



UNIVERSITAT
JAUME I

HUMANITIES AND SOCIAL SCIENCES FACULTY

English Studies Department
2006

**DISCOURSE MARKERS WITHIN THE UNIVERSITY
LECTURE GENRE:
A contrastive study between Spanish and North-
American lectures**

PhD dissertation presented by:

Begoña Bellés Fortuño

Supervised by:

Dr. Inmaculada Fortanet Gómez



UNIVERSITAT
JAUME•I

FACULTAT DE CIÈNCIES HUMANES I SOCIALS

Departament d'Estudis Anglesos
2006

**LOS MARCADORES DISCURSIVOS EN EL GÉNERO
DE LA CLASE MAGISTRAL:
Un estudio comparativo entre el español y el inglés
americano.**

Tesis doctoral presentada por:

Begoña Bellés Fortuño

Dirigida por:

Dr. Inmaculada Fortanet Gómez

INDEX

INTRODUCTION.....	3
1. ENGLISH AND SPANISH IN THE GLOBAL WORLD.	11
1.1. English as a lingua franca.....	13
1.1.1. English as the language of international academic exchange..	19
1.2. Spanish as a lingua franca	24
1.2.1. Spanish as the language of international and academic exchange	28
2. DISCOURSE ANALYSIS	33
2.1. Introduction	35
2.2. Genre analysis	39
2.3. Contrastive rhetoric	45
2.3.1. Contrastive rhetoric studies in Spanish.....	56
2.4. Corpus linguistics: a methodological approach.....	58
3. THE GENRE OF LECTURE.....	65
3.1. Introduction.	67
3.2. Research into the lecture comprehension process.....	68
3.2.1. Phonological and lexico-grammatical features in lectures.	69
3.2.2. The use of micro-skills and strategies.....	70
3.3. Research into lecture discourse.	72
3.3.1. Lecturing styles.....	72
3.3.2. Syntax of lectures.	75
3.3.3. Lecture structural patterns.	77
3.3.4. The role of discourse markers in lectures	79
3.4. Lectures in the U.S.A.	82
3.4.1. Introduction.....	82
3.4.2. The example of the University of Michigan.....	85
3.5. Lectures in Spain	86
3.5.1. Introduction.....	86
3.5.2. The example of Universitat Jaume I.....	89

4. DISCOURSE MARKERS	91
4.1. Introduction	93
4.2. Definitions of DMs: The fundamentals.....	98
4.3. Corpus studies on DMs (1990s onwards).....	112
4.4. Broad classifications of DMs	119
4.5. Purpose of the study	137
5. METHOD	141
5.1. Corpus selection	143
5.2. Corpus description.....	148
5.3. Model and method of analysis.....	154
5.3.1. DMs Classification model.....	156
6. ANALYSIS OF RESULTS: DMs Classification model application.	165
6.1. Introduction	167
6.2. DMs results in the NAC and SC: an overall view.....	168
6.3. Micro-markers.	172
6.3.1. Micro-markers in the NAC.....	172
6.3.2. Micro-markers in the SC.....	183
6.3.3. Comparison of Micro-markers in the NAC and SC.	193
6.4. Macro-markers.	194
6.4.1. Macro-markers in the NAC	194
6.4.2. Macro-markers in the SC.....	205
6.4.3. Comparison of Macro-markers in the NAC and CS.....	212
6.5. Operators	214
6.5.1. Operators in the NAC	214
6.5.2. Operator in the SC	222
6.5.3. Comparison of Operators in the NAC and the SC.....	230
6.6. DMs collocations.....	231
6.6.1. Introduction.....	231
6.6.2. Micro-markers collocations in the NAC.....	232
6.6.3. Micro-markers collocations in the SC	236
6.6.4. Macro-markers collocations in the NAC.....	239
6.6.5. Macro-markers collocations in the SC.....	242

6.6.6. Operators collocations in the NAC.....	244
6.6.7. Operators collocations in the SC	247
7. DISCUSSION	251
7.1. Introduction	253
7.2. Comments on the overall DMs results	255
7.2.1. Comments on the micro-markers results.	259
7.2.2. Comments on the macro-markers results.	262
7.2.3. Comments on the operators results.....	266
7.3. Comments on the overall collocations results.....	272
7.3.1. Comments on the micro-markers collocations	274
7.3.2. Comments on the macro-markers collocations.....	276
7.3.3. Comments on the operators collocations.....	278
8. CONCLUSION	281
8.1. Concluding remarks	283
8.2. Pedagogical implications.....	293
8.3. Limitations.....	297
8.4. Further research.....	298
9. REFERENCES	301
10. APPENDIX A: North-American Corpus (NAC)	335
LE1	337
LE2	369
LE3	403
LE4	431
LE5	463
LE6	487
LE7	511
LE8	545
LE9	565
LE10	593
LE11	637

LE12	659
11. APPENDIX B: Spanish Corpus (SC)	695
LE1	697
LE2	719
LE3	743
LE4	773
LE5	799
LE6	819
LE7	843
LE8	873
LE9	893
LE10	915
LE11	943
LE12	961
12. APPENDIX C: SC Classroom observation guide	981
LE1/SC	983
LE2/SC	985
LE3/SC	988
LE4/SC	990
LE5/SC	992
LE6/SC	994
LE7/SC	997
LE8/SC	1001
LE9/SC	1003
LE10/SC	1005
LE11/SC	1007
LE12/SC	1009

INDEX OF TABLES

Table 1. Bathia's (2002) language description stages	40
Table 2. Chaudron & Richards' (1986) micro-markers classification	120
Table 3. Chaudron & Richards' (1986) macro-markers analysed in their study	122
Table 4. Morell's micro and macro-markers samples and her proposed taxonomy	126-128
Table 5. Micro-markers classification model (Bellés-Fortuño, 2004)	130
Table 6. Macro-markers classification model (Bellés-Fortuño, 2004).....	130
Table 7. Portoles' (1998) DMs classification	135
Table 8. SC Transcription norms.....	145
Table 9. North-American English lectures description	152
Table 10. Spanish lectures description.	153
Table 11. Micro-markers results in number of occurrences and frequency rate in NAC	173
Table 12. Micro-markers results in number of occurrences and frequency rate in SC	192
Table 13. Macro-markers results in number of occurrences and frequency rate in the NAC.....	204
Table 14. Macro-markers results in number of occurrences and frequency rate in the SC	211
Table 15. Operators results in number of occurrences and frequency rate in the NAC.....	221
Table 16. Operators results in number of occurrences and frequency rate in the SC	229

Table 17. Micro-markers collocations in the NAC	236
Table 18. Micro-markers collocations in the SC.....	239
Table 19. Macro-markers collocations in the NAC.....	241
Table 20. Macro-markers collocations in the SC	244
Table 21. Operators collocations in the NAC.....	247
Table 22. Operators collocations in the SC	249

INDEX OF FIGURES

Figure 1. Classification of oral academic genres. (Fortanet 2005).....	22
Figure 2. Areas where Spanish is spoken in the world.....	25
Figure 3. Cultural Thought Patterns in Intercultural Education. (Kaplan 1984).....	49
Figure 4. A ‘symbiotic’ relation between human and machine in corpus processing (Leech 1991)	60
Figure 5. González’s (2005) proposal of the distribution of markers in the discourse structure components.....	133
Figure 6. DMs functional meanings and matching relational functions	158
Figure 7. DMs classification model for this study	161
Figure 8. DMs results in the NAC and SC	168
Figure 9. Total rate of micro-marker, macro-markers and operators in the NAC and SC.....	169
Figure 10. Total rate of Micro-markers in the NAC and SC.....	170
Figure 11. Total rate of Macro-markers in the NAC and SC	170
Figure 12. Total rate of Operators in the NAC and SC	171

To my other self

ACKNOWLEDGEMENTS

First and foremost, I am mostly indebted to my colleague and friend Dr. Inmaculada Fortanet for her comments and insights along the development of this dissertation and for introducing me to the fascinating field of academic discourse. I will not forget her fine company in Ann Arbor (MI) and the good chats at home with our kids around accompanied by an aromatic coffee in the very hot summer of 2006.

There is a person who has undoubtedly suffered along the madness of 'clearing up' a PhD dissertation; she has showed me that friendship can still be part of the wild crowds prevailing at university; she has been an example of 'giving rather than receiving'. My entire gratitude goes to my friend and colleague Merche Querol Julián, wishing her the best personally and academically.

Also, a stout support in the development of this dissertation has been my friend Dr. Juan José Ferrer Maestro for his tiredly repeated phrase 'más sabe el diablo por viejo que por diablo', his wise advice and encouraging words whenever I needed them.

A special mention goes to the members of the ELI (English Language Institute) at the University of Michigan in Ann Arbor (USA), especially to John Swales for having revised a vague draft of the thesis and showing me insights and new directions in the field of academic discourse. Special thanks to Annelie Adel, her husband Greg Garretson and Julia Salehzadeh as well as the other members of the ELI for their good company and help during my stay with them. Thanks guys!

I also wish to thank my colleagues within the GRAPE research group since they have been witnessing the steps of this dissertation. Thanks to Dr. Noelia Ruiz

Madrid, Dr. Miguel Ruiz and Dr. Ania Saorín for being there to chat every now and then.

I am also grateful to some of my colleagues at the Department of English Studies and especially to Dr. M^a Carmen Campoy for her support, Dr. Ignasi Navarro for his friendship, Dr. Jose Luis Otal for his trustworthiness and Toni Silvestre for his fruitful kindness and endless forwarded Power Points. I have to mention the English Studies Department secretaries Pilar Sebastian and Pilar Martínez for their good work and joviality. Also thanks to Dr. José Luis Blas Arroyo at the European Philology Cultures Department for his help with the Spanish literature and his rich comments. Unique and distant assistance has been that of my Chilean friend Roxana Orrego at the Universidad de Santiago de Chile whose long chats on the Skype have made my Sundays at home less dull.

There are other people within and without the Universitat Jaume I that personally or professionally have been linked to this PhD dissertation and sometimes collaborated. Thanks to Dr. Miguel Angel Fortea Bagán and his wife M^a Angeles Ferreres, to some members of the SI (Servei de Informàtica) and lecturers at the FCJE and the ESTCE, especially Dr. Enrique Belenguer Balaguer for his help and encouragement. I owe some of my most pleasant and enjoyable time to my neighbors at Mohino, especially Lidia and Jordi for adopting my kids from time to time.

All my gratitude goes to the lecturers that voluntarily allowed me to interrupt in their classrooms to record their lectures; clearly they have been key for the fulfillment of the current study, although they are guaranteed anonymity.

Also thanks to those who have been sharing their everyday life with me and suffered the side effects of a PhD dissertation, my father Antonio and my brother

Toni. I am obliged to my mother, who deserves a special mention, for her endless patience with me and my two kids, for being there when things seem to have no exit and for believing in me at all times. This is for you mum!

I cannot forget those who have very closely suffered my ups and downs, my changes of humour and my absences. I am indebted to my husband Dani and my children Héctor and Álex. My sons have certainly been the best help, support and medicine; they cheered me up whenever I needed it, spreading their joy and pleasure and giving me the strength to continue. Thanks sweeties!

INTRODUCTION

INTRODUCTION

In the current era of globalization, languages and communication seem to be paramount and the key for understanding among countries. Within this situation, English has gained the status of a *lingua franca* becoming the language of commerce, business and academic world. In the near future English will be mainly used as a second language fostering communication between non-native speakers in multilingual contexts. This phenomenon will also affect the ideological discourse about languages, linguistic competence and identity (Graddol 1999, Chew 1999).

Although not as well positioned as English, Spanish is also important in today's globalized world. Considered an international language, Spanish is the third language in the world according to its number of speakers with over 400 million, only preceded by Chinese with 1,000 million speakers and English with 500 million speakers Spanish is undoubtedly very well positioned not only because of its number of speakers, but also due to its past culture and geographical extension.

However, my interest is not primarily on the importance of a language because of its number of speakers, but because of the effects and use of English and Spanish within the academic world and, more concretely, in the study of a specific classroom genre as it is the university lecture. Regarding the use of English in academic settings around the world, I have to review Graddol's (1999) considerations about whether the authoritative norms of usage of the native speaker will prevail in the academic contexts of the English language.

This fact invites reflection towards to which extent academic discourses, and more concretely the lecture as the central instructional activity of higher education institutions, could be affected and faced. We should bear in mind the “international” background (Benson 1994: 182) of students and what important role lectures play for their learning process success. Due to the internationalization of university teaching mainly in Europe (e.g. Erasmus-Socrates programs, Leonardo da Vinci projects), and also in the United States (e.g. postgraduate education), academic discourse in English has become a main concern in higher education institutions, focusing especially on spoken academic discourse, which should be looked at as paramount for both students and faculty.

At a tertiary educational level, Spanish as an international academic language is gaining importance with the great numbers of Latin-American students who come to Europe to take Master’s or doctorate courses. At present, most European universities opt for offering courses in English for international students, but English is not exclusive, we can also encounter this type of courses in Spanish opening and facilitating enrolments in the competitive academic environment. Spain, in this respect, plays an important role. Additionally, research in Spanish is improving and growing both in Spain and in some Latin-American countries. The exchange among scholars from Spanish speaking countries is a fact, and as such a consolidating linguistic and scientific policy is being established. Often these researchers’ level of English is not at the proficiency level required for international publications and, therefore, they may prefer Spanish for

publishing in journals and reviews, knowing that Spanish will, like English, reach a vast number of scientific communities.

The objective here is the analysis of English and Spanish academic discourses as a part of discourse analysis (explained in Chapter II) and more concretely the genre of the lecture as the most extended tertiary education practice (as discussed in Chapter III). Within English and Spanish lectures I aim at the analysis of some linguistic features that have been said to aid the understanding and retention of lectures, the Discourse Markers (henceforth DMs) (as reviewed in Chapter IV). The aim is to describe their uses and functions, and the similarities and differences that I may encounter between the Spanish and English lecture discourses.

Moreover, I have in mind specific recipients that can benefit from the findings of this study. Pedagogically speaking, both native Spanish and English lecturers can be seen as beneficiaries of this research along with tertiary education students, either native English or Spanish, who want to take university lessons in their mother tongues or in a language different from their mother tongue and in countries along the territories of North-America, Latin-America or Spain to name a few. On the one hand, English and Spanish lecturers could improve their lecture discourse-content delivery with respect to DMs and also native Spanish lecturers who want to deliver their lectures in English, or vice versa. Regarding tertiary education students, the understanding and retention of lecture delivery is crucial for their university success, therefore an improved knowledge on the use and function of DMs would surely help these students in their academic education.

The following dissertation is divided into eight chapters, four of them establishing the theoretical framework, as we have already mentioned (Chapters I, II, III and IV). The last four chapters correspond to the development of the empirical study including method, result, discussion and conclusion chapters. I will briefly explain the concrete topics dealt with in each chapter.

In Chapter I, I analyze the position of English and Spanish in the current global world and support the choice for the analysis of English and Spanish as two predominant languages around the world and two main languages within the scientific and academic settings.

In the next chapter (Chapter II) I review the discourse analysis approach and the most relevant theories which our study is based on. Centered in the analysis of some linguistic features in two languages such as English and Spanish I cannot disregard previous Contrastive Rhetoric (CR) research, where the knowledge and study of cross-cultural differences seem to be fundamental for the understanding of genres; however, most of the research carried out in CR up to now has been devoted to written forms of discourse rather than spoken discourse. Still within this chapter I expand on the genre theory (Swales 1990) in order to narrow the scope of genre towards our interests in academic discourse, and what we know as EAP (English for Academic Purposes) to end up talking about a broader concept of LAP (Language for Academic Purposes) where not only English but other languages such as Spanish can fit in. The last section of this chapter explains the growing phenomenon of corpus linguistics as a methodological approach for most genre based studies. Here in this section I also describe

and review the most significant corpora currently available, the majority already on-line.

An entire chapter (Chapter III) has been devoted to the description of the lecture as the most extended classroom academic genre. Here two basic approaches towards the lecture genre are presented. On the one hand, research into the lecture comprehension process; and, on the other hand, research into lecture discourse, the point of departure for our study and under which the role of DMs in lectures is pointed out without disregarding the pedagogical aim of the current study. The two last sections of this chapter are brief introductions to lectures in the USA and in Spain. Departing from generalizations about similarities and differences between North-American and Spanish lectures, I will discuss similarities and differences in lecturing styles and practices in the two universities where the lectures for the creation of the two sub-corpora used in this study were recorded, namely, the University of Michigan (MI, USA) and Universitat Jaume I (Castellón, Spain).

Still establishing the theoretical framework, in Chapter IV I analyze in detail the use and function of DMs as they have been studied from the 1990s onwards within corpus linguistics. We will also provide an extended revision of previous classifications of DMs as the departing point for the development of a DMs classification valid for this study. It is also at the end of this chapter, and before starting the empirical approach where the purpose of the study is stated. With the analysis of DMs in the English and Spanish lecture discourses I aim at giving insight into how DMs are used for

the benefit of both native and non-native speakers in tertiary institutions, that is:

- a) help **native Spanish/English lecturers** to improve their lecture discourse both in Spanish and in English.
- b) help and benefit **English/Spanish both L2 and native undergraduate** students for the comprehension of lecture discourse in their learning process.

Starting from these goals, I establish two research questions:

- a) Is there any difference in the use of DMs between North-American and Spanish monologic lectures in the discourse of Social Sciences?
- b) What is the relation between the several types of markers? Do some specific DMs collocate?

Next, in Chapter V the method of analysis is carefully explained, the two sub-corpora described according to their attributes and characteristics, the transcription norms displayed and the problems that arose while gathering the Spanish lecture corpus presented. At the end of this chapter we also illustrate the DM classification applied in the current study.

The next chapter (Chapter VI) is the implementation of the DM classification model and the results obtained at three levels of analysis, micro-markers, macro-markers and operators. A closing section deals with the DM collocations found in the three levels of analysis, all these sections

including a comparison of the use and function of DMs in English and Spanish lectures. The results in this chapter are presented not only quantitatively but also qualitatively.

The most important results and findings on the use and function of DMs in both North-American and Spanish lectures are discussed in Chapter VII. In the conclusion chapter (Chapter VIII) I will reflect on the pedagogical and linguistic applications that can be derived from the results obtained in the study. I also comment on the limitations of the study, which can be mostly due to the characteristics and peculiarities of the corpus under study. Some future and further research is also suggested.

The last part of this dissertation includes a list of the referenced works and three appendices where the reader can find the North-American (NAC) lecture corpus transcripts, together with the Spanish (SC) lecture corpus transcripts and the Classroom Observation Guide sample developed for the gathering of the SC.

Chapter I: English and Spanish in a Global World

1. ENGLISH AND SPANISH IN THE GLOBAL WORLD

*Las creencias que nos consuelan,
las esperanzas que nos empujan al porvenir,
los empeños y los ensueños que nos mantienen en pie
de marcha histórica la misión de nuestro destino [...],
arraigan en el lenguaje común.*

Miguel de Unamuno (cited in Moreno Garcerán 2001)

1.1. English as a lingua franca

English as an international language has been coined a ‘lingua franca’, and as such it has become the language of commerce, business and the academic world. The English language seems to be bound up in the phenomenon of globalism (Chew 1999) with a growth in number of speakers, in domains of use, in economic and cultural power (Graddol 1999). Nevertheless, the number of native speakers seems to be falling these days; this phenomenon has been named ‘the decline of the native speaker’ (Graddol 1999). Even though English has an official status in 60 countries and a prominent place in 20 more (Johnson 1996), demographically speaking, the population of people speaking English as a first language is decreasing due to the birth rate decrease in developed countries; however, this demographic figure does not directly imply a decline in the number of people speaking English around the world.

