 *Move patterns and cycling of moves*

Most introductions in both corpora begin with M1 (9 and 8 in the Spanish and the English corpora, respectively). Tables 3 and 4 show the number of cycles of the most frequent move patterns.

The prototypical sequence M1–M2–M3 occurs in 6 theses in the Spanish corpus (see Table 1). Three Spanish theses present one cycle of the pattern, which makes up the introduction. This results in a straightforward way of arranging information, which may be attributed to an awareness of the need to win recognition by successfully referring to theoretical issues, defining a problem and developing it. In three other introductions this pattern is combined with other moves, particularly M1 and M3 (see Table 3). Conversely, in seven introductions in the English corpus the M1–M2–M3 pattern is present in cycles and/or in combination with other moves and patterns. Although this may be partly due to the length of the text, the repetition of moves in every thesis introduction promotes a reader-friendly structuring of texts. It allows the writer to highlight the main points while maintaining the connection between her/his claims and accepted knowledge in the field.

Another usual sequence in both corpora is M1–M3. The number of such cycles varies from 1 to 3 in the Spanish corpus and from 1 up to 11 in the English corpus. It is

the only pattern of moves present in one Spanish introduction and appears in cycles and/or with other patterns in five introductions (see [Tables 1 and 4](#)). In the English corpus, M1–M3 is also found in cycles and/or in combination with other patterns.

### 3.1.3 *Embedding of moves*

The corpora show different trends in the internal structure of moves. As we can see in [Tables 1 and 2](#), in the Spanish corpus moves do not usually contain embedded moves (3 introductions), while in the English introductions embedding is quite frequent (6 introductions). There are 5 instances of embedded moves in the Spanish corpus, and 23 in the English one.

Both sets of theses also present slight differences with regard to the moves embedded and their frequency of embedding. M1 is not found to be embedded in the Spanish corpus, while it is embedded in M3 in the English corpus (10 instances). M2 is the most usually embedded move in the Spanish corpus (3 instances out of 5 total embedded moves) whereas it is the least frequently embedded move in the English theses (3 instances out of 23 total embedded moves). In the Spanish corpus M2 is embedded in M1 but in the English corpus it is embedded both in M1 and in M3. As for M3, it is embedded in English in a higher proportion than in Spanish (see [Tables 1 and 2](#)). Example 1 in the Appendix shows a sample of embedding of M3 in M1 in the Spanish corpus.

The differences in the number of embedded moves may be linked to the difference in length and complexity of the structures of the introductions in both corpora. The English introductions are in general longer than their Spanish counterparts while the latter present simpler move structures, as [Tables 1 and 2](#) reveal. Rhetorical variation between the two corpora may reflect the different relationship between the writer and the audience in terms of expectations. The higher number of embedded cases of M1 and M3 in the English theses may be seen as a reflection of the greater degree of interaction with the audience in an effort to gain acceptance.

The comparison of the figures for the moves making up the introductions with and without counting the embedded instances yields interesting results (see [Table 5](#)). Considering the frequencies of both embedded and non-embedded moves, the most frequent move is M3 in the Spanish introductions, whereas M1 and M3 are similarly frequent in the English corpus. The presence of non-embedded M2 seems to be less

prominent in the Spanish introductions. However, when considering the frequency of embedded instances of each move, M2 appears in both corpora in the same proportion while the presence of M1 in the Spanish corpus is lowered. These data reveal a slightly different understanding of the function of the introduction in both corpora. English writers seem to devote special attention both to the preliminary, contextual and background information (M1) and the occupation of a niche of research (M3), while Spanish writers tend to emphasise the presentation of their own work (M3).

### 3.2 *Step and sub-step analysis*

Some differences emerge when considering the number of instances in which particular steps (S) and substeps (SS) are recorded, and the number of introductions presenting them. This section discusses the similarities and differences with relation to each of the moves. We first note which steps are present in both corpora and which are found only in one corpus. Among those that are present in both corpora, we distinguish (1) those steps that seem more typical than the rest, considering the number of theses where they are realised and taking into account the number of instances; (2) those that seem equally common in both corpora, and (3) those that seem to be more frequent in one of the corpora.

#### 3.2.1 *Steps and sub-steps in M1*

The steps distinguished in this move appear in both sets of theses, except *Summarising previous background information*, which does not appear in the English corpus (see Table 6).

A previous analysis of Spanish PhD thesis introductions (Carbonell-Olivares et al., 2009, pp. 162–163) showed that S2 *Making topic generalisations and giving background information* tends to be a complex step, in which a wide variety of rhetorical strategies are performed. Hence, the need to distinguish sub-steps: *Indicating a problem or need* (SS2A), *Indicating limitations* (SS2B), *Giving examples* (SS2C), *Defining terms* (SS2D), and *Giving or anticipating solutions* (SS2E). In particular, SS2A and SS2B are postulated as distinct from problems or needs indicated in M2 *Establishing a Niche*. Some statements of a problem or need or the pointing out of limitations concern issues being tackled at the *Territory* level. In other words, these issues are examined in the broad context of the state of the art as problems, needs, or

limitations discussed, providing some background to the specific thesis work. They must be distinguished from the indications of problems, needs or limitations at the *Establishing the Niche* level, that is, be directly related to the specific work or Niche the writer intends to occupy.

S2 and S1 *Claiming centrality* can be regarded as crucial in the realisation of M1, according to the number of theses in which they are found. S2 occurred in every introduction in the Spanish corpus and in 9 out of the 10 introductions in the English corpus. S1 appears in six introductions in each set of theses. The other three steps (S3 *Defining terms/classifying*, S4 *Reviewing previous research* and S5 *Explaining the institutional/research group context*) are found in both corpora, but in different proportions.

Within S2, a few sub-steps are common in both corpora: SS2A Indicating a problem or need, SS2D Defining terms/classifying and commenting on terminology and, to a lesser extent, SS2E Giving or anticipating solutions (or ways to solve problems/to tackle needs). It is remarkable that even if writers of both corpora use SS2A in most introductions (8 theses in each corpus), this sub-step is realised in many more instances in the English corpus than in the Spanish one (50 vs. 13, respectively). The same occurs with SS2E Giving or anticipating solutions, with a much higher number of occurrences in the English corpus (35 in 6 theses, vs. 15 in the same number of Spanish introductions), thus reflecting the greater complexity of English PhD introductions. However, the case is different for SS2C Giving examples. It occurs frequently in the English thesis introductions (33 cases in 7 theses) but is not found in any of the 10 Spanish introductions.

S4 *Reviewing previous research* is typical of the introductions in English (21 instances in 8 theses), whereas it is rarely found in the Spanish ones (3 instances in 2 theses). It is not that Spanish writers do not often carry out this step, but that they do so when giving background information.

### 3.2.2 Steps and sub-steps in M2

As Table 7 shows, all the possible steps realising M2 are recorded in both corpora although more instances of these steps are found in the English corpus. Most frequently used one in both languages is S1B *Indicating a problem or need*, while S1A *Indicating a gap in research* and S1D *Continuing/Extending a tradition*, are found in a similar

number of instances. It is noteworthy that the writers of theses in English tend to prefer S1B (17 instances of S1B compared to 18 instances of all other steps), while the Spanish writers do not show a clear preference for any particular step. We deduce that the writers in English tend to emphasise the availability of a niche to justify their work by indicating a problem or need repeatedly. In doing so, they continually remind the audience of the novelty of their research. The frequency of S1D is also different in both corpora: it is rarely used in the Spanish corpus (present in 2 theses) while it is commonly found in the English one (present in 7 theses).

### 3.2.3 Steps and sub-steps in M3

As can be seen in Table 8, the most usual steps found in this move in both corpora seem to be S8 *Thesis structure*, S2 *Work carried out/Announcing research*, and S1 *Purposes, aims or objectives*. Spanish writers, however, mention the aims of their work much more frequently than those writing in English (20 instances in 8 theses vs. 12 in 6 theses, respectively).

The corpora also present significant differences in the frequency with which certain steps in M3 are used. In the English corpus, writers note other specific aspects of the research undertaken in more theses and in a considerably higher number of cases than the Spanish writers do. For example, it is much more usual to find writers in English mention the materials or subjects of the research (S5) and the findings or results (S6), justify or state the significance of the work (S7), the hypotheses and research questions, and make the statement of the thesis explicit. Conversely, in the Spanish corpus more emphasis is given to indicating the field of research (S3).