English has adopted an international status and in the near future it will be mainly used as a second language fostering communication between non-

native speakers in multilingual contexts. This phenomenon will also affect the ideological discourse about languages, linguistic competence and identity. Graddol (1999) suggests the authority of the discourse of the native speaker; he advances that with the emergence of 'New Englishes' (an L2 speech community developing its own institutionalised variety of English) together with the construction of a new Europe that places European citizens in a plurilingual context, it seems logical to question whether the authoritative norms of usage of the native speaker will prevail.

As an example we can mention the case of Singapore where native speakers of several languages (Mandarin, Tamil, and Malay) use a local variety of English as the lingua franca, that is, 'Singapore English', coined and spoken in Singapore. Thus, Chua & Chew (1993) speak about the teaching of non-Anglo-American English as an international language; this would encourage tolerance for non-native norms. Chua & Chew (1993) also point out that the majority of English language learners around the world are being taught by non-native speakers of English. This is the case in Singapore; therefore we cannot expect these teachers to teach an accurate standard form of English, such as Anglo-American or British English. As a result, the 'Singapore English' model has been accepted by its government and approved by the ministry of education as a variety of English spoken by educated members of the society such as newsreaders, academics, and teachers. This language fits standard English norms, but diversity is found in "phonological patterns, vocabulary, and the structuring of information" (Chua & Chew 1993: 53).

The concept of English as an international language does not exclusively belong to the last decade. Guided by the same idea the linguist Braj Kachru

(1984) examined the linguistic and sociolinguistic bases for international Englishes and discussed a variety of situations where differences among 'native' speakers, English bilinguals or non-native speakers could be noticed.

More recent and updated research makes a distinction between the concepts of 'English as a Lingua Franca' and 'English as a Foreign Language'. A lingua franca was originally "a variety that was spoken along the South-Eastern coast of the Mediterranean between approximately the 15th and the 19th century [...] probably based on some Italian dialects in its earliest history, and included elements from Spanish, French, Portuguese, Arabic, Turkish, Greek and Persian" (Knapp & Meierkord 2002: 9). This language was probably used as a means of communication for commerce transactions; this idea is not far from the concept of English as a *lingua franca* as it is used in the 21st century.

With time, *lingua franca* has come to mean "a language variety used between people who speak different first languages and for none of whom it is the mother tongue" (Jenkins 2004). According to this definition a lingua franca has no native speakers (NSs) and therefore no native speakers' targets for its learners. Bearing in mind this idea, the concept of lingua franca is different from the concept of 'foreign language' which does have native speakers who learn the language as their mother tongue and whose educated version is a recognised target for non-native learners mainly because they are brought up and educated in an English speaking country but sharing domain with the real home mother tongue. A distinction between English as a Lingua Franca (ELF) and English as a Foreign

Language (EFL) is then clear. Speakers of EFL use their English in order to communicate with NSs of English in mainly NS settings. Their learning goal is to come closer to the norms of a NS variety of English, normally being the variety Standard British or American English. On the other hand, speakers of ELF use their English mainly to communicate with other NNSs of English and generally in NNS settings.

Following with international Englishes, with the emergence of the European Union, there has been an increasing interest towards the analysis of an ELF variety, often called 'Eurospeak' or 'Euro-English' which, together with the Asian English variety, is leading the study of ELF varieties. However, other sub-varieties are also being acknowledged, that is, German English, Korean English, Chinese English and the like.

Research into ELF is nowadays increasing the interests of applied linguists especially researchers from the so called Expanding Circle, mainly Europe and East Asia. These scholars are gathering corpora of NNS Englishes in order to describe ELF varieties and identify systematic differences between these and NS varieties (Jenkins 2004). As an example we could mention the Viena Oxford International Corpus (VOICE) (Seidlhofer 2001), Mauranen's Corpus of English as a Lingua Franca in Academic settings (Mauranen 2003a), or the Alpine-Adriatic Corpus (James 2000).

Results from research on ELF varieties aim at a language description level. However, grammars, dictionaries and other reference works will eventually come out. Although the VOICE team have already identified a number of features that indicate systematic differences between NS English and ELF Englishes, the research on lexico-grammatical features is at an early stage,

and as Seidlholfer (2004) points out, no reliable findings based on quantitative investigations can yet be reported.

It has been largely argued that scholars advocating ELF research are against any approach meaning a single English version for the world. On the contrary, one of the main characteristics of a lingua franca is diversity, EFL speakers can preserve as much as they wish from their L1 regional accents and any Expanding Circle accent is as acceptable as an NS accent (Jenkins 2006).

As opposed to this idea there are a number of non-ELF scholars who claim for the idea of a single world standard for international communication. This would be a monolithic English largely based on Standard American or Standard British English, which has mostly often been termed as 'English as an International Language', 'World standard English', or even 'Literate English'. ELF research is finding firm opponents among a number of NSs of English who advocate the pure version of English as the world's primary lingua franca (Jenkins 2006)

These NS custodians only accept their educated English variation as correct, that is to say, Standard American or Standard British English, and do not accept the world's lingua franca to develop its own norms. In addition, as Jenkins says (2006) most English Language Teaching (ELT) publishers are still disregarding ELF development following in their ELT materials the Standard American or Standard British English. In fact, more research on ELF is needed for publishers to offer learners choices between EFL and ELF.

Within the academic field, the marginalising of NNSs' contributions to academic journals has also raised discrepancies among scholars. Contributors are urged to use NS norms of academic English. Seidlholfer points out that:

[...] questions have arisen about the legitimacy of these norms, and the extent to which written English [...] should be subjected to correction to conform to native speaker conventions of use, thus allowing journals to exert a gate-keeping function based not on academic expertise but purely on linguistic criteria whose relevance for international intelligibility has not actually been demonstrated (2004: 209-239).

Jenkins (2004) in her article *ELF at the Gate: The Position of English as a Lingua Franca* sets the example of a Chinese academic, who asks in the international journal *English Today* (2004) why she, who has never entered a native English speaking country, had to adjust her China English to one of the two main NS varieties of English in order to have her article accepted for publication.

1.1.1. English as the language of international academic exchange

The globalisation of English is a *fait accompli* in the current century. However, the question that arises is whether English hegemony will become a monopole being crucial in the fields of academic and scientific production,

as Crystal (1997) mentions, or on the contrary, there will be some kind of pluralism that, although asymmetric compared to other international languages, it will, in the long run, end up with the off-spring of some other powerful countries, as Graddol (1997) foretells in his *The Future of English?*

Crucial for the survival of any language is the power of its scientific production and the language teaching in classrooms. Scientific discovering and technology can largely be spread and well-known thanks to a single global language for communication, English. This reason has led many scientists and professionals from non English speaking communities to choose English as the unique and supreme scientific language.

Adding some statistical data relevant for our purpose, it is worth to mentioning that English speaking countries such as the U.S.A and Great Britain publish together 41% of the scientific publications (journals, reviews, etc.) around the globe. However, the figures for English scientific production in general go up to 82% in the field of social sciences and reach 90% in the field of natural science (Hamel 2002).

We should bear in mind the ‘international’ background (Benson 1994: 182) of students and what important role lectures play for their learning process success. Although most of the teaching in English has been done in English speaking countries, such as Britain and the U.S. and it is also in these countries where academic discourse started to be analysed (Mauranen 2001, Swales & Malczewski 2001, among others), we cannot deny that English is expanding to other countries as the result of internationalisation.

Globalising processes are forcing higher education institutions to internationalise. Internationalisation has now changed its scope (Wilkinson 2004), it is no longer confined solely to aspects of knowledge, it also involves extensive student and staff exchanges, the joint development and implementation of new educational programmes where two or more institutions from different countries are involved and recently, a worldwide focus on student and staff recruitment.

Due to this internationalisation of university lecturing mainly in Europe (e.g. Erasmus-Socrates programmes, Leonardo da Vinci projects), and also in the United States (e.g. postgraduate education), academic discourse in English has become a main concern in higher education institutions, focusing especially on spoken academic discourse, which should be looked at as paramount for both students and faculty. Students attending a lecture need to listen and understand first to be able to take notes. On the other hand, faculty's academic life involves not only reading English publications, but attending as well as presenting papers at conferences, even research and lecturing in other universities where English is used as the primary language. All these are instances of oral academic genres which have been classified by some authors (Giménez 2000, Bellés-Fortuño & Fortanet 2004, Fortanet 2005). Fortanet (2004b) has aimed at displaying a classification of oral genres "according to criteria of purpose, rather than to interaction between speaker and listener" and where the lecture is included within the category of Classroom genres together with the seminar, tutorial interview, students' presentation and oral exams (see Figure 1 below for a detailed classification).

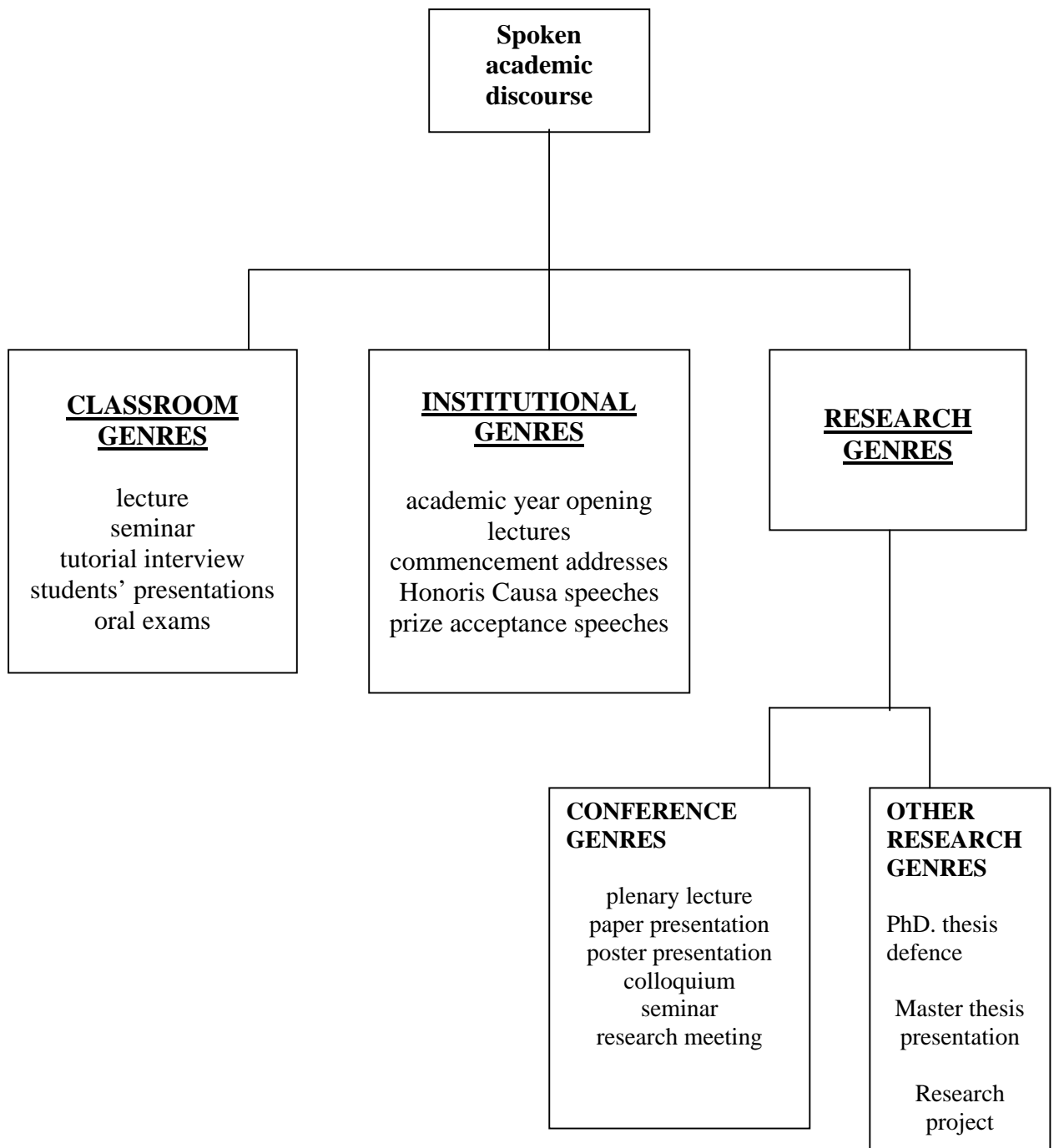


Figure 1. Classification of oral academic genres. (Fortanet 2005)

Regarding conference genres Räisänen (1999) carried out a qualitative study using a corpus of engineering conferences that she classified as Conference

Abstract (CA), the written Conference Proceedings Paper (CCP) and the oral Conference Presentation (CP). In her PhD dissertation, Räsänen highlighted the increasing mobility “over disciplinary, national and professional borders” (1999: i). Part of this scientific dissemination is given through academic conference genres.

Other academic genres that have been recently analysed are guest lectures (Crawford 2004) where the lecturers are frequently recruited from different academic, professional and cultural backgrounds “adding that ‘international flavour’ which is now an essential ingredient in the era of global education”. (Crawford 2004: 91).

Much less small-scale but very well generalised is the work of Mauranen (1998, 2001) in which English academic discourse among international academic institutions is essential. When referring to academic genres Mauranen says:

Academic speech events are commonly organised in chainlike formations either within one genre (e.g. a lecture course or a linked series of seminars) or across different genres (e.g. lectures, followed by examinations; supervision and consultations, followed by a thesis defense and its important prior text, the thesis itself) (2001: 166).

In general, classroom genres have aroused the interests of researchers, specifically the genre of lecture, being one of the most important genres within spoken academic discourse. As Flowerdew points out (1994: 14), lecture research:

[...] can indicate to teachers and course designers what linguistic and discursal features learners need to be familiar with in order to understand a lecture and what, therefore, should be incorporated into ESL courses. In addition a knowledge of the linguistic/ discursal structure of lectures will be of value to content lecturers in potentially enabling them to structure their own lectures in an optimally effective way.

Whether ELF or EFL, academic or business, in native or non-native speaking settings, English as the means of communication in the global world is an underlying reality.

1.2. Spanish as a lingua franca

With the discovery of the Americas in the year 1492 Spain's presence and therefore the spread of the Spanish language became a reality. Ferdinand and Isabel, the 'Catholic King and Queen' who united the kingdoms of Castilla and Aragón declared Castilian, *castellano*, the official language of Spain. It was then when Antonio de Nebrija published the first Grammar of the Spanish language and advised monarchs that Spanish should become the official language of the Spanish empire.

Since then the Spanish language has witnessed a large expansion far beyond the Spanish borders being nowadays spoken not only in Latin-America (Argentina, Bolivia, Colombia, Costa Rica, Cuba, Chile Ecuador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Uruguay, Perú, Puerto Rico, República Dominicana, El Salvador and Venezuela) but

also in some parts of the U.S. such as California, Florida, New Mexico, Texas, Arizona, Colorado or New York. Spanish also extends to Africa (Guinea, Western Sahara) as well as to other minor countries such as the Philippines, Israel, or some areas in Minor Asia (see Figure 2).



Figure 2. Areas where Spanish is spoken in the world.

Bearing this in mind, we can statistically prove that Spanish is the third language in the world according to its number of speakers with over 400 million, preceded by Chinese with 1,000 million speakers and English with 500 million speakers. Undoubtedly, the number of speakers of a language promotes and adds value to its linguistic community, the more speakers a language has, the greater its prestige may be. However, giving an exact number of speakers of a language is very complicated and figures are usually approximate (Décsy 1986-1988).

Although Spanish is the third most spoken language in the world, it is in fact the world's second largest international language after the English language; Spanish is also the official language in twenty-one nations being the

language spoken by a 90% of the population in those nations. Most of the population speaking Spanish lives in America; there are over 30 million speakers of Spanish in the USA (11.5% of the population) and Latin America is currently competing successfully in many industries in the world market. But Spanish is not only gaining position in the world of business and commerce. The number of students choosing Spanish is also growing; Spanish is offered as a second/ foreign language in most European countries as well as in the U.S.A where the nation's Hispanic population expands dramatically to such an extent that legislators and civic leaders are confronting new questions about how, or whether, to regulate the emergence of Spanish as a lingua franca in American life (Branigin 1999).

Spanish has a predominant place in the world and takes part in the globalisation process, even though some authors such as José Jiménez Lozano (2001) try to separate the concept of globalisation from languages stating that globalisation is fundamentally a political-economic praxis and vision. However, economy and policy are closely tied up to linguistic policy and cannot be detached from it. The economic power of a country crucially positions the language spoken in that same country around the world as proves the English language hegemony.

Regarding Spanish in the academic world, it is worth to point out that 60% of students in the U.S. choose Spanish as a second language. In this current year there will be more people speaking Spanish in the U.S. than in Spain. In Brazil where there are 180 million inhabitants, Spanish is going to be declared the second official language and a standard Spanish version will be taught in educational institutions.

El Instituto Cervantes, whose aim is to consolidate and spread Spanish by means of educational institutions where Spanish is taught, is promoted by the Spanish government, and promotes the Spanish language around the world. It is the Spanish reference *par excellence* and has already developed a learning method for Spanish as a foreign language that suits different languages and situations in every *Instituto Cervantes* centre in the world (total of 34).

Some authors agree on what the ideal features of Spanish to be learned and taught are (Marcos Marín 2001, Marqués de Tamarón 2006). Spanish is easy and coherent, its orthography is by no means dense when compared to French, it is the most widely spoken language as a mother tongue and it is unitary among its diversity, a Spanish speaker from Seville can understand without many difficulties a Spanish speaker from Buenos Aires.

Marques de Tamarón (1995) analyses the role Spanish plays in the world. In his work he predicts that in a century or so, a 90% or 95% of the 6,000 languages currently spoken in the world will disappear. The standardisation of the mass media will lead to the decline of minor languages, while only stronger or the most widely spoken ones will survive. Those languages technologically advanced will be able to compete in a world dominated by networks, oral interfaces and all kind of developed linguistic software. Spanish is nowadays very well positioned due to its number of speakers, its cultural and geographical extension.

The larger the presence of Spanish in the academic world, the longer its international prevalence would be. Along this line, Del Castillo (2001) explains that by raising the number of situations in which Spanish is present

in both scientific and technological academic research, and by linking Spanish to new technologies, we will encourage the learning of Spanish and the spreading through publishing houses (Del Castillo 2001).

1.2.1. Spanish as the language of international and academic exchange

*Como son las palabras que conservan las ideas y las transmiten,
no se puede perfeccionar el lenguaje,
sin perfeccionar la ciencia, ni la ciencia sin el lenguaje.*

Cited in Hamel 2002

As commented has a privileged position in the world and also as a scientific language. As opposed to what seems the unstoppable hegemony of English in the scientific and academic fields, Spanish is the second international language and has all the conditions needed in order to be present in strategic international contexts such as international relations, business, commerce and the scientific world.

Authors such as Hamel (2002) postulate the idea of a plurilingual model in the field of sciences, especially the plurilingual nature of the social sciences where the Spanish language would be preserved and enhanced as one of the vital slots of science. The model presents two main reasons for this positioning:

1. The reduction to a unique and exclusive language in the scientific and academic research environments would lead to a hazardous impoverishment of the scientific field, especially that of the social sciences.
2. English hegemony would emphasise already existing differences in the access to international science as well as in the spread of our own scientific production. It goes without saying the important value of science as a means of production and how the abandoning of these fields by a speaking community would weaken in the mean or long run the economy of a country.

Up to now we can find that in Spain 85% of the scientific journals, reviews, etc. are issuing articles in Spanish; there is also a large production of scientific books in the field of social sciences supplying the vast Hispanic market (Spain, Argentina and Mexico to name a few). This is due to the prestigious and well recognised Spanish literary work which knows no barriers in the global world.

However, Spanish is also a preferred second and foreign language chosen by learners in educational institutions around the world maybe attracted by Spanish cultural issues.