When introducing the work carried out (S2), introductions in English present a more elaborate set of strategies: stating the focus of research and explicitly mentioning the work or aspects beyond the scope of the thesis (SS2B) are strategies found principally in the English corpus. We also find an important difference in the frequency of appearance of the sub-step that describes the work done (SS2A) in both corpora. This sub-step is more common in the introductions in English (22 instances in 7 English theses vs. 8 instances in 6 Spanish theses).

### 3.2.4 Embedding of steps and sub-steps

We consider a step or sub-step  $x$  to be embedded in another step or sub-step  $y$  when the realisation of  $x$  is part of  $y$ . In other words,  $y$  includes a distinct realisation of  $x$ . Tables 9 and 10 display the steps and sub-steps that are embedded in other steps and sub-steps. Embedding between steps and sub-steps is widely found in M1 and M3 in both corpora, whereas this phenomenon is very rare in M2. This reflects the extensive development of M1 and M3 in contrast with the conciseness in the realisation of M2 in both sets of theses. The Spanish and English introductions present embedding of steps and sub-steps in substantially different proportions (17 cases vs. 49 cases, respectively), and the nature of the embedded steps and sub-steps is also different in the corpora, both for M1 and M3.

In M1, we find 12 instances of embedding in the Spanish introductions, and 25 in the English ones. The corpora basically coincide in the complexity of S2 *Making topic generalisations and giving background information* (see Table 9). It is the most frequently embedded step (9 and 14 cases of embedded S2 in the Spanish and in the English corpora, respectively). This reveals the importance attached by writers to making constant references to background information when claiming the centrality of the topic (S1), giving definitions and classifying (S3) and/or reviewing previous research (S4). S2 is also the step that most often contains an embedded step (in the Spanish corpus it is the only step that has embedding; in the English corpus, 16 out of 25 steps that have embedding are S2). For instance, the presentation of background information and topic generalisations may contain S4, S1, and also S8 *Thesis structure and Application of product* of M3. These results for M1 reveal the concern of authors with establishing a link between the state-of-the-art of the topic and the present research at different moments throughout the development of the thesis introduction.

There are also some differences between the corpora with regard to which steps present embedded steps. While the Spanish introductions show embedding only within S2, as already stated, the English corpus does so also in S1, S3 and S4.

In M3 embedding differs widely in both corpora. It is practically non-existent in the Spanish corpus, but frequent in the English one (4 vs. 24 cases, respectively). Only one Spanish text presents embedding: that is, when stating the purposes, aims or objectives of the thesis (S1), the writer also announces the work carried out (S2). In the English texts S2 *Work carried out*, S4 *Method/Parameters of research* and S8 *Thesis structure* contain the greatest number of cases of embedding, consisting usually in

*Defining terms* but also in sub-steps indicating aspects of the thesis structure and the work carried out. Example 2 (see appendix) shows S6 with an embedded *Defining terms* step.

These results show the greater complexity of the structure of the English introductions in comparison with the Spanish texts. Overall, the steps containing embedded steps as well as the variety of embedded steps are more diverse in the English corpus. In this sense, the English introductions reflect greater concern with providing some information that completes, extends or supports the range of informational strategies deployed throughout the introduction. The aim underlying the resulting complexity is to facilitate the reader's understanding of the connection between the different information units of the introduction.

### 3.2.5. *Overlapping of steps and sub-steps*

Occasionally, one can find steps and sub-steps which are realised in combination with other steps or sub-steps, i.e. within the same sentence or text segment. We refer to this phenomenon as overlapping of steps/sub-steps. It is found in M1 and M3, and is more frequent among the English introductions.

In M1, mostly the same steps and sub-steps are found to overlap in the Spanish and the English introductions. In both languages, some reference to the review of previous research (S4) may be carried out at the same time as background information (S2) is provided. In the Spanish introductions the review is also found while definitions of terms and classifications (S3) are given. In English other steps and sub-steps overlap as well: S1 *Claiming centrality*, SS2A *Indicating a problem/need*, SS2C *Giving examples* and S5 *Explaining the institutional/research group context*.