There is then a communicative need departing from economic reasons and the phenomenon of globalization where Spanish is the second language internationally used.

Spanish is taught around the world as a second or foreign language sharing some problems with the English language. Spanish teachers and lecturers

face a dilemma. They hesitate about which Spanish to teach in the classroom, that is, his/her own native Spanish, the Spanish used in films and TV or ‘Disneyland Spanish’ (Moreno 2000: 79-82). Whichever Spanish is taught, it will strongly depend on students’ needs and idiosyncrasy, the truth is that the teacher/lecturer will definitely bring the most of his/her native model.

Spanish teachers and lecturers around the world know that the language is a powerful communicative tool and as such students should be provided with all the communicative features of the Spanish language such as “las características fonéticas, morfosintácticas y léxicas de un modelo de español general, o internacional [...] porque con una lengua internacional no existen fronteras lingüísticas” (López 2003: 30).

At a tertiary educational level Spanish as an international academic language gains importance with the great amount of Latin-American students who come to Europe to take master or doctorate courses. Up to now most European universities opt for offering courses in English for international students, but English is not exclusive, we can also encounter this type of courses in Spanish opening and facilitating enrolments in this competitive academic environment. Spain, in this respect, plays an important role. Many students from Latin-America prefer a Spanish university because they do not have problems with language production and understanding. Additionally, research is also improving and reaching a high level both in Spain and in some Latin-American countries. The exchange among scholars from Spanish speaking countries is a fact, and as such we are creating a consolidating linguistic and scientific policy.

Often the researchers' level of English is not at the proficiency level required for international publications and therefore, they may prefer Spanish to publish in journals and reviews, knowing that Spanish will, as well as English, reach a vast number of scientific communities. Additionally, academic exchanges with non-Hispanic scholars and/or researchers have to give these researchers the opportunity to learn and/or improve their knowledge of Spanish as well as to enlarge scientific production in Spanish.

In the present chapter I have tried to show why English and Spanish are two very widespread international languages, usually but exclusively being a first choice for international students who want to learn a language different from their mother tongue. Researchers around the world seemed to be concerned about the spoken language and its different varieties; as a consequence some considerations about the existence of Englishes as opposed to single unique English spoken around the world, or even the existence of different Spanishes, mainly distinguishing between Peninsular and Latin-American Spanish have been argued. However, I believe that for the development of new Englishes or Spanishes the learning of a standard basis of the language that should be seen as the norm is linguistically speaking quite necessary, the other Englishes or Spanishes would be deviations of somehow standardised norms. This is the reason for the development of this study I have analysed native speakers' versions of English and Spanish. In this way, by analysing spoken English and Spanish linguistic elements and investigating the kinds of discourse practices of

English and Spanish university settings, I could better help both lecturers and tertiary education students.

The internationalisation of the university implies an internationalisation of its lecturers. As far as I know there are some Spanish universities that have started projects in order to offer core subjects lectured exclusively in English, although the lecturer may often be a native Spanish. Therefore, Spanish lecturers are required to be better prepared in English not only to be able to deliver lecture sessions in English but also to develop their personal research in English (conferences, journal publications, etc.). With this, my interest and focus is upon Language for Academic Purposes (LAP).

In the present study some Spanish and English language features are observed and analysed at a tertiary level education, in this way I will be able to work upon some academic language generalisations on how English and Spanish discourses, as two of the main and most important world global languages are used and structured. In the following Chapter I will review the most relevant studies on academic discourse carried out up to now and the paradigms used.

Chapter II: Discourse Analysis

2. DISCOURSE ANALYSIS

It would be nice if we could squeeze all we know about discourse into a handy definition. Unfortunately, as is also the case for such related concepts as 'language', 'communication', 'interaction', 'society' and 'culture', the notion of discourse is essentially fuzzy.

T.A. van Dijk 1997: 1

2.1. Introduction

In this chapter, I offer a review of the most important theoretical approach in the field of discourse analysis, especially research into academic discourse..

The object under study in the present work is academic discourse and what we broadly understand here as Language for Academic Purposes (LAP). As has been argued in Chapter I, English is now the *lingua franca* of a global world, and as such it is more and more often used as the language of instruction in educational institutions as a result of the globalisation of education. The teaching of English is mainly addressed to the preparation of non-native speakers (NNSs). The analysis of native English academic discourse allows researchers to use it with pedagogical purposes.

Regarding Spanish in the global world, it is gaining a position in the sectors of business and commerce. The number of students is also growing, Spanish being offered as a second language in most European countries as well as in the U.S.A where the nation's Hispanic population is expanding dramatically

to such an extent that legislators and civic leaders are confronting new questions about how, or whether, to regulate the emergence of Spanish as a lingua franca in American life (Branigin 1999).

Some of the earliest analyses of academic discourse date back to the 1960s and focus on the quantitative study of the formal feature of language varieties, or registers (Barber 1962; Halliday, Strevens & McIntosh 1964). After these first studies, the work done in academic discourse has become “narrower and deeper” (Swales 1990: 3). According to Flowerdew narrower “in the sense that it has focused on specific genres and deeper in so far as it has sought to investigate communicative purposes, not just formal features” (2002: 2). This work tries to fit into this description and it can be said to be narrow, since we centre our attention on the lecture genre, and deeper because it tries to analyse the communicative purpose of discourse markers within spoken academic discourse.

The research tradition of genre analysis has been framed within English for Specific Purposes (Hyon 1996, Yunick 1997) drawing special attention to Hallidayan ideas about the relationship between language and its social function. Seminal in this approach are the studies carried out by Swales 1981, 1990; Dudley-Evans and Henderson 1990a, 1990b; Henderson and Hewings 1990; Bathia 1993 and Skulstad 1996, 2002, among others. Swales’ work, although focused on EAP, provides a definition of genre mainly based on the relationship between text and discourse community:

A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert member of the parent discourse community,

and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and styles.(Swales 1990: 58)

We can also find in the field of EAP Tarone et al.'s early study (1981), "On the use of passive in the astrophysics journal papers"; which is significant as both a deeper and narrower approach; deeper because it focused on the communicative value of a syntactic feature, the passive, and narrower because it analyzed this feature as it was contextualized within one particular genre. This earlier approach to genre analysis culminated years later with the studies of Swales (1990) and Bhatia (1993). Swales and Bhatia worked on genre analysis, the former in the field of academic discourse, the latter in business, academic and legal genres always with pedagogical purposes. The result has been the genre analysis approach and the very much cited Swalesian Move-Step analysis.

Very influential in North-American academic writing has also been the New Rhetoric School (Miller 1984; Bazerman 1988; Yates & Orlikowski 1992; Freedman & Medway 1994a, 1994b; Berkenkotter & Huckin 1995) where context is emphasized. These paradigm fundamentals are based mainly on the theories developed by Foucault (1970, 1980) and Vygotsky (1978, 1986), among others, where the focus is on L1 teaching setting and studying disciplines such as rhetoric, composition studies and professional writing.

Along with the genre analysis work, other discourse analysts have been working in different approaches although with the same pedagogical purposes in mind. Some of the approaches relevant for the present research

are contrastive rhetoric and corpus linguistics. The latter, mainly used as a methodological approach, has been devoted one section in this chapter.

Contrastive rhetoric is the study of the similarities and differences between two languages and how the influence of the L1 may affect the way individuals express themselves in the L2. This approach was firstly defined and adopted by Kaplan (1966) founding a large school (Clyne 1987, Connor 1996, Hinds 1983, Mauranen 1993, Ventola 1992).

I will now discuss the Contrastive Rhetoric approach and later in the next section I will move towards Corpus Linguistics as a methodological approach in the field of discourse analysis. These two approaches are probably among the most influential in academic discourse analysis and have had direct pedagogical applications to the tertiary educational level which is the aim of our study.

The objective is analyzing the lecture as an academic genre by means of corpus linguistics as a methodological approach and from the contrastive rhetoric point of view as we compare two different language corpora (English and Spanish).

Therefore, a detailed description of the latest work carried out in academic discourse based on the paradigms described above is presented in the following sections.

2.2. Genre analysis

There have been different stages from the earliest analysis of linguistic data to the concept of genre analysis. In this latest approach the study of social structures, social identities and discourse systems is important. The attempt to study and analyze descriptions of language use has been known as genre analysis, where according to Bathia an effort “is made to offer a grounded description of language use in educational, academic, or professional settings.” (2002: 21)

Already in the 1980s Swales established the general value of genre analysis as a means of studying spoken and written discourse for applied ends. He was convinced that a genre-centred approach could offer a workable way of making sense of the myriad of communicative events that occur in the contemporary English-speaking academy (Swales 1990).

Later, Bathia (2002) introduces the different language description stages and how these are based on the way to relate texts to contexts. Language can be described as text, genre, or social practice (see Table 1 below). Along these three descriptions an ideal positioning would be somewhere in the middle, looking at the use of language as a genre in context without a pure grammatical grounding. This would keep a balance between the study of linguistic form and the study of context in a broad sense of socio-cultural factors.

<p style="text-align: center;">Language description</p> <p style="text-align: center;">as</p> <p style="text-align: center;">Text</p> <p style="text-align: center;">What features of lexico-grammar are statistically and/ or functionally distinctive?</p> <p style="text-align: center;">Context: narrowly configured in terms of textual links</p> <p style="text-align: center;">Genre</p> <p style="text-align: center;">Why do we use the language the way we do and what makes this possible?</p> <p style="text-align: center;">Context: more specifically configured in terms of disciplinary cultures</p> <p style="text-align: center;">Social Practice</p> <p style="text-align: center;">How do we relate language to social structures, social identities, and social practices?</p> <p style="text-align: center;">Context: broadly configured in terms of socio-cultural realities</p>

Table 1. Bathia's (2002) stages of language description.

Bearing the use of language as a genre in mind, Bathia has defined Genre Analysis as:

[...] the study of situated linguistic behavior in institutionalized academic or professional settings, whether in terms of *typification of rhetorical action*, as in Miller (1984), and Berkenkotter and Huckin (1995); *regularities of staged, goal oriented social processes*, as in Martin, Christie, and Rothery (1987), and Martin (1993); or *consistency of communicative purposes*, as in Swales (1990) and Bathia (1993) (2002: 23).

Genre analysis as a view of language use has some features considered common ground; Bathia (2002: 23) describes them as follows:

1. Genres are reflections of disciplinary cultures and, in that sense, those of the realities of the world of discourse, in general.
2. Genres focus on conventionalized communicative events embedded within disciplinary and professional practices.
3. All disciplinary or professional genres have integrity of their own, which is often identified with reference to textual and discursive (text-internal) factors, or contextual and disciplinary (text-external) factors. However, it is not always fixed and or static but often contested, depending upon the rhetorical context it tends to respond to.
4. Genres are recognizable communicative events, characterized by a set of communicative purpose(s) identified and mutually understood by members of the professional or academic community in which they regularly occur.
5. Genres are highly structured and conventionalized constructs, with constrains on allowable contributions in terms of the intentions one can give expression to, the shape they can take, and also in terms of the co-grammatical resources one can employ.
6. Established members of a particular professional community will have a much greater knowledge and understanding of

generic practices than those who are apprentices, new members, or outsiders.

7. Although genres are viewed as conventional constructs, expert members of the disciplinary and professional communities are often in a position to exploit such conventions to express 'private intentions' within the structures of socially acceptable communicative norms.

The main feature obtained from this view of language is that genres are communicative events that are structured and conventionalized and common to a particular professional community and culture. However, these conventionalized constructs are open and sensitive to deviations and therefore not static but dynamic. In this respect Berkenkotter and Huckin point out that:

Genres are inherently dynamic rhetorical structures that can be manipulated according to conditions of use, and that genre knowledge is therefore best conceptualised as a form of situated cognition embedded in disciplinary cultures (1995: 6).

This dynamic side of genres is explained by Bathia (2002) in the terms of socio-cognitive needs. For him genres do change over time in response to socio-cognitive needs or 'private intentions', although requiring communities to negotiate their response in the light of recognizable or established conventions. These established conventions can be manipulated

or new ones can be generated, which gives what Bathia calls ‘tactical freedom’ to members of the same community.

As generic practices are widely known by established members of a particular professional community, genres can significantly overlap as is the case of research article introductions (Swales 1981, 1990), abstracts (Bhatia 1993), textbooks (Myers 1992), among many others. However, it is true that these genres are also tied to disciplinary variations (Biber 1988, Fortanet et al. 1998, Hirvela 1997, Holmes 1997, Bathia 1999, Hyland 2000). According to Bathia (2002) these variations seem to be more significant when related to lexico-grammatical resources and rhetorical strategies, these two concerns being really important in the evolution of ESP and EAP creating tension between pedagogic convenience and pedagogic effectiveness.

It is precisely the field of Academic Purposes what frames this study. Traditionally English for Academic Purposes (EAP) has been the standard approach in response to the English communication needs of tertiary level students in the academy. However, in the early 1990s aspects related to EAP programs were based on experiences rather than on any systematic research. However, and according to Swales (2004), over the past dozen years or so there have been a surprising number of PhD. theses and dissertations in the area of EAP as well as a growing accumulation of research findings, which have been published in book chapters and in journals such as *English for Specific Purposes*, *English for Academic Purposes*, *Applied Linguistics*, and *Written Communication*.

Within the broad concept of EAP all areas of academic concern such as teaching (undergraduate and postgraduate text books, lectures, seminars, academic discussions), research (articles, conferences, academic research reports, research grant applications, etc.), examination (examination papers and answers, project reports and theses, essays, and other written work) are covered. Sometimes academic administration issues (course and program description) can also be included. Going one step further within EAP one can find the concept of English for Research Purposes (ERP) in an attempt to better understand the research world and its discursal products and processes (Swales 2004). As far as we know, such a broad investigation in Academic Purposes has not been developed in the Spanish language; however, the settings in Spanish tertiary education are similar to the areas of academic concern we have just described above. Not having such an extended literature in the Spanish language within the register of academic discourse has made me discuss and adopt the concept of Language for Academic Purposes (LAP) in an attempt to embrace both English and Spanish academic discourse issues.

In addition to genre analysis there is also another approach gaining impetus among discourse studies researchers, contrastive rhetoric (CR). The success of CR according to Swales (1990) has been a built-in assumption that discourse is indeed both socially situated and designed to achieve rhetorical goals. In the following section I review the origins of CR and describe the most relevant studies.

2.3. Contrastive Rhetoric

Studies on the basis of contrastive rhetoric have been up to now devoted to the teaching of reading and writing to foreign students (Kaplan 1984, 1987, Enkvist 1997, Connor 1996, 2002, 2003, Atkinson 2003, Moreno 2004). This research perceives a close and needed relationship between cultural differences and the different teaching approaches. However, we need to understand the concept of pure rhetoric to establish a connection between rhetoric and second language teaching approaches. Authors such as Oliver (1965) defined and described the concept of rhetoric almost forty years ago:

Rhetoric is a mode of thinking or a mode of “finding all available means” for the achievement of a designated end. Accordingly, rhetoric concerns itself basically with what goes in the mind rather than with what comes out of the mouth... Rhetoric is concerned with factors of analysis, data gathering, interpretation, and synthesis (1965: 10-11).

If rhetoric is ‘a mode of thinking’, and each culture has its own and intrinsic ‘mode of thinking’, cultural variation should be taken into consideration in second-language teaching. Nevertheless, as far as I am concerned the connection has only been limited to the level of grammar, vocabulary and sentence structure. Scarce attention has been paid up to now to the level of discourse and as a result not to spoken discourse.

Some studies comparing the Indian and the Western languages (Sapir 1912, Whorf 1956) noticed the language diversity and established the Sapir-Whorf hypothesis of linguistic relativity , that is, the Whorfian hypothesis, upon

which the origins of Contrastive Rhetoric is based. The hypothesis suggests that different languages affect perception and thought in different ways. It asserts that one's native language influences and controls thought, therefore barring fluent second language acquisition. However, the Sapir-Whorf hypothesis has been largely criticized by linguists and psychologists; Fishman (1977), a sociologist, pointed out the large number of bilinguals who have no problem switching between the grammar and lexica of their languages, questioning thus the Whorfian hypothesis. Several psychologists have also criticized the hypothesis positioning themselves for or against any of these two versions. The first version, the weaker one, maintains that language influences thought; the latter, called the stronger version, states that language controls both thought and perception. Following with language diversity, Dufrenne claimed that this diversity:

affects not only the languages, but also the cultures, that is to say the whole system of institutions that are tied to the language [...] and language in its turn is the effect and the expression of a certain world view that is manifested in the culture. [...] The types of structures characteristic of a given culture would then, in each case, be particular modes of universal laws (1963: 35-37).

Analyzing this postulate in detail, our attention has been drawn by the concept of institution, considering that the university as such is a higher education institution 'tied to the language' spoken in its community, being the language spoken the expression of the prevailing culture. As a consequence, cultural variation, language diversity and the different

educational institutions where a language is spoken and taught seem to come together.

Kaplan (1984: 43-62) reinforced the Whorfian view that each language imposes a world view on its users, and mentioned the link between logic and rhetoric, he understood logic in the popular sense, rather than the logician's sense of the word, and claimed that not only spoken language but also logic and rhetoric are culture specific; he wrote:

Logic [...] is evolved out of a culture; it is not universal. Rhetoric, then, is not universal either but varies from culture to culture and even from time to time within a given culture. It is affected by canons of taste within a given culture at a given time (p. 44).

Kaplan (1984) referred then to the sequence of thought in English as a Platonic-Aristotelian sequence, descended from the philosophers of ancient Greece and shaped subsequently by Roman, Medieval European, and later Western thinkers. This sequence of thought takes the form of a concrete language pattern. He goes on claiming that we cannot assume that "because a student can write an adequate essay in his/ her native language, he/ she can necessarily write an adequate essay in a second language" (p.44). Kaplan sets up the example of those instructors/ lecturers/ teachers that, having corrected foreign students' term papers, theses and/ or dissertations, have written comments such as: "The material is all here, but it seems out of focus", "Lacks organization", "Lacks cohesion". Presumably the student paper is out of focus "because the student is employing a rhetoric and a

sequence of thought which violate the expectations of the native reader” (Kaplan 1984: 44).

In fact, the relationship between sequence of thought and grammar in a given language is not something new; even ancient rhetoricians as Cicero or more recent ones such as Brooks and Warren have indicated the link between thought sequence and rhetoric, these last two claim:

In his *Cultural Thought Patterns in Intercultural Education*, Kaplan (1984) carried out a comparative analysis among seven hundred student compositions written by students from different nationalities and countries whose mother tongue was other than English. After disregarding one hundred of these for not being significant and representative enough, six hundred examples, representing four basic language groups, were examined. Departing from the English sequence of thought which is dominantly linear, he found that in the Arabic language, paragraph development is based on a series of parallel constructions, both positive and negative; such a development would strike an English reader as archaic or awkward. This Arabic paragraph sequence could be, according to Kaplan, extended to all Semitic languages.

The second language group was the type of Oriental writing. This is marked by a somewhat directional approach, in Kaplan words the development of the paragraph may be said to be “turning and turning in a widening gyre.” The circles turn around the subject but it is never looked at directly. Romance language writing was analyzed with samples from French and Spanish students. Kaplan observed that there is one thing which characterizes Romance language paragraphs: digressions. French and

Spanish ways of writing tend to digress or to introduce extraneous material more often and freely than English writing. Finally, Kaplan observed the complexities of Russian paragraphs. The structure of the Russian sentence turned out to be completely different from the structure of the English sentence. Russian paragraph structure is made up of a series of parallel constructions and a number of subordinate structures or “parenthetical amplifications” which are not sometimes important to the central idea (p.51) Kaplan represented his findings graphically as shown below:

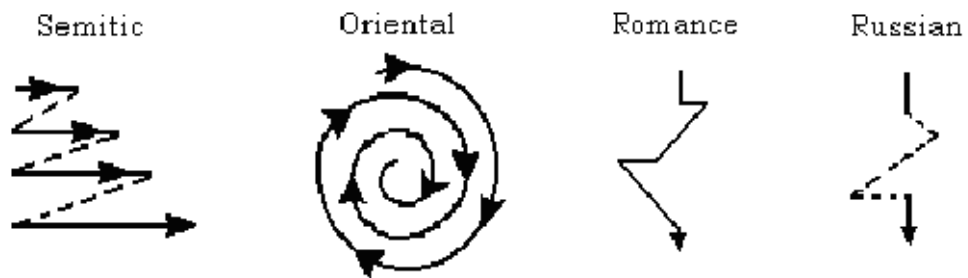


Figure 3. Cultural Thought Patterns in Intercultural Education (Kaplan 1984).

With the results of his investigation Kaplan revealed that “each language and each culture has a paragraph order unique to itself, and that part of the learning of a particular language is the mastering of its logical system” (p.51). This reveals the pedagogical side of the study; language teaching to foreign students whether in terms of reading and writing must take into consideration these paragraph sequencing differences not only for the teacher him/ herself but in order to present this in the classroom and make students become aware of it. In Kaplan’s words: “[...] contrastive rhetoric

must be taught in the same sense that contrastive grammar is presently taught” (p.52).

However, Kaplan emphasis relies on the teaching of reading and writing composition but he fails to mention spoken discourse as a means of language communication that is present in language classrooms as much as reading or writing. I believe that contrastive rhetoric is not only a vivid element in the sequence of paragraphs across cultures as a reflection of the sequence of thought. Contrastive rhetoric is also present in spoken discourse in classroom settings being also a consequence of the diversity of thought patterns across cultures.

Kaplan’s Contrastive Rhetoric was the first in the field of language acquisition that focused only on the rhetoric of writing, studying differences in writing styles across cultures. However, Kaplan’s Contrastive Rhetoric lacked some of the elements included in traditional Aristotelian rhetoric where invention, memory, arrangement, style and delivery were taken into account. Liebman (1992) criticized Kaplan’s approach (1966, 1972) since it reduced rhetoric to one element: arrangement or organization. Going one step ahead, Kaplan in his 1972 volume adopted a limited but popular view of rhetoric as understood in English-speaking countries where discourse was classified according to four elements: description, narration, argumentation and exposition. However, the major component of classical rhetoric, persuasion, was absent from this classification and in fact substituted by argumentation. In 1971, with the publication of Kinneavy’s book, *A Theory of Discourse*, persuasion was included as one of the four major concerns of discourse. The result was the development of new approach within

contrastive rhetoric. Kaplan's 'traditional' contrastive rhetoric has also been criticized by some authors in the last 20 years. Matalene (1985) for example commented on the extremely ethnocentric profile of Kaplan's CR as well as on the fact of privileging the writing of native English speakers, Raimes (1991) criticized Kaplan for considering transfer from a first language a negative influence on second language writing and Mohan and Au-Yeung Lo (1985) for examining only L2 products and ignoring educational and developmental process variables. Even Kaplan himself has a number of more recent publications (Kaplan 1987, 1988) where some new postulations are added to the original 1966 article as it is the idea that rhetorical differences do not necessarily reflect different patterns of thinking, but they may reflect different writing conventions that are learned in a culture.

The fact is that very little or scarce research has been done in the area of contrastive rhetoric in the field of spoken discourse across cultures. Most of the research carried out up to now follows the premises given by Kaplan (1984) thirty years ago and most studies have been devoted to the analysis of writing and reading influenced by cross-cultural aspects.

Contrastive rhetoric has been defined as "an area of research in second language acquisition that identifies problems in composition encountered by second language writers and, by referring to the rhetorical strategies of the first language, attempts to explain them. [...] Contrastive rhetoric maintains that language and writing are cultural phenomena." (Connor 1996). However, through the last decade the study of applied linguistics has seen the light of new trends which highlight the influence of discourse analysis going far beyond the sentence level and taking into consideration different

speech patterns of speakers. It is precisely the offspring of these new trends in linguistics where the basis of the present study lays, being contrastive studies paramount for the understanding of cultural particulars as well as linguistic universals as Péry-Woodley points out:

Contrasting and comparing are basic to any form of anthropological investigation, and this includes of course linguistic investigation. It is the contrastive light which shows a particular practice as specific to a group; conversely, it is the contrastive approach which allows the identification of universals. Not only is a contrastive stance a superlative way of gaining precise descriptive knowledge about individual languages and cultures, it is invaluable in the quest for a general understanding of language-based communication, [...] (1990: 143).

In the same way we find a ‘new rhetoric’ (Perelman & Olbrechts-Tyteca 1969) that focuses on the achievement of a particular effect on the audience as opposed to classical rhetoric; there has been a development of Kaplan’s ‘traditional’ contrastive rhetoric resulting in a new extended approach of contrastive rhetoric built on Kaplan’s which spreads towards interdisciplinary boundaries.

Most research carried out on CR up to now, which is based on Kaplan’s initial study, is concerned with students writing as a reflection of traditional rhetoric. However, trying to center the scope of teaching towards university classroom settings and more concretely to the United States’ university system we find complementary perspectives on traditional rhetoric where

instruction in oral presentation is part of university courses (Russell 1991). In his *Writing in the academic disciplines, 1870-1990: A curricular history* (1991) Russell reminds about how rhetoric courses in the United States University included among other things exercises in which original speeches and essays were delivered orally. However, this difficult task for teachers would soon be replaced as the university standards changed towards a text-based teaching discipline. As Russell points out; “the requirements shifted further away from oral performance and toward text-based, discipline specific scholarship” (1991: 60).

CR studies on written texts are many; however, to our knowledge, little or scarce research has been carried out on the study of spoken academic text. Spoken academic discourse seems to be more and more relevant for university settings not only in the States but also around Europe, where university students claim on mobility to other countries where English would be the language used in nearly all communicative acts (spoken and written), the university way of teaching, that is to say, the lecture genre is changing to a more communicative approach (Benson 1994) where students are asked to participate and enroll classroom discussion. Only a few rhetoricians have carried out spoken language research up to now by means of examining a small number of easily identifiable surface structures such as the individual morphemes (Dulay & Burt 1974). Although syntactic issues in writing have been the most studied issue in CR, new trends in which this study is framed upraise to compare discourse structures and genres.

In the present study I emphasize the importance of spoken academic English within the academic world, taking the lecture genre as the model to be

analyzed and examining some cross-cultural differences which may be found between the North-American and the Spanish lectures such as lecture sequencing or discourse coherence.

As Connor explains, CR in the context of applied linguistics “is taking new directions in five domains” (Connor 1996: 19). These five domains are: contrastive text linguistics, the study of writing as a cultural activity, contrastive studies of the classroom dynamics of L2 writing, contrastive rhetoric studies conducted in a variety of genres in a variety of situations for a variety of purposes and contrastive rhetoric studies dealing with the inculcation of culturally different intellectual traditions and ideologies. If we were to frame our study within one of these five domains, not only one could be chosen. The present study can be included within contrastive text linguistics. However the definition of text linguistics has raised some discrepancies among European researchers, taking into consideration the Finnish text linguist Nils Enkvist (1984) who points out that the term ‘text’ can refer to both spoken and written language. As we analyze spoken language comparing discourse features across languages, this study can be framed in the domain of contrastive text linguistics concerned with the processes that speakers and listeners (instead of readers and writers) go through to produce and comprehend texts. However, we also aim at analyzing a specific genre (lectures) in a concrete situation (Spanish and North-American undergrad classrooms) considering the approach of CR studies in a variety of genres and situations for a variety of specific purposes.

There have been previous classroom-based contrastive studies similar to the one carried out here and in which classroom talk is linked to culture, for example, and to mention one, the sociolinguists Scollon and Scollon (1981) studied cross-cultural interaction between Athabascan Native Americans and native English speaking North-Americans finding differences in the amount of talk as well as in the roles speaker and listener adopt. Athabascan for example, expects the speaker to adopt a more dominant role than native English speaking North-American do. This type of studies has proven the importance of cultural variation in the classroom setting taking into account behavior and conversational patterns. I will now review some relevant CR studies in the frame of the Spanish language.

2.3.1. Contrastive rhetoric studies in Spanish

In this section I aim at reviewing only the most significant studies on Contrastive Rhetoric in Spanish in order to highlight generalizations about rhetorical patterns of language. For an exhaustive reference list on CR studies, Silva's (1993) work *Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications* includes reports involving twenty-seven different languages from a large number of unpublished dissertations and research reports dealing with the rhetoric of writing.

Spanish contrastive rhetoric studies date back to the 70s (Santiago 1970, Santana-Seda 1970). However and bizarrely these first empirical studies

have not been gathered in CR anthologies (Kaplan 1983, Connor & Kaplan 1987, and Purves 1988).

In the year 1988, Reid's dissertation compared the English writing of 184 native Spanish speakers with the English writing of native Arabic, Chinese, and English speakers using an essay-writing task on the Test of English as a Foreign Language. Her findings showed that Spanish L1 writing included longer sentences and a higher use of pronouns, as compared to native English speakers' writing, showing preference towards 'loose coordination'. Montaña-Harmon (1988, 1991) studied variations among four groups of students: Mexican secondary school students writing in Spanish, recently arrived immigrant Mexican-American secondary school students in the United States, Spanish L1 students grown up in the States and native English-speaking students in the U.S. Results demonstrated that Spanish speaking students wrote longer sentences and fewer simple sentences, in contrast to the English speaking students. The native English speakers American students "used simple vocabulary, few synonyms, and no flowery language" (1991: 424).

Later, Reppen and Grabe (1993) demonstrated the differences between Spanish and native English-speaking writing styles. They analyzed a total of 545 texts (234 Spanish L1 and 311 English L1) from which findings corroborated previous studies' results, that is, how the phenomenon of language transfer from Spanish L1 affected students when writing in English, "particularly with regard to the use of elaborate style".

All studies presented so far deal with Latin American Spanish and not Spanish written in Spain. Spain Spanish CR studies have covered different

features (e.g. proverbs, Ibáñez & Ortigosa 2004) and written genres (e.g. hotel brochures, Cortés de los Ríos & Cruz 2004). Within the register of academic discourse the work of Moreno (1996, 1997, 2004) has been seminal in CR, although once more the focus is on written modes of discourse. She analyzed several discourse features such as causal coherence and premise-conclusion sequences in the academic genre of research articles (RAs) by means of comparing English and Spanish RAs within business and economic disciplines. She aimed at analyzing the values of similarity or disparity which characterize the particular rhetorical context of the RA claiming that there is significant intercultural variation in the rhetorical preferences of national cultures (Moreno 1997). With her work she opened a way for applied contrastive studies that wish to explore the origins of the phenomenon of the linguistic transfer through quantitative methods in the fields of EAP and translation (Moreno 2004).

Spanish-English rhetoric is currently gaining interest among CR practitioners concerned with a wide genre diversification. Unfortunately, the focus is still on written text (writing patterns and styles) rather than on spoken text. Generally, CR studies are carried out by means of comparative corpora which have been gathered for a specific study. The corpus gathering technique has given name to corpus linguistics as one of the most often used linguistic methodological.

2.4. Corpus linguistics: a methodological approach

An approach gaining impetus in discourse studies is corpus linguistics. It is concerned with the collection, structuring, and analysis of a large amount of discourse, always with the assistance of computers. Computer processing allows operations of quantifying (counting the number of given words or sentences), concordancing (producing lists of linguistic items and their immediate linguistic context in order to determine syntactic, semantic and pragmatic properties), parsing (separating sentences into grammatical parts) and labeling (syntactic analysis) or tagging parts of speech. Studies based on this approach, although not focused specifically on academic discourse, are the work of Sinclair and Collins Cobuild project (Sinclair 1987) and the one of Biber (1988) in register analysis.

Several book collections on corpus linguistics have been published both by European and North-American linguists with a wide impact (Aijmer & Altenberg 1991, Simpson & Swales 2001, Leistyna & Meyer 2003, among others). Aijmer and Altenberg have defined corpus linguistics as “the study of language on the basis of text corpora” (1991: 1). They established the beginnings of corpus linguistics with the development of two events that occurred around 1960s. One was Randolph Quirk’s (1960) launching of the Survey of English Usage (SEU) with the aim of collecting a large varied corpus of spoken and written English. The other was the advent of computers which made it possible to store, scan and classify large bodies of data.

More recently other authors have defined corpus linguistics essentially as a technology (Simpson & Swales 2001) that implies the use of a computer that can store a collection of text files and then apply software to those files to produce frequency lists, lists of key words, and strings of words showing which words co-occur with others. However, this has not always been so, although nowadays corpus linguistics is linked to computerized corpora, there was a time previous to the computerized era where corpora were not computerized. A pioneer linguist who gathered and studied oral non-computerized corpora in North-America was Charles C. Fries who wrote about the necessity to study spoken language:

In the meantime, however, beginning in 1946, it became possible to obtain an entirely different kind of evidence. Instead of the letters collected and studied for the American English Grammar (1940) I procured the means and the opportunity to record mechanically many conversations of speakers of Standard English in this North Central community of the United States. Altogether these mechanically recorded conversations amounted to something over 250,000 running words (1952: vii-viii).

However, nowadays we cannot understand corpus linguistics but as an interaction between the human mind and the machine (computer). Leech (1991) pointed out that successful analysis depends on a division of labor between the corpus and the human mind. There is a truly interactive relation between analyst, software and corpus as Figure 4 below shows.

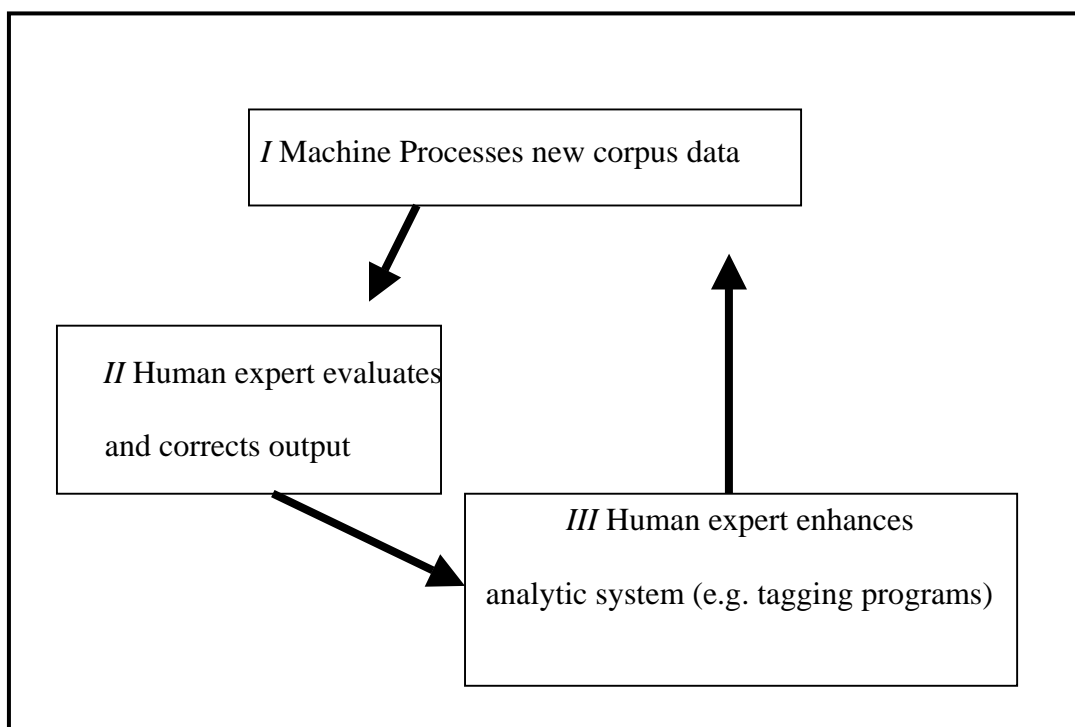


Figure 4. A 'symbiotic' relation between human and machine in corpus processing (Leech 1991).

Over the last 20 years the developments in corpus linguistics have been the focus of European as well as North American researchers. As to European researchers, there are particularly active groups in the United Kingdom and Scandinavia as the prolific publications and creation of several corpora show. Take for instance the COBUILD project (earlier Bank of English project) started in the year 1991; it covers over 450 million words of British English (spoken and written), a specialized corpus devoted to single genres. Another important European corpus is the BNC (British National Corpus), a 100 million word collection of samples of mainly written but also spoken

English language from a wide range of sources which was also started in the year 1991. It also has to be pointed out the existence of the ICE (International Corpus of English) project which began in 1990 with the main aim of collecting material for comparative studies of English worldwide. The British component of the corpus was published in 2001 with 1 million words. At the moment 13 research teams are participating in this project following Greenbaum's (1991) initial idea. Also important is the CIC (Cambridge International Corpus); started in 1992, it holds 600 million words but this time including British and American English discourse. Within the ICI we can find corpora such as the British CANCODE (Cambridge and Nottingham Corpus of Oral Discourse in English) or the CAMSNAE (Cambridge Corpus of Spoken North-American English). The primary aim of the ICI is to help in writing books for learners of English and is only available to authors and writers of books working for Cambridge University Press. The result of a more recent project is the BASE (The British Academic Spoken English Corpus) developed at the University of Warwick under the leadership of Hilary Nesi and Paul Thompson; it consists of 160 lectures and 40 seminars recorded in a variety of university departments, this corpus aims to be a counterpart of the North-American MICASE (Michigan Corpus of Spoken Academic English); I will talk about this corpus below.

Although not so actively as European linguists, some North-American scholars have made developments in corpus linguistics as it is the case of Douglas Biber et al. at Northern Arizona University with the Longman Grammar of Spoken and Written English (1999) (the majority of co-authors

being European, though), Michael Barrow at Rice University and other corporist groups at The University of California or the University of Pennsylvania. Regarding North-American English corpora, the ANC (American National Corpus) stands out. It was started in 1990 and includes a massive electronic collection of American English including texts of all genres and transcripts from spoken data that is still in progress. When completed, the ANC will contain a core corpus of at least 100 million words, comparable across genres to the BNC described above. Undoubtedly important in the growth of corpus linguistics within the EAP field in North-America is MICASE (Michigan Corpus of Spoken Academic English). It started in 1997 at the University of Michigan and was designed to digitally audio-record and transcribe about 1.8 million words of contemporary academic speech at a major American research university, covering speech events ranging from freshman advising to doctoral defenses along with traditional university lectures and class discussion.

The study undertaken here uses the MICASE for the North-American English corpus as a starting point in the field of EAP. To date, there is some relevant research within corpus linguistics that has dealt with the field of EAP and the analysis of spoken language (Swales & Malczewski 2001; Mauranen 2001, 2003b; Biber, Reppen, Clark & Walter 2001; Swales & Burke 2003; Biber 2003; Flowerdew 2003), however little research has been exclusively focused on spoken monologic lectures and the use of discourse markers as this study does. In the next section I will review relevant research about the lecture genre within the LAP field.

Chapter III: The Genre of Lecture

3. THE GENRE OF LECTURE

*The research article has commanded the greatest amount of attention among academic discourse analysts [...].
The lecture, as a genre, while still pre-eminent in undergraduate education, remains relatively neglected.*

J. Flowerdew 2002: 5

3.1. Introduction

Teachers involved in the field of academic study have a wide range of instructional material available, namely speech events such as seminars and tutorials; materials such as videos; or activities such as writing assignments, among others; but the lecture “remains the central instructional activity” (Flowerdew 1994). Waggoner (1984) characterises lectures as having “paradigmatic stature” and other authors such as Benson (1994) define lecture as “the central ritual of the culture of learning”. Lecturing is a widely accepted practice in higher education in American institutions and throughout the world (Dunkel & Davy 1989). However, lectures are not homogeneous. The lecture class is changing (Waggoner 1984), so that traditional methods of learning coexist with newer interactive methods; both lecturers and students feel the influence of a greater egalitarianism than in former times. Students see teachers at a closer distance and the role of a helper, a counselor or a facilitator for the learning process better fits their perspectives. On the other hand, teachers seem to invite students to interact and participate more than in previous times, what may be understood as an attempt to narrow distances and avoid formalisms.