In M3, the Spanish introductions present the particularity of combining the realisation of the goal statement (S1) with a variety of other steps and sub-steps (S2 *Work carried out/Announcing research*, S4 *Method/Parameters of research*, S7 *Justification/Significance*, SS8D *Chapter goal*). In contrast, among the English texts S1 is found to overlap with another step in just one instance (with S2). The combinations S1 + S2 and S1 + S7 tend to provide a straightforward starting point for the presentation of the thesis research. It is also noteworthy that, unlike the Spanish texts, the English ones realise S4 *Method/Parameters of research* more often in combination with other

steps, such as *Application of product* and *Defining terms*. The steps that frequently overlap with other steps in the English introductions are those only occasionally present in the Spanish corpus, thus reflecting the higher expectations in the Anglo-American market society when it comes to producing applicable and well-defined research outputs.

However, there are also some similarities between both corpora in the performance of certain steps in M3. A salient case is that involving the presentation of findings or results (S6), which is realised at the same time as other key steps: announcing the work carried out (S2, see Appendix, example 3), explaining the methodology employed and the research parameters (S4), providing the thesis statement, and mentioning the application of the product developed. .

#### **4. Conclusion**

We have carried out a contrastive analysis of Spanish and English PhD thesis introductions in computing in order to determine if there are qualitative and quantitative differences in the use of the rhetorical strategies adopted by each discourse community. At the formal level it has been found that English introductions are more complex as regards the presentation and organisation of the information provided.

The move analysis has revealed that not all the introductions conform fully to the CARS model, although the M1–M2–M3 arrangement is the structural pattern most generally followed. This is usually combined with other patterns and moves, resulting in frequent cycles and the embedding of moves. In the Spanish corpus M2 is not always present. Spanish academic conventions seem to establish M1 and M3 as obligatory moves in PhD introductions and do not stress the need to establish a niche which would justify the work done. This finding agrees with Burgess's (2002, p. 198) hypothesis that certain writers of RA introductions "view the problem they address as entirely uncharted territory for their readers", which explains why Spanish PhD candidates put so much effort into describing the territory and occupying a particular niche. This comment may make clear why showing knowledge of the field of research and defending a novel and specific contribution to it seem to have higher priority than establishing a gap in previous research, which explains the non-antagonistic stance in the Spanish introductions. Their aim is to provide a broad contextualisation of their research and a description of the findings so as to establish "the scholarly credibility of

the author as a worthy member of the research community” (Yakhontova, 2002, p. 231). According to this view, the Spanish introductions conform to Swales’s (2004, p. 244) flexible OARO (Open a Research Option) model, which reflects a more relaxed world in which there is less competition for research space. In contrast, the structure of the English introductions reveals a concern with establishing a niche. Yet, though they follow the CARS model more closely, combinations of moves and patterns, cyclicity and embedding tend to make their organisation more complex.

The step analysis suggests that the structure of Spanish introductions is mainly motivated by the presentation of background information (S2 of M1) and the work carried out (S2 of M3), as well as the deployment of that information in a sequential and orderly fashion. The English introductions show great concern for the inclusion of background information and a separate step dedicated to the review of previous research (S2 and S4 of M1), but they also tend to stress the writer’s own work, its originality and contribution to the field of study (S1, S2, S6, S7 of M3). The English writer’s interest in referring repeatedly to these aspects accounts for constant alternation of the corresponding steps and sub-steps. Such alternation produces a number of embedded and overlapping steps and sub-steps. It allows the enhancement of the writer’s claims while involving the audience in the reasoning.

This study has been intended as a modest genre analysis of two comparable corpora of PhD theses. The relatively small size of the corpora leads us to view the results of the analysis with caution, and our findings should thus be corroborated with a larger corpus. Our study has also attempted to contribute to the understanding of academic writing in different cultural and linguistic traditions, and thus to CR studies. We believe that insights into cultural, linguistic and generic conventions will help teachers to guide novice academic writers when writing up their research.

## Appendix

### Example 1. Embedding of moves in the Spanish corpus: M1 [M3]

M1–S2. En los sistemas distribuidos que controlan un proceso o aplicación que pueda entrañar un cierto riesgo, [...], un requerimiento que se hace cada vez más necesario es la mejora de la confiabilidad [...] A nivel de sistema se puede mejorar la garantía de funcionamiento mediante la inclusión de estrategias de Puntos de Recuperación entre los distintos nodos.[embedded M3–S3 begins]. Esta tesis se centrará en el último de los puntos, es decir, en la introducción de técnicas de Puntos de Recuperación. [...] [end of embedded M3–S3]. La Recuperación por Vuelta Atrás es una de las técnicas más conocidas para recuperar un sistema después de la ocurrencia de errores. (T7. Rubio Moreno, A., 2002. *Propuesta de una*

nueva técnica de Puntos de Recuperación a dos niveles para sistemas distribuidos de control industrial, pp. 6-7)

#### Example 2. Embedding of steps in M3 in the English corpus: S6 [*Defining terms*]

M3-S6 (Results in terms of *model proposed*) The thesis presents a model of the requirements process which provides guidance for identifying and developing descriptions of the perspectives, and resolution of conflicts between them. [*Defining terms* begins] A perspective can be thought of as a consistent view of the world arising from the context of a particular role. Perspectives do not necessarily correspond to people, as one person may use several perspectives [...] Perspectives are represented using viewpoints, which are formatted descriptions in some appropriate representation scheme. [end of *Defining terms*]. No restriction is placed on the form of those descriptions, nor on the degree of formality. Hence the model may be used in conjunction with existing specification languages and knowledge representation schemes (end of S6) (T5. Easterbrook, S, 1991. *Elicitation of requirements from multiple perspectives*, pp. 10-11)

#### Example 3. Overlapping of steps in M3 in the English corpus: S6 + SS2A *Work done*

M3-S6 (*Contribution*) [...] we contribute to the area of type-based approaches to security (+ SS2A) by presenting type and effect systems which incorporate a machinery for tracing the flow of values in a distributed setting where functions are the essential element of computation. (T4. Dilsun Kirli, Z. 2001. *Mobile computation with functions*, p. 3)

## References

- Ädell, A. (2008). Metadiscourse across three varieties of English: American, British, and advanced-learner English. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 45-62). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Ahmad, U. K. (1997). Research article introductions in Malay: Rhetoric in an emerging research community. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 273- 304). Berlin/New York: Mouton de Gruyter.
- Árvay, A. & Tankó, G. (2004). A contrastive analysis of English and Hungarian theoretical research article introductions. *International Review of Applied Linguistics*, 42, 71-100.
- Bunton, D. (2002). Generic moves in Ph.D. thesis introductions. In J. Flowerdew (Ed.), *Academic discourse* (pp. 57-75). London: Longman.
- Bunton, D. (2005). The structure of PhD conclusion chapters. *Journal of English for Academic Purposes*, 4, 207-224.
- Burgess, S. (2002). Packed houses and intimate gatherings: Audience and rhetorical structure. In J. Flowerdew (Ed.), *Academic discourse* (pp. 196-215). London: Pearson Education.
- Carbonell-Olivares, M., Gil-Salom, L. & Soler-Monreal, C. (2009). The schematic structure of Spanish PhD Thesis introductions. *Spanish in Context*, 6(2). 151–175. doi 10.1075/sic.6.2.01car
- Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge: Cambridge University Press.
- Connor, U. (2002). Contrastive rhetoric and academic writing: Multiple texts, multiple identities. *Forum: Applied Linguistics Newsletter*, 23(1), 1- 6.
- Connor, U. (2004). Intercultural rhetoric research: Beyond texts. *Journal of English for Academic Purposes*, 3; 291-304.