A great part of university discourse research focuses on the lecture (Johns 1981, Richards 1983, Benson 1989) and more specifically on the lecture comprehension process. Knowing the best way for students to internalise and comprehend lecture content seems to be paramount for university success; that is why there is some research on spoken academic language centred on different aspects of lectures (Flowerdew 1994, Chaudron & Richards 1986, Thompson 1994, Jones 1999, Khuwaileh 1999, Kerans 2001).

In this section we are going to review two main approaches within lecture research. On the one hand, research into the lecture comprehension process will show the relevance of the *how* of teaching and learning in relation to lectures. On the other hand, research into the lecture discourse will show the relevance of the *what* of teaching and learning. This last aspect will shed some light on the design of ESL (English as a Second Language) courses as well as provide lecturers with valuable information in order to structure their own lectures in an effective way.

3.2. Research into the lecture comprehension process

Research into lecture comprehension process has been said to be the *how* of teaching and learning (Flowerdew 1994), and therefore being of great help in the field of applied linguistics. Lecturing methodology has benefited from research on the lecture comprehension process; both lecturers and learners can make the most of their teaching and learning knowing how the listening

comprehension process and distinctive features of lectures work. Next, we present relevant studies undergone on the lecture comprehension process.

3.2.1. Phonological and lexico-grammatical features in lectures

Spoken text has its own lexico-grammatical features, which require the application of particular sets of knowledge on the part of listeners (Biber 1988). Regarding phonological features in lectures, native or non-native students must recognize unit boundaries phonologically, irregular pausing, false starts, hesitations, stress and intonation patterns. These features are particularly hard to be identified by non-native speakers that might not have been exposed to lecture speech. Lectures, as part of oral academic discourse, can become a handicap to non-native speakers who have learned English in a much more traditional written style, not being used to spoken discourse. Moreover, listeners are also presupposed to have “the ability to distinguish what is relevant to the main purpose and what is less relevant” (Flowerdew 1994: 11), following the theories of relevance of Grice (1975) and Sperber and Wilson (1986). Lecture listening comprehension implies functions of interactive discourse such as asking for repetition, negating meaning or using repair strategies. Whenever the lecturers allow questions from the audience or ask questions themselves, turn-taking conventions have also to be considered (Flowerdew 1994).

3.2.2. The use of micro-skills and strategies

Some researchers have identified a set of micro-skills which are assumed to be necessary for the comprehension of lectures in a second language. The first of these micro-skills taxonomies was the one proposed and designed by Munby (1978), becoming a departing point in any needs analysis and course design. Based on Munby (1978), Richards (1983) provided a second taxonomy much more closely related to academic listening. Within Richard's list of 18 skills we find relevant information for the purpose of this study as DMs seem to be considered valuable reference for the lecture listening comprehension process. This taxonomy includes among others:

- i) the ability to identify topic of lecture and follow topic development
- ii) the ability to recognize the role of discourse markers for signalling the structure of the lecture, and
- iii) the ability to recognize the function of intonation to signal information structure (e.g. pitch, volume, pace, key)

Regarding the role of strategies within the research of L2 lecture listening comprehension, it might be thought that if students handle the relevant strategies for listening and note-taking, they will be able to apply them to the various lectures they encounter. Olsen and Huckin (1990: 42) suggest that students in Engineering need to be taught appropriate strategies to understand the “discourse-level pragmatics of academic lectures”. They state that students have to be conscious of the difference in the strategies

needed to understand two different ways of presenting lectures. On the one hand, for a lecture that aims to present information students will need the information-driven strategy, that is, a strategy that allows students to distinguish the most relevant information in a lecture. On the other hand, a mode of lecture that builds an argument based on a number of points will require a “more context-sensitive point-driven strategy” that allows students in the field of Engineering, but also in other disciplines such as the Humanities and Social Sciences, to understand a problem-solving lecture. For Benson (1994: 193) strategies (linguistic, cognitive or social) “represent the upper, often observable and sometimes teachable layer of a learning culture which also contains a lower, hidden layer of unspoken values, assumptions and beliefs”. He later on argues that listening skills and strategies “are a necessary but no sufficient part” of a listening exposure. The lecture comprehension process can never be unconscious to learners. The present study aims at analysing some distinctive discourse features that help lecture comprehension, that is, discourse markers. However, they constitute just a tiny piece of the whole puzzle, which can in no case be enough for the successful comprehension of lectures. Olsen and Huckin (1990: 33) point out that students: “may understand all the words of a lecture (including lexical connectives and other discourse makers) and yet fail to understand the lecturer’s main points or logical arguments”. Therefore, the use of strategies is relevant for the comprehension of lectures, both for teachers and learners.

3.3. Research into lecture discourse

As mentioned above in this project, research into lecture discourse has primarily had effects on the *what* of teaching and learning, providing information about the linguistic and discoursal features of lectures. The sections presented below refer to those linguistic and discoursal features of lectures that have been of valuable help for both lecturers and learners in the learning process of content lectures. We will revise the different types of lectures classified by authors such as Goffman (1981) or Dudley-Evans and Johns (1981); I will also look into how syntax is presented in lectures as a type of spoken text. Next, the organization of lectures will be examined and some of the structural patterns proposed shown (Cook 1975). Finally, we aim at presenting a thorough review of the role DMs have played in lectures by examining previous research on signalling devices (Cook 1975, Murphy & Candlin 1979, Chaudron & Richards 1986).

3.3.1. Lecturing styles

Several studies have identified a number of lecture styles. Morrison (1974, reported in Jordan 1989: 153), studied science lectures and divided them into two kinds: i) formal and ii) informal. The former refers to “close spoken prose”, and the latter is defined as “high informational content, but not necessarily in high formal register”. This first classification, although somehow useful, is too simplistic. More complete classifications of lecture

styles are those proposed by Goffman (1981), Dudley-Evans and Johns (1981) and Dudley-Evans (1994). The study points out that: “the key to the understanding of lectures is an appreciation of lecturers’ individual styles” (Dudley-Evans 1994: 148). Goffman (1981) talks about three *modes* of lectures, namely, ‘memorization’, ‘aloud reading’ and ‘fresh talk’, whereas Dudley-Evans and Johns (1981: 134) distinguish three styles:

- i) *reading style*, “in which lecturers either read the lecture or deliver it as if they were reading it” (Dudley-Evans 1994: 148). It is characterised by short tone groups, and narrowness of intonational range, in which falling tone predominates; although level tone may also occur
- ii) *the conversational style*, “in which lecturers deliver the lecture from notes and in a relatively informal style with a certain amount of interaction with students” (Dudley-Evans 1994: 148). It is characterised by longer tone groups and key-sequences from high to low. When the lecturer is in ‘low-key’ at the end of a key sentence, the speaker may markedly increase tempo and vowel reduction, and reduce intensity, and
- iii) *the rhetorical style*, “in which the lecturers give a performance with jokes and digressions” (Dudley-Evans 1994: 148). It is characterised by the wide intonational range. The lecturer often exploits high key, and a ‘boosted high key’. There are frequent asides and digressions marked by key and tempo-shift,

sometimes also by voice quantity-shift. (Dudley-Evans & Johns 1981: 134).

We may establish a parallelism between the classifications proposed by Goffman (1981) and Dudley-Evans and Johns (1981). The ‘reading style’ in Dudley-Evans and Johns’ (1981) classification could be similar to Goffman’s ‘aloud reading’. Goffman’s ‘memorization’ resembles ‘conversational style’; and ‘fresh talk’ could be compared to Dudley-Evans and Johns’ ‘rhetorical style’. Quoting Goffman (1981: 165), conventional lectures are:

[...] institutionalized extended holdings of the floor in which one speaker imparts his views on a subject, these thoughts comprising what can be called a ‘text’. The style is typically serious and slightly impersonal, the controlling intent being to generate calmly considered understanding, not mere entertainment, emotional impact or immediate action.

There is no written evidence about the frequency of use of lecture styles, but there seems to be a general agreement on identifying the informal conversational style (based on notes or handouts) as the predominant mode of lecture presentation not only for native, but also for non-native audiences (McDonough 1978, DeCarrico & Nattinger 1988, Dudley-Evans 1994). Along this line, Frederick (1986) talks about a “participatory lecture” closer to discussion. More recent work (Benson 1994) perceives a move towards a more interactive style of lecturing. This trend seems more predominant in

the United States universities rather than in European ones; this fact may cause some problems for non-native speakers that have been trained in a much more traditional lecturing style, producing a default or precarious listening comprehension. Problems of a cultural nature, the role and status of university lecturers, degree of deference between lecturers and students or pure lecture content problems are the ones students may encounter. It is evident that the social norms of a lecture vary according to cultures, for example, in some Asian countries lecture attendance is regarded as more important than interaction; whereas in the United States the student who interacts whenever allowed is more appreciated than the mere spectator student (Benson 1994). Depending on the position lecturers adopt students may feel comfortable or at a loss: international students, for example, may well be alarmed with the figure of an egalitarian concept of teacher, a lecturer as a 'facilitator' and 'guide' (Benson 1994), expecting a more traditional concept of lecturer as a more authoritative and demanding person, as it could be the case of Japan (Nakane 1970: 59).

3.3.2. Syntax of lectures

Regarding the syntax of lectures, these, as a type of spoken text, might be seen as characterized by typical spoken syntactic features rather than by written features (Tannen 1982, Halliday 1985/ 89). Far from this idea, Biber (1988) points out that there is no single parameter of linguistic variation that distinguishes spoken and written texts. Instead, we find what he names

dimensions, that is to say, clustering of features which work together to fulfil some underlying function within the various spoken and written genres, namely formal/ informal, restricted/ elaborated, contextualised/ decontextualized, and involved/ detached. According to these characteristics, spoken texts can sometimes be informal, restricted, contextualised and involved. Nevertheless, different types of spoken texts may vary their characteristics. In this sense, lectures as formal and strictly planned speech events are thought to share many of the features of written texts, although this is not always so.

There are several mechanisms in spoken discourses which facilitate learners' comprehension. The use of linguistic repetition, as an example, plays an important role. Giménez points out that in ancient discourses, whether social or pedagogical “no sólo se enseñaba la repetición lingüística sino que también se utilizaba como mecanismo pedagógico y didáctico” (1998: 302). Some research considers the linguistic repetition to be relevant as a means of cohesion and global structuring of the discourse (van Dijk et al. 1972, Hymes 1981, Ventola 1987). In a recent study on lectures, Giménez (2000) analyses the effects of linguistic repetition on the academic genre of lecture within the discourse of Social Sciences. Giménez (2000) presents stated evidence of the importance of linguistic repetition in the genre of lecture for a logical understanding.

3.3.3. Lecture structural patterns

The structuring and organization of a lecture plays an important role for the listening comprehension process. However, compared to other genres such as the research article (Swales 1990), very little research has been carried out on this aspect. The main interest takes into consideration those aspects of lecture structure that might be relevant in training non-native speakers. Thus, much of the research done in the 70s and early 80s examines how the information organised in a lecture is signalled (Cook 1975, Montgomery 1977, Murphy & Candlin 1979, Coulthard & Montgomery 1981). Cook (1975) distinguishes two structural patterns within a lecture: the macro-structure and the micro-structure of a lecture. The macro-structure of a lecture is made up of a number of 'expositions'. An exposition consists of different classes of episodes, namely, an optional episode of expectation, an obligatory focal episode, an obligatory developmental episode together with optional developmental episodes, and an obligatory closing episode. Within the micro-structure, episodes are described in terms of 'moves'. Just to illustrate this, a concluding move is a justificatory statement, a focal episode with a concluding function, or a summary statement. A summarizing move gives a summary of the immediately preceding discourse. Cook's (1975) attempt describes the boundaries of these units but fails to give detailed information about their internal structure.

More recent work by Young (1994) tries to "describe the macro-structure of university lectures and to identify some of the more prominent micro-features that contribute to this structure" (1994: 159). Young departs from

some research on the macro-structure spoken monologic discourse. For the development of the study she gathers a corpus made of seven two-hour university lectures from third and fourth year courses. Three are lectures delivered to non-native speakers of English from a Western European university on disciplines such as Soil, Physics, Sociology and Economics. The other four lectures were delivered to native speakers of English in North American Universities. This specific selection of the corpus could give an identifiable macro-structure across disciplines and across levels. Young describes the macro-structure of a lecture in terms of 'strands' or 'phases'. She distinguishes six phases split in two groups: three metadiscoursal strands which comment on the discourse itself; and the other three which mark university lectures. The first three metadiscoursal phases proposed by Young (1994: 166) are:

- a) *Discourse structuring phase*: addressors indicate the direction that they will take in the lecture.
- b) *Conclusion*: where lecturers summarize points they have made throughout the discourse.
- c) *Evaluation*: the lecturer reinforces each of the other strands by evaluating information which is about to be, or has already been transmitted.

The two former metadiscoursal phases are more frequent than the latter. The phases which mark university lectures are (Young 1994: 167):

- a) *Interaction*: indicates an important feature of this registerial variety.
- b) *Theory or Content*: used to reflect the lecturer's purpose, which is to transmit theoretical information.
- c) *Examples*: it is in this phase in which the speakers illustrate theoretical concepts through concrete examples familiar to students in the audience.

In her study, Young comes to the conclusion that a more accurate schema of university lectures is presented when using phases rather than presenting the macro-structure of a lecture in terms of outlines (Woods 1978: 42), where the structure of a lecture is understood as beginning, middle and end configuration. As Young indicates (1994: 173), when referring to the genre of lecture: "phasal analysis seems to offer a more realistic portrayal of the nature of this particular genre".

3.3.4. The role of discourse markers in lectures

Following the lecture discourse literature, many researchers have suggested that an understanding of the role of discourse markers and the relationships between different parts of the text is fundamental for the comprehension of lectures (Morrison 1974, Coulthard & Montgomery 1981, Chaudron & Richards 1986). In his work, Cook (1975) examines the functions of connectives, which serve as indicators of topic continuation. Other authors

identify a number of markers of the rhetorical organization of lecture discourse (Murphy & Candlin 1979).

Several studies have examined features of discourse organization. Chaudron (1983) in an early study analysed the effects of topic signalling in experimental lectures on ESL learners' immediate recall of the topic information. Sawa (1985) studied two factors in recorded lectures: repetition and paraphrasing of information, and signalling of major segments and emphasis. In a later study, Chaudron and Richards (1986) investigated the effect of pragmatic signalling devices on comprehension. They provided four groups of subjects with samples of four versions of the same lecture. The first version included no signalling devices at all, the second one included some, as Chaudron and Richards call them, 'micro-markers' (lower-order markers linking clauses and sentences). The third version contained, as they coin them, 'macro-markers' (higher-order markers marking major transitions), and the last version was a combination of versions 2 and 3, that is to say, it included a combination of macro and micro-markers. Four different groups of subjects were tested after listening; the main findings showed that macro-markers proved to be "more conducive to successful recall of the lecture than micro-markers" (Chaudron & Richards 1986: 122). Previous text studies (Kintsch & Yarbrough 1982) had also pointed out that subjects are better able to answer gist and main-idea questions for texts that contain evident rhetorical cues (discourse markers) than for texts that, although having the same content, do not include evident rhetorical cues. Kintsch and Yarbrough (1982) findings showed that a combination of a transparent rhetorical organisation and the

presence of evident rhetorical cues help the global comprehension and recall of information. The rhetorical cues proved to activate appropriate rhetorical schemas and provide a way of structuring the content of incoming information.

A more recent lecture-oriented input study carried out by Dunkel and Davis (1994) tested the differences between the lecture information recall of first-language listeners and second-language listeners relative to the presence or absence of rhetorical signalling cues (discourse markers) in the discourse. Subjects were provided with two forms of the same lecture: one form contained explicit cues indicating the rhetorical structure of the lecture (evident form), in the second form these cues were deleted (non-evident form). Two groups of native speakers of English and two groups of non-native speakers of English listened to either the evident and non-evident form of the lecture. Subjects were tested after listening; the results showed no statistical difference between the language proficiency and rhetorical cuing variables, that is to say, the presence of the rhetorical signalling cues (discourse markers) had slight impact on the proportional number of words written in the protocols. In fact, subjects that listened to the non-evident form recorded more words in their protocol than those who received the evident form containing the signalling cues. Dunkel and Davis' findings seem to contradict previous studies on the effects of discourse markers in the comprehension process. Whereas previous researchers (Kintsch & Yarbrough 1982, Chaudron & Richards 1986) think that the listener benefits from the presence of signalling cues in discourse messages, Dunkel and Davis' work found no support to this argument if the materials and

procedures used in their study (recall protocols) are proved to be valid for measuring comprehension and retention of lecture information. As Dunkel and Davis (1994: 68) state:

[...] unlike Chaudron and Richards who used cloze, true-false, and multiple choice tests, we used a different measure, written recall protocols (Bernhardt, 1983). It is possible that this difference in method caused our results to diverge from those of Chaudron and Richards.

They point out a limitation in their study: the lack of a verification of subject's prior background knowledge of the content of the lecture. They suggest that further research is needed of the interaction of text type (content and structure) with signalling devices.

3.4. Lectures in the U.S.A.

3.4.1. Introduction

As already reviewed there is a wide literature on the study of lecture genre in order to be able to differ from various lecturing styles, linguistic features affecting lecture delivery, etc. However, we also have to take into consideration assumptions about lecture culture and how these may differ significantly. These cross-cultural differences have been up to now part of the study of CR as I revised in the previous chapter. Students attending a lecture will need to take into consideration cross-cultural differences in

order to select which listening strategies to use for a better comprehension of the lecture. Lecturers giving a lecture will also have to take into consideration cross-cultural lecture delivery characteristics, especially a lecturer who is a non-native speaker of the language.

There is this idea in the United States and Canada that lectures are becoming more interactive (Benson 1994, Swales 2002, Salehzadeh 2006) and therefore the students will probably have to adapt to lecture comprehension strategies they had used in the past and adjust them to the new lecture culture. Equally, lecturers that want to deliver a lecture in a foreign country will have to adapt their lecture delivery strategies to the new culture.

Especially in the United States, there seems to be a trend for lectures to adopt a more 'open' lecture style, as Swales (2002) has defined it. This style is characterised by lecturers not reading anymore from a lecture written text, rather, they speak from notes or an outline. Therefore, much of what they say is constructed on the spot. According to Salezahdeh (2006) this 'open' lecture style has some advantages for the listeners and the speakers since they both have a more direct connection. However, there are also some disadvantages such as for example the interactive nature of the lecture, which can be difficult for a listener to follow or the use of more ungrammatical phrases.

Regarding other socio-cultural differences that we might encounter in the United States tertiary classroom we can think of some non-professional behaviour as compared to Asiatic tertiary education idiosyncrasies. In the United States most professors wear casual clothes, or they can even drink while lecturing, or might sit on top of a table or desk as they talk. These

attitudes would be considered as rude and impolite lecture practices in other countries, for example in China.

Many professors in the United States are likely to address their students by their first names, and they may encourage students to address them by their first names. With this North-American lecturers attempt to balance the relationship and avoid the tension that could arise between professors and students. As for the student audience in the United States, most lecturers will accept students eating and drinking in the classroom while in countries such as Great Britain or Australia only drinking but not eating would be acceptable. North-American tertiary education students can also adopt some relaxed comfortable postures in the classroom, students might put their feet on the chair in front of them while they take notes, and these practices could be considered very rude in other countries. However, in the North-American lectures setting a student sleeping in class or talking with a friend during a lecture is not accepted and considered very rude, unless they are talking about the content of the lecture.

From the lecture corpus in this study; I have observed that the North-American lectures within the monologic lecture style tend to be longer than the Spanish lectures gathered (Tables 9 and 10 in Chapter V) not being this fact connected to the number of students, that is, to whether the lecture is large or small. This might be linked to the type of subject being delivered; probably core subjects with a stronger theoretical content are prone to be longer than free electives with practical credits.

Undoubtedly, these socio-cultural features have to be taken into consideration when we are attending (as a listener) or delivering (as a

lecturer) a lecture in the United States. Now, I will approach the University of Michigan as the model for the North-American lecture corpus.

3.4.2. The example of the University of Michigan

(U of M) (U.S.A.)

The University of Michigan is one of the best-known public research universities in the United States and offers degrees in practically all fields except for agriculture and veterinary medicine. It has about 36,000 students (data from 2001), about 10,000 of whom are graduate students. The university is highly successful in obtaining research grants and produces about 700 doctoral graduates a year. About 70 percent of its undergraduates are from Michigan, the remainder being from other states or international.

The percentage of graduate students who are non-native speakers of English varies greatly from one department to another, being highest in Science and Engineering and lowest in departments like English. Considerable efforts have been made in recent years to increase the representation of minorities (especially African Americans and Hispanics) within the student body, although these efforts are being challenged at court at present.

The university is a member of the Big Ten, a consortium of other flagship public universities in the Midwest plus the private North-western University. The university's main campus is in Ann Arbor, a fairly large university town of 120,000 situated in the south-east corner of the state.

Within this university setting the ELI (English Language Institute) at the University of Michigan started in 1997 the MICASE (Michigan Corpus of Academic and Spoken English) project, becoming one of the most useful English spoken academic corpora used nowadays by corpus linguists. MICASE has been restricted to a single institution that seems to have a profile that differs in degree rather than in kind from that of many other major public research universities. Extrapolations to other broadly comparable U. S. academic contexts cannot, of course, definitely be made but can certainly be suggested.

3.5. Lectures in Spain

3.5.1. Introduction

The lecture as a genre has not been as widely studied in Spanish settings as the English lecture. To my knowledge only a few Spanish linguists have approached the genre of lecture or its translation equivalence in Spanish *clase magistral*; however, some pedagogy and psychology researchers have attempted to analyse the Spanish lecture within the teaching psychology field (De la Cruz 1996; Doménech 1999, 2004; Aldana 2003). Doménech (1999) speaks of a tertiary education instructor who has to accomplish two main tasks as an excellent researcher and teacher, although these two traits can imply completely disparate things. Doménech proposed the MISE (Modelo Instruccional de Situación Educativa) tertiary education

instructional model taken from Rivas (1993, 1997). The MISE model is basically based on different instructional/ psychological theories of information processing. As I am dealing with linguistic theories this model is not applicable for this study. Other authors have considered the Spanish *clase magistral* (lecture) as a very widely known teaching method usually accepted by Spanish tertiary education students who consider the lecture as a very well structured and clear synopsis. Spanish students seem to dislike very digressive lectures and with those that imply the reading of a book considered as the textbook (Aldana 2003). Aldana (2003) also points out that some lecturers prefer the monologic lecture because they believe their students are not mature enough to undertake an autonomous learning process.

As a result, we can say that the lecture or *clase magistral* remains the most extended teaching practice within Spanish universities, especially for subjects that imply a large proportion of theoretical content, more likely to be used within the fields of Humanities and Social Sciences rather than in the Engineering Sciences.

According to the Spanish lecture corpus gathered for the purpose of this study and what I observed, I could say that Spanish lectures tend to be less interactive than the North-American ones. However, we can appreciate an evolution in the Spanish tertiary education which is in continuous change; the lectures based on the *magister dixit* are no longer used except for some scarce old-fashioned supporters. Spanish students are prone to a more active, Socrates-like style, in which they feel as main characters as opposed to mere spectators. As Gutiérrez pointed out: “El profesor seguirá

enseñando meras técnicas, pero deberá recibir el influjo positivo de los alumnos” (2003: 75).

Obviously, we cannot forget cultural differences affecting the genre of lecture and that may arise between countries. As mentioned before for the North-American ‘open’ lecture styles (Swales 2002), the informal, casual-style behaviour is also present in the Spanish university classroom, and lecturers do normally wear casual clothes and may also take something to drink while delivering the lecture. In fact, they might really need it if the lecture is too long. However, within the Spanish lecture setting it would be considered very rude if a student would take something to eat in the classroom, or adopt a too relaxed posture in the classroom, which can be very impolite and unpleasant practices. Sleeping or talking in the lecture classroom is also seen rude within the Spanish university context, apart from disturbing the lecture delivery in the case of chatting to a friend during a lesson. Regarding the use of first names, this is quite common among Spanish tertiary students and lecturers, who with this, aim at shortening distances for a better communication, without denying the lecturer his/ her leading position.

The Spanish lectures gathered at Universitat Jaume I (Spain) for the purpose of this dissertation are slightly shorter than the North-American lectures taken from the MICASE, maybe due to the type of lecturing, being the North-American lectures more interactive and ‘open’ as they can expand for a longer time, whereas if it is the lecturer who mainly monopolises the floor, it may happen that the lecturer’s exhaustion forces the lecture to be shorter.

In the next section a brief profile of Universitat Jaume I on the East coast of Spain and where the Spanish lecture corpus was recorded is given.

3.5.2. The example of the Universitat Jaume I (Spain)

The Universitat Jaume I is one of the five public research universities of the Valencian Council in Spain. It was founded in the year 1991 and has about 13,000 students, a number that allows a closer and more personal attention to their students. The university offers 28 degree courses which are under constant assessment and improvement. All degrees offer students the possibility to realize an internship in previously agreed institutions and companies once they are in the upper-division. Statistics show that one every four students are recruited by the company/ institution after they finish their work placement.

The Universitat Jaume I is the pioneer in the development and adaptation of the European harmonisation process started with the Bologna agreement, adapting pilot programmes to the degrees offered in the three faculties and taking new challenges for the successful integration of the ESHE (European Space for Higher Education).

Since the very beginning in the early 90s, the Universitat Jaume I has also been the pioneer in the integration of ICTs in the university classroom, being the first Spanish public university in having its own Internet server (www.uji.es) and the first one in getting a center for new technologies and

education known as CENT (Centre d'Educació i Noves Tecnologies). Since then multimedia teaching has become a reality in the classroom, all classrooms being provided with audiovisual and computer equipment for the development of lessons. Moreover, the Universitat Jaume I has recently signed the Aula Virtual pilot project opening a new horizon towards on-line education and lecturing styles.

The Universitat Jaume I has signed more than 200 agreements with other Spanish but also foreign higher education centers in Europe, the United States and Latin-America. International exchange students from all over the world come to study at Universitat Jaume I through Socrates-Erasmus grant programs, and other European programs such as Tempus, Alfa or Leonardo da Vinci, all these programs managed by the International and Educational Cooperation office at Universitat Jaume I.

The Spanish corpus consisting of twelve Spanish lectures has been recorded within these university settings, constituting a pioneer project for the study of Spanish lectures and the starting point for the Ministry project entitled “Análisis de las características pragmáticas y lingüísticas del inglés y el español académico hablado de los géneros docentes y de investigación” (HUM2004-02599/ FILO) held by some members of the GRAPE (Group of Research on Academic and Professional English) at Universitat Jaume I.

Chapter IV: Discourse Markers

4. DISCOURSE MARKERS

*[...] as a teacher of English, my own experience
leads me to suppose that students can get a better handle
of communicative affairs by concentrating,
at least, initially, on the sui generis features
of particular genre texts.*

J. M. Swales 1990: 18

4.1. Introduction

In the previous chapter I have reviewed different aspects concerning research into lecture discourse. It has been said that a conscious knowledge of lecturing styles and different syntactical elements can potentially aid the lecture comprehension process (Flowerdew 1994). Discourse markers are part of that gathering of linguistic features enhancing and fostering successful lecture comprehension. Thus, from the 70s onwards researchers into the lecture comprehension process whether in L1 or L2 have pointed out the effectiveness of learning about discourse markers for the comprehension of connected discourse (Cook 1975, Murphy & Candlin 1979, Kintsch & Yarbrough 1982, Chaudron & Richards 1986).

Discourse markers have been the focus of many studies, gaining importance from the 70s onwards. At that time, Text Linguistics began to co-exist with new theories as Pragmatics. Therefore, the boundary of sentence (Bloomfield 1933, Chomsky 1970) was trespassed to span a wider and longer space, the text. Within this frame the approaches to DMs started to change and were described as having two different representations or meanings: the grammatical and the pragmatic meaning (Newmeyer 1980).

Already at the beginning of the 20th century Charles Morris (1938) had conceived the Theory of Signs departing from three disciplines: syntax, semantics and pragmatics. Pragmatics rises then to differentiate between what we say and what we interpret from what has been said. At an earlier stage, communication was seen as a sentence statement coding and decoding process (Saussure 1916, Jakobson 1960). Later, Grice (1975) postulated that we can communicate much more than we say and that there is a new element in the communicative process called ‘implicature’, verbal communication consists of two parts, one that is coded and the other one which is the result of inferences, that is, mental processes that lead to understanding. As a consequence, human communication is an inferencing process; decoding is not enough for the completion of verbal communication. As applied to the topic of the present work some DMs mean to have the ability to relate the discourse member or segment in which they appear with a previous element or segment in the discourse:

(1) Ana es de Teruel y, *por tanto* es habladora. (Portolés 1998: 30)

The Spanish DM *por tanto* (therefore) makes us look for another member in the previous discourse, in this case it is “Ana es de Teruel [...]”, this is an example of discursive deixis. *Por lo tanto* can help us infer a consequence for the first segment or member of the discourse; however, the use of the DM *por lo tanto* is not enough to get the right implicatures as Llorente (1996) criticises. Rather, the DM *por lo tanto* is functioning here as a linguistic unit that gives cohesion to a discourse. Although DMs are aimed

at giving cohesion, this is not their aim but a result of their use. DMs cannot therefore be considered as mere cohesive devices. Thus, we can think of the use of DMs in human communication not with the unique aim of producing cohesive messages, but with the aim of communicating the best way possible to make the hearer/reader understand our message. This idea of cataloguing DMs as ‘cohesive links’ was previously taken by Llorente (1996). She criticises Portolés’ earlier work (1993) pointing out that even though DMs are not only used to produce a cohesive piece of text but to enhance inferences in the communication process, she discusses that coherence and inferences are the same thing, we get meaningful communicative acts as the result of the links units have in discourse (Llorente 1996: 26).

I agree with Llorente (1996) and her postulates, we cannot consider a text cohesive without getting the meaningful inferences for the communicative process taking into consideration other aspects such as, settings, discourse context, genres or cross-cultural differences among others. We definitely need to develop other larger scope strategies for the ongoing of the communicative act, in van Dijk and Kintsch’ (1983) words, these would be ‘global strategies’.

Take the example showed above (1), *por tanto* is a DM that gives a causal semantic meaning and relation in a text. This causal relation cannot be established unless we get the message “Ana es de Teruel”, we infer that people from Teruel are all very talkative but that might be far from being true. The lexico-semantic relation the DM *por lo tanto* expresses cannot be understood unless we get some background inferences from the message.

Following with the marker *por lo tanto* and taking an example from the Spanish lecture corpus in the present study we find:

- (2) Efectivamente, desde luego, se produce una retroalimentación, si se consigue el éxito en esas campañas militares tendremos otros territorios de los que obtener recursos para seguir financiando nuevas campañas militares, *por lo tanto*, tenemos ya la pescadilla que se muerde la cola.
(LE1/ SC)

In this example the marker *por lo tanto* could be considered as a DM classified as *conector* (Llorente 1996) since it gives a logico-semantic relation among elements in the discourse, that is, causal meaning. However, if we look for the former member in discourse to establish the logico-semantic relation, we realise that the meaning here is not as clear as with the previous example “Ana es de Teruel”. The reason is we might be missing some inferences. The lecturer establishes the causal relation between elements in the discourse with a Spanish idiom “*por lo tanto* tenemos la pescadilla que se muerde la cola”. Unless we are native Spanish or get a broader vision of the context (discourse), we will probably not understand this metaphorical expression. The activation of global strategies is here fundamental and therefore this example taken from spoken academic discourse shows those aspects such as settings, discourse context and cross-cultural differences have to be taken into account in order to get a successful communicative message. The marker *por lo tanto* per se cannot make the message meaningful unless we take the notion of pragmatic inferences.

According to Grice, an implicature is a clause (proposition) conveyed implicitly by an utterance (sentence). Conversational implicatures are implicit propositions (clauses) which can be inferred from what is said, based on pragmatic principles. However this is not always the case, sometimes implicatures are conventionally encoded in a particular linguistic expression or DM rather than inferred. The following example shows the use of the linguistic unit *on the other hand*.

- (3) My brother-in-law lives on a peak in Darien; his great aunt, *on the other hand*, was a nurse in World War I. (Grice 1989: 362)

For Grice, what is 'said' by the speaker is only that his/her brother-in-law lived on a peak in Darien, and that the brother-in-law's aunt was a nurse in World War I. There is, however, another implicit proposition that is conveyed by the linguistic expression *on the other hand*. *On the other hand* expresses in some way contrast between the two explicitly conveyed propositions (clauses). Whether the linguistic unit *on the other hand* is considered to be meaningful or non-truth-conditional (Shourup 1999) will depend on context. To be successful with the inferencing process, context is a sine qua non; context in verbal communication is always mental and made of beliefs that are kept in the mind of the participants and their immediate perceptions (Sperber & Wilson 1986).

As a result, human communication is essentially inferential, although there are some linguistic units or expressions whose meaning definitely conditions discourse processing in relation to context. These units are what we call here DMs.

Considering the broad investigation in DMs, there seems to be convincing evidence that listeners (L1 or L2) benefit from the presence of rhetorical signaling cues incorporated to written or spoken text. Most of the research carried out up to now has analyzed signaling cues as they appear in written discourse; however, to our knowledge little or scarce attention has been paid to the role and function of DMs in spoken academic discourse. Goffman (1981) discusses on lectures mentioning that ‘footing’ is communicated through cues and markers but he fails to examine in detail those linguistic units. Flowerdew and Tauroza (1995) have studied the effect of the use of DMs in lectures, although concerned with second language lecture comprehension rather than with the role and function of DMs in the lecture situation. As far as I know, little research focuses on the analysis of discourse markers in monologic lecture talk as the study presented here does.

4.2. Definitions of Discourse Markers: the fundamentals

Discourse markers have been largely studied by researchers and they are still focusing their interest. Nevertheless, the term discourse marker has aroused some discussion. There is no consensus among researchers in understanding what the term discourse marker implies or refers to. On the other hand, researchers may happen to agree on the underlying concept of discourse markers, but they use different names to refer to that very same concept. Thus, we find labels such as: cue phrases (Knott & Dale 1994),

discourse connectives (Redeker 1990), discourse signaling devices (Polanyi & Scha 1983), pragmatic connectives (van Dijk et al. 1978, Stubbs 1983), or pragmatic markers (Schiffrin 1987; Fraser 1988, 1990), among others.

Regarding the theoretical status of discourse markers we want to focus on what they are, what they mean, and what functions they manifest. To do this we will review four research efforts that have been of great impact in the field of discourse analysis. The first approach is the work undergone by Schiffrin (1987), who studied elements, which mark “sequentially-dependent units of discourse”. The second approach is the one defined by Fraser (1999), who approached discourse markers from solely a grammatical-pragmatic perspective. The third perspective is the one provided by Blakemore (1987, 1992), who works with the Relevance Theory Framework (Sperber & Wilson 1986). The last approach is centered on the field of discourse coherence and has been followed by authors such as Mann and Thompson (1987, 1988) or Hobbs (1985).

In her book *Discourse Markers*, Schiffrin was concerned with the ways in which DM function to “add to discourse coherence” (1987: 326). Schiffrin maintains that coherence is constructed through relations between adjacent units in discourse (1987: 24). She basically sees DMs as serving an integrative function in discourse and therefore contributing to discourse coherence. She also points out the different nature of DMs, while some DMs relate only the semantic reality (the facts) of the two clauses, others, including *so*, may relate clauses on a logical (epistemic) level and/ or speech act (pragmatic) level.

In Schiffrin's (1987) view, DMs have both semantic and pragmatic meaning. This idea differs from Chaudron and Richard's (1986) definition of DMs, who argue that DMs simply indicate problems of on-line discourse production, that is, they act as filled pauses in order to give the speaker time to organise his/ her thoughts, and to give the listener time to process the spoken signal. However, Chaudron and Richard's (1986) do not attribute DMs signposting relations between different parts of the discourse.

Schiffrin was aware of the limitations of her research since she analysed only 11 expressions in the first instance, namely: *and, because, but, I mean, now, oh, so, then, well,* and *y'know*, as they occur in unstructured interview conversations. She clarifies that "except for *oh* and *well*...all the markers I have described have meaning", which she calls 'core meaning' (1987: 314). Later, she suggests other categories to be considered as DMs and that were not taken into consideration in her study in a first stage. These are perception verbs such as *see, look,* and *listen*, deictics such as *here* and *there*, interjections such as *gosh* and *boy*, meta-talk such as *what I mean is* and quantifier phrases such as *anyway* or *anyhow* (1987: 328). In any case, Schiffrin's research on DMs has been particularly relevant in the field of discourse studies and extremely influential for this ongoing research since she examined DMs in the spoken discourse of ordinary conversation.

Another study within the same approach is that of Redeker (1990, 1991), who defines a 'discourse operator' (1991: 1168) as:

[...] a word or phrase that is uttered with the primary function of bringing to the listener's attention a particular kind of linkage of the upcoming utterance with the immediate discourse context. An

utterance in this definition is an intonationally and structurally bounded, usually casual unit.

She proposes a revised model of discourse coherence based on three components: Ideational Structure, Rhetorical Structure and Sequential Structure. Redeker (1991: 1170) points out that “any utterance... in a discourse is then considered to always participate in all three components, but one will usually dominate and suggest itself as the more relevant linkage of this utterance to its context”. She revises Schiffrin’s notion of ‘core meaning’ and expands on this (1991: 1164) suggesting that “the core meaning should specify the marker’s intrinsic contribution to the semantic representation that will constrain the contextual interpretation of the utterance”.

Other approaches analyze and study DMs from a grammatical-pragmatic perspective. An example of this is Fraser’s study (1999). At an early stage Fraser speaks about ‘pragmatic formatives’ (1987) to finally arrive at the label of ‘pragmatic markers’ (1996). He characterized DMs as linguistic expressions. According to Fraser (1999: 936) this linguistic expression (or DM):

- (i) has a core meaning which can be enriched by the context
- (ii) signals the relationship that the speaker intends between the utterance the DM introduces and the foreign utterance (rather than only bringing up the relationship, as Schiffrin suggests).

Fraser goes on defining discourse markers as:

[...] a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases. With certain exceptions, they signal a relationship between the interpretation of the segment they introduce, S2, and the prior segment, S1 (1999: 937).

Fraser agrees with Schiffrin saying that DMs have a ‘core meaning’, and he adds at this point that their meaning is procedural and not conceptual. He (1999: 950) classifies two types of DMs: “those that relate the explicit interpretation conveyed by S2 with some aspect associated with the segment, S1; and those that relate the topic of S2 to that of S1”.