- Connor, U. (2008). Mapping multidimensional aspects of research: Reaching to intercultural rhetoric. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 299-315). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Cooley, L. & Lewkowicz, J. (1997). Developing awareness of the rhetorical and linguistic conventions of writing a thesis in English: addressing the needs of EFL/ESL postgraduate students. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 113-129). Berlin/New York: Mouton de Gruyter.
- Duszak, A. (1997). Cross-cultural academic communication: A discourse-community view. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 11- 39). Berlin/New York: Mouton de Gruyter.
- Feng, H. (2008). A genre-based study of research grant proposals in China. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 63-86). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Gnutzmann, C. & Oldenburg, H. (1991). Contrastive text linguistics in LSP-research: Theoretical considerations and some preliminary findings. In H. Schröder (Ed.), *Subject-oriented texts* (pp. 103-137). Berlin: Walter de Gruyter.
- Kwan, B. (2006). The schematic structure of literature reviews in doctoral theses of applied linguistics. *English for Specific Purposes*, 25, 30-55.
- Kaplan, R. B. (1966). Cultural thought pattern in inter-cultural education. *Language Learning*, 16(1), 1- 20.
- Lee, S. (2000). Contrastive rhetorical study on Korean and English research paper introductions. *Pan-Pacific Association of Applied Linguistics*, 4(2), 316-336.
- LoCastro, V. (2008). Long sentences and floating commas. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 195-217). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Loukianenko Wolfe, M. (2008). Different cultures- different discourses? Rhetorical patterns of business letters by English and Russian speakers. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 87-121). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Martín-Martín, P. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22, 25-43.
- Martín-Martín, P. & Burgess, S. (2004). The rhetorical management of academia criticism in research article abstracts. *Text*, 24(2), 171-195.
- Melander, B., Swales, J.M. & Frederickson, K.M. (1997). Journal abstracts from three academic fields in the United States and Sweden: National or disciplinary proclivities? In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 251-272). Berlin/New York: Mouton de Gruyter.
- Moreno, A. I. (1997). Genre constraints across languages: Causal metatext in Spanish and English RAs. *English for Specific Purposes*, 16, 161-179.
- Moreno, A. I. (1998). The explicit signalling of premise-conclusion sequences in research articles: A contrastive framework. *Text*, 18, 545-585.
- Moreno, A. I. (2004). Retrospective labelling in premise-conclusion metatext: An English-Spanish contrastive study of research articles on business and economics. *Journal of English for Academic Purposes*, 3, 321-339.
- Moreno, A. I. (2008). The importance of comparable corpora in cross-cultural studies. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric:*

- Reaching to intercultural rhetoric* (pp. 147-168). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Pak, C. S. & Acevedo, R. (2008). Spanish-language newspaper editorials from Mexico, Spain and the U.S. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 123-145). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Paltridge, B. (2002). Thesis and dissertation writing: An examination of published advice and actual practice. *English for Specific Purposes*, 21, 125-143.
- Ridley, D. (2000). The different guises of a PhD thesis and the role of a literature review. In P. Thompson (Ed.), *Patterns and perspectives: Insights into EAP writing practice* (pp. 61-76). Reading: University of Reading.
- Suárez, L. & Moreno, A. I. (2008). The rhetorical structure of academic book reviews of literature: An English-Spanish cross-linguistic approach. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 147-168). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M. (2004). *Research genres: Exploration and applications*. Cambridge: Cambridge University Press.
- Taylor, G. & Chen, T. (1991). Linguistics, culture, and subcultural issues in contrastive discourse analysis: Anglo-American and Chinese scientific texts. *Applied Linguistics*, 12(3), 319-336.
- Thompson, P. (2001). *A pedagogically-motivated corpus-based examination of PhD theses: macrostructure, citation practices and uses of modal verbs*. Unpublished doctoral dissertation. Reading: University of Reading.
- Wang, W. (2008). Newspaper commentaries on terrorism in China and Australia: A contrastive genre study. In U. Connor, E. Nagelhout & W. V. Rozycki (Eds.), *Contrastive rhetoric: Reaching to intercultural rhetoric* (pp. 169-191). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Yakhontova, T. (2002). Selling or telling? The issue of cultural variation in research genres. In J. Flowerdew (Ed.), *Academic discourse* (pp. 216-232). London: Pearson Education.

## Tables

Table 1.  
Move structure of the Spanish PhD thesis introductions

Thesis	Move sequence	M1	M2	M3	Total*
T1	M1[M2]-M3-M1[M2]-M3	2	0 [2]	2	4[2]
T2	M1-M3-M1-M3-M1-M3	3	0	3	6
T3	M1[M2]-M3	1	0 [1]	1	2[1]
T4	M1-M2-M3	1	1	1	3
T5	M1-M3-M1-M3-M1-M2-M3-M2-M3	3	2	4	9
T6	M1-M2-M3	1	1	1	3
T7	M1[M3][M3]-M3-M1-M3	2	0	2 [2]	4[2]
T8	M1-M2-M3	1	1	1	3
T9	M3-M1-M2-M1-M2-M3	2	2	2	6
T10	M1-M2-M3-M1-M3	2	1	2	5
Total		18	8 [3]	19 [2]	45[5]

[...] indicates a move embedded within another move.

\*The total figures reflect separately the number of moves and the number of embedded moves.

Table 2.  
Move structure of the English PhD thesis introductions

Thesis	Move sequence	M1	M2	M3	Total*
T1	M1-M2-M1-M3-M1-M3	3	1	2	6
T2	M1-M2-M3-M1[M3]M1-M2-M3-M1-M3-M1-M2-M3-M2-M3	5	4	5 [1]	14[1]
T3	M1-M2-M3-M2-M3	1	2	2	5
T4	M1-M3-M1-M2-M3-M1-M2-M3-M1-M2-M3	4	3	4	11
T5	M1-M2-M3-M1-M2-M3-M1-M2-M3-M1-M2-M3	4	4	4	12
T6	M1-M3-M1-M3-M1-M3-M1-M3-M1-M3-M1-M3-M1-M2-M1-M2-M1[M3]-M3-M1-M3-M1[M3-M3]-M3-M1-M3[M1]-M1-M3	13 [1]	2 [1]	11 [3]	26[4]
T7	M2-M3-M2-M3-M1-M3-M1-M2-M1-M2-M3[M1-M1-M2-M3]-M1-M2-M1-M2-M1-M3-M1-M3-M1[M2]-M3	8[2]	6[2]	7[1]	21 [5]
T8	M1[M2]-M3-M1-M3-M1-M2-M3-M1-M3-M2-M3-M1-M2-M3[M1-M3-M1-M3-M1-M3]	5 [4]	3 [1]	6 [4]	14[9]
T9	M3-M2-M3[M1-M1-M1]	0 [3]	1	2	3 [3]
T10	M1[M3]-M3-M1-M2-M3-M2-M3-M1-M2-M3	3	3	4 [1]	10[1]
Total		46[10]	29[3]	47[10]	122[23]

[...] indicates a move embedded within another move.

\*The total figures reflect separately the number of moves and the number of embedded moves.

Table 3.  
Number of cycles of the sequence M1-M2-M3 in the Spanish and the English corpus

Number of cycles of the sequence M1-M2-M3	Number of Spanish theses	Number of English theses
1 cycle	6	2
2 cycles	-	2
3 cycles	-	2
4 cycles	-	1

Table 4.  
Number of cycles of the sequence M1-M3 in the Spanish and the English corpus

Number of cycles of the sequence M1-M3	Number of Spanish theses	Number of English theses
--	--------------------------	--------------------------

1 cycle	2	3
2 cycles	3	1
3 cycles	1	1
6 cycles	-	1
11 cycles	-	1

Table 5.  
Frequency of instances of moves in relation to the total number of moves

Moves	Spanish (total n. 45, excluding embedded instances)	English (total n. 122, excluding embedded instances)	Spanish (total n. 50, including embedded instances)	English (total n. 145, including embedded instances)
M1	40%	37.7%	36%	38.6%
M2	17.8%	23.8%	22%	22.1%
M3	42.2%	38.5%	42%	39.3%

Table 6.  
Distribution of steps and sub-steps in Move 1: Establishing a Territory

Steps (S) and sub-steps (SS)	Number of instances		Number of theses	
	Spanish	English	Spanish	English
S1: Claiming centrality (importance of topic)	8	8	6	6
S2: Making topic generalisations and giving background information	19	52	<b>10</b>	<b>9</b>
SS2A: Indicating a problem/need	13	<b>50</b>	8	8
SS2B: Indicating limitations	8	8	4	3
SS2C: Giving examples	0	<b>33</b>	<b>0</b>	<b>7</b>
SS2D: Defining terms/classifying and commenting on terminology	37	50	8	8
SS2E: Giving or anticipating solutions (or ways to solve problems/to tackle needs)	15	<b>35</b>	6	6
S3: Defining terms/classifying	13	9	3	5
S4: Reviewing previous research	3	<b>21</b>	<b>2</b>	<b>8</b>
S5: Explaining the institutional/research group context	2	3	2	1
/Summarising previous background information/	1	0	1	0