In a more recent publication, Fraser describes the canonical form for a DM SEQUENCE, that is, S1-DM+S2, where the S1 and S2 are discourse segments consisting of clauses, or the remain of clauses from which portions have been elided. In this article he defines a DM as:

a lexical expression, not necessary restricted to a single word and it need not be in S2-initial position [...]. In addition the S2 can generally be uttered by the speaker of S1 or a second speaker (2004: 15).

Based on his earlier research, Fraser (2004) discusses the syntactic, semantic and pragmatic properties of these units. According to DMs syntactic properties, he states there are five separate and distinct syntactic categories that contribute primarily to DMs:

- a) Coordinate conjunctions: *and, but, or, so, yet...*

- b) Subordinate conjunctions: *after, although, as, as far as, as if, as long as, assuming that, if, immediately ...*
- c) Adverbials: *anyway, besides, consequently, furthermore, still, however, then...*
- d) Prepositional phrases: *above all, after all, as a consequence, as a conclusion, in fact, in general, in contrast (to that)...*
- e) Prepositions: *despite, in spite of, instead of, rather, than ...*

These categories, although useful sometimes, fail to explain some cases in which alternative forms occur, especially in spoken discourse. Fraser (2004) goes on showing the semantic properties of DMs displaying four basic semantic relationships in the use of DMs, under the semantic point of view he proposes a marginal DMs classification where the most general DM of each sub-class appears in bold (the primary DM for the sub-class) and classifications are left open:

- a) CONTRASTIVE MARKERS (CDMs): **but**, *alternatively, although, contrariwise, contrary to expectations, conversely, despite (this/ that), even so, however, in spite of (this/ that), in comparison (with this/ that), in contrast (to this/ that), instead of (this/ that), nevertheless, nonetheless, notwithstanding, on the other hand, on the contrary, rather, (than this/ that), regardless (of this/ that), still, though, whereas, yet...*
- b) ELABORATIVE MARKERS (EDMs) **and**, *above all, also, alternatively, analogously, besides, by the same token,*

correspondingly, equally, for example, for instance, further(more), in addition, in other words, in particular, likewise, more accurately, more importantly, more precisely, more to the point, moreover, on that basis, on top of it all, or, otherwise, rather, similarly, that is (to say)...

c) IMPLICATIVE MARKERS (IDMs) *so, after all, all things considered, as a conclusion, as a consequence, (of this/ that), as a result (of this/ that), because (of this/ that), consequently, for this/ that reason, hence, it follows that, accordingly, in this/ that/ any case, on this/ that condition, on these/ those grounds, then, therefore, thus...*

d) TEMPORAL MARKERS (TDMs) *then, after, as soon as, before, eventually, finally, first, immediately afterwards, meantime, meanwhile, originally, second, subsequently, when...*

Although Fraser states that all DMs can be fitted into one of these four semantic categories according to their core meaning, we think these sub-classes can be misleading and confusing as for example the **Elaborative** class where we find instances of DMs that can function as additive markers (*and, in addition*) while others can be used as *reformulators (accurately, in other words)* (Del Saz 2003). Therefore, these four DMs distinctions are too broad and loose trying to gather and join too many DMs of a diverse nature. From his previous studies Fraser (2004) observed that only a few DMs can fall into more than one semantic relationship being *such* and *rather* some of these. However, from our point of view this classification of DMs based on

semantic relations does not take into consideration the pragmatic meaning of DMs, the speaker's intentions and illocutionary force that is so frequently conveyed in spoken discourse and sometimes expressed with the use of DMs.

Fraser agrees with Sweetser (1990) that many DMs which are syntactically conjunctions are pragmatically ambiguous and he goes on distinguishing three domains to which DMs apply:

- a) Semantic/ propositional content: John is very hungry, *so* he is eating a sandwich. (Facts of S1 justify Conclusion stated in S2)
- b) Epistemic/ logical: John is very hungry, *so* he must be very grouchy. (Knowledge of S1 justifies Conclusion stated in S2)
- c) Speech Act/ Pragmatic: John is very hungry, *so* go get him some food, please. (Report of S1 justifies the Request stated in S2) (Fraser 2004: 30-31)

The idea is that DMs are polysemous and that they may function in more than one domain not being restricted to a single and unique domain.

Blakemore (1987, 1992, and 1995) brings another theoretical perspective. She approaches DMs from the Relevance Theory Framework, dealing with DMs as a type of Gricean conventional implicature (Grice 1989), and focuses on how DMs impose constraints on implicatures. According to Blakemore, DMs have no representational but procedural meaning, agreeing with Fraser and Schiffrin. She maintains that DMs should be analyzed as linguistically specified constraints on contexts. Blakemore (1992: 138-141)

proposes four ways in which information conveyed by an utterance can be relevant:

- (i) It may allow the derivation of a contextual implication
(e.g. *therefore, too, also*)
- (ii) It may strengthen an existing assumption, by providing better evidence for it (e.g. *after all, moreover, furthermore*)
- (iii) It may contradict an existing assumption (e.g. *however, still, nevertheless, but*)
- (iv) It may specify the role of the utterance in the discourse
(e.g. *anyway, incidentally, by the way, finally*).

Researchers working in the field of discourse coherence have also approached the study of DMs from this perspective. We could mention studies such as those of Hobbs (1985), Mann and Thompson (1987, 1988), Sanders et al. (1992), Knott and Dale (1994) and Hovy (1995). According to these authors the discourse relations are sometimes made explicit by the use of DMs (they call them ‘cue phrases’). The discourse coherence perspective has focused on the number of discourse relations and their justification. A first approach within this perspective identifies and justifies a ‘standard’ set of relations relying on DMs. In this sense, the taxonomy of coherence relations mirrors DM differences of meaning. A second approach in this perspective identifies discourse relations as “psychological constructs that people use to create text” (Fraser 1999: 937). Knott and Dale (1994) tried to

combine these two approaches by analyzing ‘relational phrases’ (cue phrases) taken from written text, and then using a test based on substitutability, they established a taxonomy of relational phrases (cue phrases) hierarchically organized according to their different functions for signaling discourse relations. As Fraser (1999) states, the discourse coherence approach to the study of DMs is opposed to the other three approaches explained above, since here the discourse relations are seen as tools for text analysis, whereas in previous approaches discourse markers were seen as linguistic entities and their effect on the interpretation of discourse had a second place.

The proposal presented here is not strictly based upon the discourse coherence model; however, we agree that coherence is constructed through relations in the discourse and that these relations are frequently expressed by linguistic units (cue phrases, discourse particles, connectors, etc) I call here DMs. Along with Fraser and Schiffrin I agree on DMs having a core meaning, however I believe that this meaning is strongly context-dependent rather than semantic. All researchers mentioned above base their studies upon relations within discourse and how DMs affect or are affected by these relations. Schiffrin (1987) proposes three levels according to DM nature, Fraser (1999, 2004) distinguishes three domains following DM properties and Blakemore (1992) suggests four ways in which information can be conveyed by DMs within discourse. All these models are primarily based on the relations DMs signal. As a result, I consider a classification of DMs based on explicit discourse relations as commonsense and consistent. However, all classifications seem to be more semantic than pragmatic, since

there is no consideration of context/co-text, that is, the situation in which the discourse is produced.

In relation to Spanish authors whose focus of study has been DMs, the labeling of these linguistic units has also been controversial. They have been named in many different ways for the last fifteen years, most of them translated from English studies: *enlaces extraoracionales*, *organizadores discursivos*, *conjunciones*, *operadores pragmáticos*, *marcadores del discurso*, *señales discursivas*, *conectores de discurso*, *enlaces textuales*, *partículas modales*, etc. Whichever name given to these linguistic units, they have been studied according to their function in discourse, that is, pragmatic-discursive function, syntactic or lexico-grammatical function.

Bearing in mind that the present study takes into account the use of these linguistic units I name here DMs (as to put them together in a single group) within the spoken academic text, we are aware of the importance pragmatic-discursive and logico-semantic DMs have in the discourse rather than the syntactic and grammatical meanings conveyed.

Some Spanish authors clearly make a distinction between *operadores* and *conectores* explaining that *operadores* refer to a single utterance, whereas *conectores* relate two or more propositions (Escandell 1993: 115, Gutiérrez 1993: 21); however, we think this is a poorly explained and weak distinction to be taken into consideration.

Portolés (1993: 160) makes a preliminary shallow distinction between *conectores textuales* and *marcadores textuales* (a later and comprehensive classification would be provided in his work of 1998, read section 4.5 for DMs classifications). For him the so called *conectores textuales* serve to

process context by means of linking clauses at a semantic and pragmatic level, these are instances such as *además*, *por lo tanto*, *sin embargo*. On the contrary, *marcadores textuales* give rise to conversational inferences (Portolés 1993). What distinguishes *marcadores* from *conectores* is whether they convey conventional and controlled inferences or conversational ones, as it is the case of *conectores*. Portolés (1998: 11) is aware of the importance DMs have for the ongoing of the human communication phenomenon. He says:

Los marcadores constituyen, por ello, el primer paso en una dirección distinta de los estudios lingüísticos, son como unas nuevas Islas de los Galápagos, un espacio pequeño en el que se descubre una realidad diferente de la habitual, pero, por eso mismo, extremadamente iluminadora.

Later on, in his 1998 book he writes that DMs are:

[...] unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional y poseen un cometido coincidente en el discurso: el de guiar, de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas, las inferencias que se realizan en la comunicación (1998: 25-26).

Although DMs are not given a syntactic function in this definition, they are described as having morphological, syntactic, semantic and pragmatic properties when used in the inferential communication. He goes on to say that DMs are not fundamentally grammatical but semantic. He also

conceives of DMs as having a core meaning and that this meaning is procedural, coinciding with Schiffrin, Fraser and Blakemore. However, Portolés definition of *marcador* stands closer to Grice's Relevance Theory (1989) based on conventional implicatures among the discourse segments, in this sense coinciding with Blakemore (1987, 1992, 1995). In contrast to previous approaches which distinguished semantic, logical and pragmatic domains or levels (see Schiffrin 1987 and Fraser 2004) of DMs, Portolés (1993, 1998) fails to describe the logical relations conveyed by some DMs discriminating only between semantic and pragmatic levels.

Other Spanish authors are also concerned about the relations DMs can bring in discourse or segments within a sentence statement; this is the case of Llorente (1996: 14). She makes a distinction between DMs that signal logic and semantic relations and those which signal discursive and pragmatic relations, coinciding with Schiffrin's levels (1987) in some way, since she joins together in a single category DMs that express logico-semantic relations. Therefore we have two domains or levels, the former link meanings in, for instance, cause-effect, temporal or addition discourse relations, what she calls *conectores*. The latter are DMs which link communicative acts taken by discourse participants, organising and interlacing them; she calls them *operadores*. However, Llorente centres the scope of her study on the analysis of *operadores discursivos* disregarding the *conectores* that provide semantic discourse relations. She explains:

[...] el rasgo definidor pertinente de lo que llamo "operador discursivo" es su capacidad de servir a la realización de actos pragmático-discursivos, es decir, de actos necesarios para hacer

avanzar la interacción, de actos que regulan el desarrollo del discurso (Caron 1977), relacionan entre sí otros actos discursivos y, en resumen, se destinan a facilitar el procesamiento de la información (1996: 14).

Along with Schiffrin's research, Llorente (1996) aims at the study of spoken discourse by analysing colloquial, conversational language. She gathered and transcribed one hundred telephone conversations from a radio station and tried to analyse instances of what she calls *operadores discursivos* very relevant in the case of real life conversations, and aiming at facilitating discourse communication processes in direct speech under promptness and feedback constraints as in telephone conversations. She analyses DMs such as *Hola buenas tardes, vamos a ver, yo quiero, y resulta que, ya le digo, etc.* Some of these instances can only occur and be considered as DMs when analysing conversational talk, we would not give *hola* the status of DMs within the lecture genre and we would rarely find instances of *ya le digo*, for example. In any case, Llorente's definition of *operador* relies on the underlying concept of discursive act, in that sense and taking spoken corpora for the analysis of DMs, the notion of *operador* is undoubtedly fundamental. On the contrary, from our point of view the distinction between *conector* and *operador* seems weak for a proper classification of DMs.

I have looked at the notions of DM and which role or function they may have in the discourse as some authors have presented them. Next, I aim at revising relevant and recent studies on DMs that have approached spoken discourse based upon lecture corpora. Later, studies providing

classifications of DMs will be presented and discussed. I will first review previous classifications of DMs constructing the basis for our own classification model, which has to be equally pertinent to English and Spanish DMs and take into consideration the peculiarities and characteristics of the corpus under study.

4.3. Corpus studies on DMs (1990s onwards)

As said elsewhere in this chapter, scarce research relating the spoken lecture discourse and the use and function of DMs has been carried out to date. Recently, the genre of lecture as well as other academic genres (seminar, conference presentations, etc) has aroused the interest of researchers. DMs were considered to be paramount for the understanding of written texts; on the same basis, DMs are vital for the creation of a meaningful and coherent message in the communication process of oral discourse. Undoubtedly, the difference between spoken and written discourse affects the use and function of DMs. Those DMs that tend to be more recurrent in written texts, are less frequently used in spoken discourse (*on the contrary/ por el contrario, to sum up/ en resumen*, etc); instead there are other more recurrent ones such as *Ok, well/ bien, bueno*. In spoken discourse, speakers gain a richer context; they have prosody and phonology as well as non-verbal communication or interaction with external physical objects.

Research on DMs can be approached from second language lecture comprehension (L2) or departing from L1 DM role and function in lectures.

The studies reviewed here have used spoken or written corpora as the method of study and they have centred on the use of a single or various DMs. Furthermore each study aims at developing a wide and broad classification of DMs.

Del Saz (2003, 2005) for example, deals with the notion of reformulation and the lexical units that explicitly convey reformulation, what she calls Discourse Markers of Reformulation (DMs of RF). She argues that what has been defined in the English language as *reformulator* can clearly be considered a DM. To do so, her efforts focus on the inclusion of DMs of RF within Fraser's (1999, 2004) wider classification of DMs fitting them in the sub-class of Elaborative DMs. She bases her study on naturally occurring instances of language collected in the British National Corpus (BNC) (see Chapter II, section 2.4). She goes on to explain that *reformulators* are DMs since they have the defining properties of a DM according to Fraser (1999): connectivity and non-truth conditionality. Moreover the DMs of RF strictly follow Fraser's canonical form for a DM SEQUENCE, that is, S1-DM+S2. Del Saz' study (2003), although bringing into research the concept of DMs of RF, does not aim at finding a new category for these markers, but at including them within the previous wider classification provided by Fraser (1999). Fraser (1999) fails to explain some instances of DMs, besides his canonical form of DMs proposed, this being also too broad, as I argued in section 5.2, especially when working with a conversational spoken corpus.

Dealing with spoken academic discourse, Swales and Malczewski (2001) focus on what they call "a cluster of features that constellate around discourse management across a wide range of university speech events"

(2001: 146). They go on to define speech events as ‘activity types’ (Levinson 1979) in which language is the prime vehicle for getting things done. Swales and Malczewski (2001) are aware of the peculiarities of academic spoken discourse; they distinguish between that academic speech that is monologic (often lectures, conference presentations, etc) and dialogic talk (telephone workshops, interviews, etc). In their study they use MICASE (Michigan Corpus of Academic Spoken English) (R.C. Simpson, S.L. Briggs, J. Ovens, and J.M. Swales 2002) (see Chapter II, section 2.4) and analyse the use of footing changes they name New Episode Flags (NEFs). Their focus is on the linguistic resources (or NEFs) used by participants in a wide variety of university events to move from lecture format to discussion (or the reverse) or change the direction of the lecture or discussion. Then, they deal with linguistic resources such as group vocatives (*folks, gang, friends, guys*), directive or vocative verbs (*say, listen, look*) and exhortative or jussive imperative *let*, all these not with a large number of occurrences. In contrast, there are other recurrent and vital NEFs in the MICASE they analyse such as *okay, so* and *now*. These three last instances are considered in their study to be DMs as well as some uses of the exhortative or jussive *let* and its variations *let me (lemme), let us (let’s)*. Swales and Malczewski’s (2001) research on NEFs is relevant for our study because of the use of an academic spoken corpus and how they look at the functions and meanings of these linguistic units in the specific context these appear. However, from our point of view Swales and Malczewski make an overuse of syntactical or grammatical functions (adjectival and adverbial forms, group nominations, vocative verbs, etc) to label NEFs, rather than semantic, logical and

pragmatic relational meanings across the discourse, which are less used (fillers, back-channels, indicators of assent).

Another study dealing with academic discourse is one by Rendle-Short (2003), who analyses the use of the DM *so* in seminar talk within the computer science discipline. Based on Schiffrin (1987), she explains how *so* can function in two distinct ways. Semantically, *so* ties adjoining clauses together conveying causal relations. At this level *so* can mark structure at two levels. First, it can mark the overall structure by indicating its relationship to a whole stretch of discourse or what we understand as relation between one part of discourse and another part of discourse. Secondly *so* marks structure locally, by referring to the immediately preceding clauses, in this case *so* is marking lower-levels of structure. Moreover, *so* can also function pragmatically; in this case *so* marks the potential speaker discourse transitions functioning as such as a topic-shifter. Rendle-Short (2003) takes these three functions of *so* and analyses them in a corpus composed of six first video-taped and then transcribed computer science seminar talks. She found that the DM *so* is commonly used in seminar talk with different functions depending on its position in the talk. She is particularly attracted by the way in which DMs are used together with intonation, pitch, volume, gesture and tools considering all these resources for the speaker. Also, she has shown that monologic talk is not continuous but divided into smaller parts or sections which follow a finely organised and well-structured discourse pattern. The question is whether this finely organised discourse pattern can be universally applied or, on the contrary,

spoken discourse patterns mainly depend on the genre under analysis as well as on the speaker's performance.

Many Spanish authors (Marsà 1992, Llorente 1996, Portolés 1998, Lahuerta & Pelayo 2003, Lahuerta 2004, González 2004, 2005) have studied the use of Spanish DMs in discourse, primarily written discourse, but as far as we know not many overall classifications have been provided, instead peculiarities of some concrete DMs have been analysed (pragmatic markers, markers of reformulation, marginal uses of DMs,...). Lahuerta and Pelayo (2003) for instance, analyse the marginal uses of discourse markers and how they affect reading comprehension in Spanish as a foreign language. They select randomly some discourse markers such as: *por lo tanto*, *por otro lado*, *en primer lugar*, *por una parte/ por otra parte* and *sin embargo* as they appear in four journalistic texts in terms of, what Lahuerta and Pelayo have called, 'unorthodox uses' in contrast to 'orthodox uses' of these DMs. Lahuerta and Pelayo (2003) understand the unorthodox use of a marker when it has two or more different argumentative meanings. This occurs, of course, depending on the context some DMs appear in the discourse and the writer/ speaker's intention. They conclude that the marginal uses of some DMs can provoke reading comprehension errors in Spanish L2 readers and therefore discourse context is needed to lead inferences in the text.

Searching for contrastive research on DMs between English and Spanish languages we find Fraser and Malamud-Makowski (1996) and Malamud-Makowski (1997) research. Malamud-Makowski (1997) carries out a semantic analysis of Spanish contrastive and inferential DMs based exclusively on Fraser's (1990) domains. In contrast to previous corpus

research, she does not base her study on any corpora but rather on her intuition as a native speaker of Spanish from Argentina. She identifies Spanish contrastive and inferential DMs' core meaning and then proposes an organisation into groups but only from the semantic point of view. Fraser and Malamud-Makowski (1996) present an analysis of contrastive DMs in English to then compare them with counterparts in Spanish. They divide English contrastive DMs in six randomly chosen groups without labels and then try to find Spanish DM counterparts equally dividable into those six groups. Fraser and Malamud-Makowski's (1996) English and Spanish contrastive DM study is based on DM meaning according to a single unique previously provided classification based on semantic relationships, that of Fraser (1990). Additionally, DMs are randomly listed relying on NS intuition making no difference in the role and function of DMs between spoken or written discourse or among disciplines. To my understanding, a stronger basis is needed to be able to generalise the conclusions of research like this.