Table 7.  
Distribution of steps and sub-steps in Move 2: Establishing a Niche

Steps	Number of instances		Number of theses	
	Spanish	English	Spanish	English
S1A: Indicating a gap in research	4	9	<b>4</b>	<b>7</b>
S1B: Indicating a problem or need	6	<b>17</b>	<b>4</b>	<b>8</b>
S1C: Question-raising	1	1	1	1
S1D: Continuing/Extending a tradition	4	8	<b>2</b>	<b>7</b>

Table 8.  
Distribution of steps and sub-steps in Move 3: Occupying the Niche (Announcing the present research)

Steps (S) and sub-steps (SS)	Number of instances		Number of theses	
	Spanish	English	Spanish	English
S1: Purposes, aims or objectives	<b>20</b>	12	<b>8</b>	<b>6</b>
S2: Work carried out/Announcing research	5	<b>29</b>	<b>8</b>	<b>8</b>
SS2A: Work done	8	<b>22</b>	6	7
SS2B: Work or aspects out of scope	3	<b>13</b>	3	<b>9</b>
SS2C: Previous requirements	0	3	0	1
S3: Field of research	<b>5</b>	0	<b>4</b>	0
S4: Method/Parameters of research	<b>10</b>	<b>21</b>	<b>5</b>	<b>6</b>
S5: Materials or Subjects	1	<b>5</b>	1	<b>4</b>
S6: Findings or Results: Product of research/Model proposed/Contributions/Solutions	<b>10</b>	<b>22</b>	4	<b>9</b>

S7: Justification/Significance	<b>3</b>	<b>10</b>	3	<b>6</b>
S8: Thesis structure	9	21	<b>9</b>	<b>10</b>
SS8A: Overall thesis structure	4	4	4	3
SS8B: Chapter structure	9	12	7	5
SS8C: Chapter contents	<b>50</b>	<b>78</b>	<b>9</b>	<b>10</b>
SS8D: Chapter goal	3	6	<b>2</b>	<b>5</b>
/Research questions or Hypotheses/	1	<b>9</b>	1	<b>5</b>
/Application of product/	3	3	3	3
/Evaluation of product/	1	0	1	0
/Defining terms/	2	10	1	5

Table 9.  
Embedding of steps and sub-steps in M1: Establishing a Territory

Spanish corpus		English corpus	
Steps (S) and sub-steps (SS)	N. of instances	Steps (S) and sub-steps (SS)	N. of instances
S2 [S1]	1	S1 + SS2A [SS2C]	1
SS2D [S4]	1	S2 [S1]	2
SS2D [SS8B of M3]	1	S2 [S4]	5
SS2E [SS2B]	4	S2 [S2 + S4]	1
SS2E [SS2D]	5	S2 [SS8C of M3]	1
		S2 [Application of product of M3]	1
		S2 + S4 [SS2C + Application of product of M3]	1
		SS2E [SS2C - SS2E]	2
		SS2E [SS2A]	1
		SS2E [SS2B]	1
		SS2E [SS2C]	1
		S3 [S2]	1
		S3 [SS2A]	1
		S3 [SS2C]	1
		S4 [SS2A]	2
		S4 [SS2C]	1
		S4 [S3]	2

Table 10.  
Embedding of steps and sub-steps in M3: occupying the Niche (Announcing the present research)

Spanish corpus		English corpus	
Steps (S) and sub-steps (SS)	N. of instances	Steps (S) and sub-steps (SS)	N. of instances
S1 [S2]	4	S2 [SS8D]	1
		SS2A [SS2C]	1
		SS2A + S6 [Defining terms]	1
		S4 [Defining terms]	3
		S4 [SS2D]	1
		S5 + S4 [SS8D]	1
		S6 [Defining terms]	2

S8 [2A]	7
S8 [Application of product]	3
S8 [S7]	1
SS8D [SS8C]	1
Defining terms [S6]	1
Defining terms [SS2A]	1

---