Other interesting studies on Spanish DMs have also been carried out by Latin-American researchers such as Curcó (2004) or Carranza (2004). Curcó (2004) analyses the adverb *siempre* as a DM in the varieties of Spanish spoken in México City (México) from a relevance theory perspective (Wilson & Sperber 1993) and based on Blakemore's procedural (1996) view of DMs. Carranza (2004) explores the use of *che*, *qué*, *qué sé yo*, *ah*, *¿no es cierto?* and the combination *bueno pero* in Spanish from Córdoba (Argentina) in terms of how they contribute to the oral text and social context in courtroom discourse. She analyses data from eighteen trials in the

Criminal Courts of a major city of Argentina. Her findings showed that the DM *ahora* acts mainly at an ideational level (Redeker 1990), *ahora* introduces new sub-topic or aids argument development in courtroom talk. *Bueno* is the most frequent marker in Argentinean Spanish and has many diverse applications as a ‘bracket’ or also known as, pause-filler. It can signal major shifts in the task at hand which may constitute a change in the speaker’s alignment. Participants in the study happened to mark a self-repair or a direct quote with *bueno*. Especially interesting is the observation of *bueno* co-occurring with the DM *pero* (*bueno...pero*), Carranza recognises that *bueno...pero* form a DM cluster where it is recognised as “bracketing specific interactional moves, self-repairs” (2004: 217).

The research on DMs above reviewed is relevant for the study here undertaken; however the specificities of the corpora (Peninsular SC and NAC) along with the novelty on the field lead towards the conception and development of a new DMs taxonomy. Other contrastive studies presented here do not compare Peninsular Spanish to English and moreover, not all of them deal with the lecture as an academic genre. Additionally, we have seen the specificities of some DMs such as *so*, whose functional and meaning variety is undoubtedly context-dependent. As a consequence, the development of a new DMs classification model based on the corpus under study is seen as a necessity. In the next chapter, relevant DMs classifications in literature are discussed.

4.4. Broad Classifications of Discourse Markers

As seen above, defining discourse markers has been a difficult task for researchers; however, classifying them has also caused discrepancies since they do not seem to agree on an exclusive and unique classification. Disagreements arise with regard to how the DM class should be delimited, whether the items comprise a single grammatical category, what type of meaning they express and the sense in which those units may be said to relate elements of discourse.

Dealing with the lecture discourse, I will to first mention one of the most cited classifications of DMs; the one by Chaudron and Richards (1986). They propose a distinction between micro markers (lower-order DMs) and macro-markers (higher-order DMs). Micro-markers indicate links between sentences within the lecture, or function as fillers. They fill pauses giving listeners more time to process individual segments of a piece of discourse; they hence provide more opportunities for bottom-up processing. Macro-markers signal the macro-structure of a lecture through highlighting major information in the lecture and the sequencing or importance of that information. These discourse signals help top-down processing. Under this two-folded category Chaudron and Richard's aim at providing a broad taxonomy for DMs.

They provide a taxonomy for micro-markers according to five semantic categories, namely, **Segmentation**, **Temporal**, **Causal**, **Contrast**, and **Emphasis**. **Segmentation** is understood as a semantic category including DMs which frame the segments of a discourse such as *well* or *OK*.

Contrastive relationships among the discourse are represented by the **Contrast** category. **Temporal** and **Causal** categories embrace DMs of intersectional relations, whether temporal or causal links. Relative emphasis can be expressed with DMs such as *of course*, *you see* or *in fact*, all these under the semantic category of **Emphasis**. Table 2 below includes the micro-markers that were taken into consideration by Chaudron and Richards (1986) and the categories established for their study.

Micro- markers

Categories:

Segmentation	Temporal	Causal	Contrast	Emphasis
<i>Well</i>	<i>At the time</i>	<i>So</i>	<i>Both</i>	<i>Of course</i>
<i>OK</i>	<i>And</i>	<i>Then</i>	<i>But</i>	<i>You can see</i>
<i>Now</i>	<i>After this</i>	<i>Because</i>	<i>Only</i>	<i>You see</i>
<i>And</i>	<i>For the moment</i>		<i>On the other hand</i>	<i>Actually</i>
<i>Right</i>	<i>Eventually</i>			<i>Obviously</i>
<i>All right</i>				<i>Unbelievably</i>
				<i>As you know</i>
				<i>In fact</i>
				<i>Naturally</i>

Table 2. Chaudron and Richards' (1986) classification of micro-markers.

This taxonomy, however, is exclusively based on semantic relations across the discourse, obviating other domains provided by the state-of-the art research on DMs (Schiffrin 1987; Blakemore 1987; Fraser 1990, 1999, 2004) and therefore clearly misinterpreting the signposting role carried by DMs within interaction (Schiffrin 1987, Shourup 1999). Moreover, labeled categories are not morphologically and syntactically homogeneous. Compare **Emphasis** with **Contrast**, for example. From this view, we could

assume that when we establish a contrast relation across utterances the aim is not to emphasize, I digress from this idea as will be discussed later on.

Regarding the classification of macro-markers and contrarily to what they had done with micro-markers, Chaudron and Richards (1986) did not distinguish any semantic category. On the contrary, a list of those macro-markers contained in the lecture established for the development of their study was provided. The list included signals or metastatements about the major propositions within the lecture, or the important transition points in the lecture. The macro-markers as listed by Chaudron and Richards in their work can be seen in Table 3 below.

Macro-Markers

<i>What I'm going to talk about today is something</i>	<i>Another interesting development was</i>
<i>You probably know something about-already</i>	<i>You probably know that</i>
<i>What [had] happened [then/after that] was [that]</i>	<i>The surprising thing is</i>
<i>We'll see that</i>	<i>As you may have heard</i>
<i>That/this is why</i>	<i>Now where are we</i>
<i>To begin with</i>	<i>This is how it came about</i>
<i>The problem [here] was that</i>	<i>You can imagine what happened next</i>
<i>This/that was how</i>	<i>In this way</i>
<i>The next thing was</i>	<i>It's really very interesting that</i>
<i>This meant that</i>	<i>This is not the end of the story</i>
<i>One of the problems was</i>	<i>Our story doesn't finish there</i>
<i>Here was a big problem</i>	<i>And that's all we'll talk about today</i>
<i>What we've come to by now was that</i>	

Table 3. Chaudron and Richards' (1986) macro-markers analyzed in their study.

In *The Effect of Discourse Markers in the Comprehension of Lectures* (1986), Chaudron and Richards' findings appear to show that macro-markers "are more conducive to successful recall of the lecture than micro-markers. Micro-markers do not aid the learner's retention of the lecture" (1986: 123). They claim two main reasons for these findings. On the one hand, micro-markers do not add enough content to make the subsequent information meaningful or salient. On the other hand, the quantity of the markers scattered probably results in making the entire lecture appear less well organized, a notion already pointed out in Hiller et al.'s work (1969). Although Chaudron and Richards' study deals with the university lecture genre as our study does, and how DMs affect students lecture comprehension, the classification they propose is the result of analysing a small sample corpus, a single American history lecture in four different versions; therefore not significant enough to be taken as a single model for a classification of DMs as the lack of semantic categories for the macro-makers shows.

Earlier, Murphy and Candlin (1979) had already provided a complete taxonomy for the classification of macro-markers, based on the analysis of engineering lectures. However, Chaudron and Richards' classification lacks this categorization for the macro-markers; the reason could be that the macro-markers found in Chaudron and Richard's (1986) study did not clearly fit into any of the categories Muphy and Candlin (1979) proposed.

Murphy and Candlin (1979) distinguished first among three types of discourse markers within micro-markers division: i) **Markers**, they include signaling devices such as *well, right, now*, providing a clearer discourse

segmentation; ii) **Starters**, for example, *Well now, let's get on with*, which establish links among discourse; and iii) **Metastatements**, used to emphasize important information in the discourse as for example *I want to mention three types of generator*. With respect to macro-markers, and contrary to Chaudron and Richards (1986) who did not label categories, Murphy and Candlin (1979) developed the following macro-marker divisions: **Starter**, to begin the discourse; **Elicitation**, which includes the words or expressions eliciting information; **Accept**, in order to show approval; **Attitudinal**, where the speaker takes positions about the discourse content; **Informative**, words used to emphasize important information; **Comment**, to express additional information; **Aside**, considered as an attempt to deviate from the ongoing discourse; **Metastatement**, which includes all the words and expressions used to strengthen and validate points in the discourse; and **Conclusion**, including final remarks.

Although these two divisions of micro and macro-markers proposed above are a daring attempt to classify DMs (Murphy & Candlin 1979, Chaudron & Richards 1986), these taxonomies are mainly based on semantic categories. I agree with the underlying division of micro and macro markers, being the definition of macro-marker probably much more clearly identifiable as linguistic units highlighting major information in the lecture, whereas the definition of micro-marker seems quite narrow and weak as lower-order markers or simply pause-fillers. However, individual categories seem to fail to express pragmatic or discursive discourse relations.

Based on these broad classifications previously mentioned and still on lecture discourse, Morell (2001) examines the role of discourse markers and

personal pronouns in the discourse of two lecture styles: an interactive (participatory) lecture and a non-interactive (explanatory) lecture. When analysing DMs she also makes the distinction between macro and micro-markers. In order to classify micro-markers, she uses Chaudron and Richards' (1986) taxonomy (see Tables 2 and 3), but she adds a new category to the micro-markers classification: **Elicitation**, already present in Murphy and Candlin's (1979) classification of macro-markers, as seen above. Within the **Elicitation** semantic category we can find DMs that are used to elicit information from the students or to involve them in the discourse, for example, *Anything else?* This new category was not found in Chaudron and Richards' classification as they analyzed reading style lectures that do not include elicitation, whereas in a more interactive kind of lecture, elicitation is to be expected. In terms of macro-markers, Morell (2001) categorizes them according to some of the divisions cited by Murphy and Candlin (1979), disregarding three categories from the original classification, namely, *Informative*, *Comment* and *Aside*, since they were not relevant for her study. Table 4 below includes Morell's classification of micro and macro-markers based upon authors such as Murphy and Candlin (1979), and Chaudron and Richards (1986). Morell (2001) primarily displays micro-markers and later macro-markers, then she distinguishes micro and macro-markers as they appear in the two types of lectures analyzed. The following first two tables present micro-markers in the non-interactive and interactive lectures, the latter shows macro-markers as they appear in the non-interactive and the interactive lectures.

MICRO-MARKERS

Non-interactive

Segmentation	Temporal	Causal	Contrast	Emphasis	Elicitation
<i>ok</i>	<i>and then (after)</i> <i>that</i>	<i>so</i>	<i>but</i>	<i>in fact</i>	<i>ready?</i>
<i>and</i>		<i>that (which)</i> <i>means</i>	<i>although</i>	<i>of course</i>	
<i>or</i>		<i>because</i>		<i>Such as</i>	
<i>now</i>		<i>so (that)</i>		<i>Note</i>	
<i>well</i>		<i>therefore</i>			

Interactive

Segmentation	Temporal	Causal	Contrast	Emphasis	Elicitation
<i>ok</i>	<i>then</i>	<i>so</i>	<i>but</i>	<i>of course</i>	<i>anything else?</i>
<i>and</i>			<i>unless</i>	<i>as you know</i>	<i>any others?</i>
<i>or</i>					<i>anyone else?</i>
<i>now</i>					<i>what?</i>
<i>anyway</i>					<i>why not?</i>
					<i>why is that?</i>
					<i>louder, please</i>

MACRO-MARKERS

Non-interactive

Starter	Elicitation	Accept	Attitudinal	Metastatement	Conclusion
<i>today I'm going to talk about</i>	<i>what do we mean by...</i>		<i>I would dare to say</i>	<i>I have a quote for you</i>	<i>what you have seen in this lecture first was</i>
<i>first I'll talk about</i>	<i>Remember...</i>		<i>I believe that</i>	<i>I'd like to read this to you</i>	<i>and then we have seen...</i>
<i>...and then about</i>	<i>any questions</i>		<i>I do believe</i>	<i>It says</i>	
<i>the reasons</i>			<i>I wouldn't doubt it</i>		
<i>to begin...</i>					
<i>another piece of data(which you might find interesting)</i>					
<i>the second item in this lecture</i>					
<i>the third item of this lecture</i>					

Interactive

Starter	Elicitation	Accept	Attitudinal	Metastatement	Conclusion
<i>we are going to get started</i>	<i>I have a question for you</i>	<i>that's right that is true</i>	<i>I think</i>	<i>to back up that statement</i>	<i>to finish today's lecture</i>
<i>can I have your attention?</i>	<i>do you think...?</i>		<i>it's a difficult question to answer</i>		<i>that finishes up today's lecture</i>
<i>today we're going to start to talk about</i>	<i>what makes you think so?</i>				<i>we'll continue with this tomorrow</i>
<i>before we begin</i>	<i>what was that like?</i>				
<i>we will begin now</i>	<i>do you remember when we talked about...?</i>				
<i>this first quote</i>	<i>do you remember?</i>				
<i>another interesting fact</i>	<i>what do we mean by...?</i>				
<i>another important piece of information</i>	<i>how do you think...?</i>				
<i>i'd like to give one more fact or piece of information</i>	<i>for what reason do you think?</i>				
	<i>which one do you think...?</i>				
	<i>let's take a mini survey here</i>				
	<i>another question for you</i>				
	<i>the next question I have for you</i>				
	<i>does anyone have an answer for that? or think they have an answer?</i>				
	<i>do you agree?</i>				
	<i>how about...?</i>				
	<i>what's the difference?</i>				
	<i>remember we've talked about this before</i>				

Table 4. Morell's (2001) micro and macro-marker samples and her proposed taxonomy according to interactive and non-interactive DMs.

Significant in Morell's (2000, 2001) research is the inclusion of a new category, **Elicitation**. Although our study deals with monologic lecture talk; we have observed that a grading of monologic talk does exist related to students' interventions in lectures and interaction lecturer-hearer (Bellés-Fortuño 2004). Lectures can be considered highly monologic or mostly monologic when instances of students' interventions occur. In any case, Morell (2000, 2001) does not aim at providing a broad taxonomy of DMs, but instead she uses previous classifications inheriting the gaps and lacks already mentioned.

In a previous study (Belles-Fortuño 2004) I carried out a contrastive analysis between North-American and British English lectures and the role and function of DMs in them. Contrary to Chaudron and Richards' (1986) results, the findings showed that the use of micro-markers was more relevant and recurrent than macro-markers due to the spoken lecture corpus under study. I concluded that a more general use of micro-marker in both North-American and British English lectures could be due to the steadiness and invariability of these markers. Micro-markers (*so, well, the, because, etc.*) (see tables below) are types of more fixed and invariable signaling cues than those DMs signaling the macro-structure of a lecture, that is macro-markers. Contrary to micro-markers, macro-markers can vary according to the type of discourse, disciplinary variations or even lecturers' personal style.

Along with Morell, I did not focus the study on the development of a broad taxonomy of DMs, rather I departed from the previously fixed taxonomies described above (Murphy & Candlin 1979; Chaudron & Richards 1986;

Morell 2000, 2001). However, the DMs presented in previous classifications were not equally recurrent and significant for our analysis, therefore instances of micro and macro-markers within the corpus substituted previously classified DMs. The English DM classification model is as follows:

Micro-markers

Categories:

Segmentation	Temporal	Causal	Contrast	Emphasis	Elicitation
<i>Ok</i>	<i>and then</i>	<i>so (that)</i>	<i>but</i>	<i>in fact</i>	<i>why is that?</i>
<i>and</i>	<i>after this</i>	<i>because</i>	<i>although</i>	<i>of course</i>	<i>anything else?</i>
<i>now</i>	<i>after that</i>	<i>therefore</i>	<i>unless</i>	<i>as you know</i>	<i>anyone?</i>
<i>well</i>	<i>eventually</i>				<i>why not?</i>

Table 5. Micro-markers classification model (Bellés-Fortuño 2004).

Macro-markers

Categories:

Starter	Elicitation	Accept	Attitudinal	Metastatement	Conclusion
<i>today i'm/ we're going to talk about, i'll/ we'll talk about</i>	<i>(wh-) do you think?</i>	<i>that's right</i>	<i>I think</i>	<i>let me (lemme)</i>	<i>finally</i>
<i>to begin with</i>	<i>any questions</i>	<i>right</i>	<i>I believe that</i>	<i>let's try, go back, find, focus</i>	<i>the last thing</i>
<i>the second thing</i>	<i>how about...?</i>	<i>excellent</i>	<i>we believe</i>	<i>it says</i>	<i>to end up/with</i>
<i>firstly, secondly, thirdly</i>				<i>I wanna/ want to mention, get back to, do...</i>	

Table 6. Macro-markers classification model (Bellés-Fortuño 2004).

At that time, I realized that categories in our classification based on previous ones somewhat lacked stability, resulting in a rather unsteady categorization

of DMs, not responding to the general literature basis presented before (domains or levels), that is, semantic reality (the facts), a logical (epistemic) level and/ or speech act (pragmatic) level, representing only semantic relations between clauses. Moreover, we encountered some difficulties in classifying some DMs, partly due to the lack of coherence in the criteria used to distinguish categories (eg. the categories Temporal and Causal are closely semantic, whereas Attitudinal is more pragmatic).

Spanish DMs have also aroused the interest of researchers especially during the last 20 years (Fuentes 1987, Casado 1991, Portolés 1998, Martin & Portolés 1999, to name a few). However, most of their studies have been focused on the use of some specific and more concrete types of DMs and how they functioned in written texts rather than spoken language, without aiming at providing an exhaustive classification of DMs. It is Portolés (1998) who aimed at and tried to display a Spanish DMs classification model that could be widely, but not exclusively, useful for a vast number of discourse contexts.

As said before in this section, most of the studies carried out by Spanish authors used their corpus on written discourse; still, there is a study carried out by González (2004, 2005) which draws our attention, since her study shares traits with our ongoing research. González's study is a contrastive analysis between Catalan and English based on an oral corpus. She analyses 40 oral narratives, 20 in English and 20 in Catalan. The informants from whom the oral narratives were elicited were all native speakers of Catalan and English. González claims the different nature of DMs. She distinguishes those DMs that set up logico-semantic argumentative relations (of cause,

result, reason, concession, contrast, time, etc) from pragmatic markers. She states these DMs have “descriptive or lexical meaning and have been traditionally called in the literature ‘argumentative connectors’” (González 2005). DMs that are included in this category are for instance, *therefore (por lo tanto)*, *in contrast (en contraste con)*, *on the other hand (por otra parte)*, *nevertheless (sin embargo)*, *because (porque)*. Additionally, she mentions those markers:

[...] whose main functions are rhetorical signal the speaker’s intentions and goals and basically help convey the illocutionary force of the story. Markers found in the sequential structure delimit segments boundaries and sustain the discourse network; they highly facilitate the in-and-out shift of the narrative segments. In the case of markers that have a dominant inferential role, the link that is set up between the cognitive domain of the speaker and hearer is fundamental to understand and grasp the point of the story (González 2005: 54).

This DM category is what González calls pragmatic markers under the three structures: rhetorical, sequential and inferential. She disregards semantic markers for her study paying attention only to the so called pragmatic markers, in particular she takes for her study markers such as *well, so, then, I mean, you know* and *anyway*, together with their Catalan counterparts *bé, bueno, clar, doncs, pues, llavors, aleshores, no* and *eh*. She displays a proposal for a discourse coherence model based upon Schiffrin (1987) and Redeker’s (1990) discourse coherence models and on the semantic versus pragmatic source of coherence (González 2005:57). Figure 5 below shows González discourse coherence relations model.

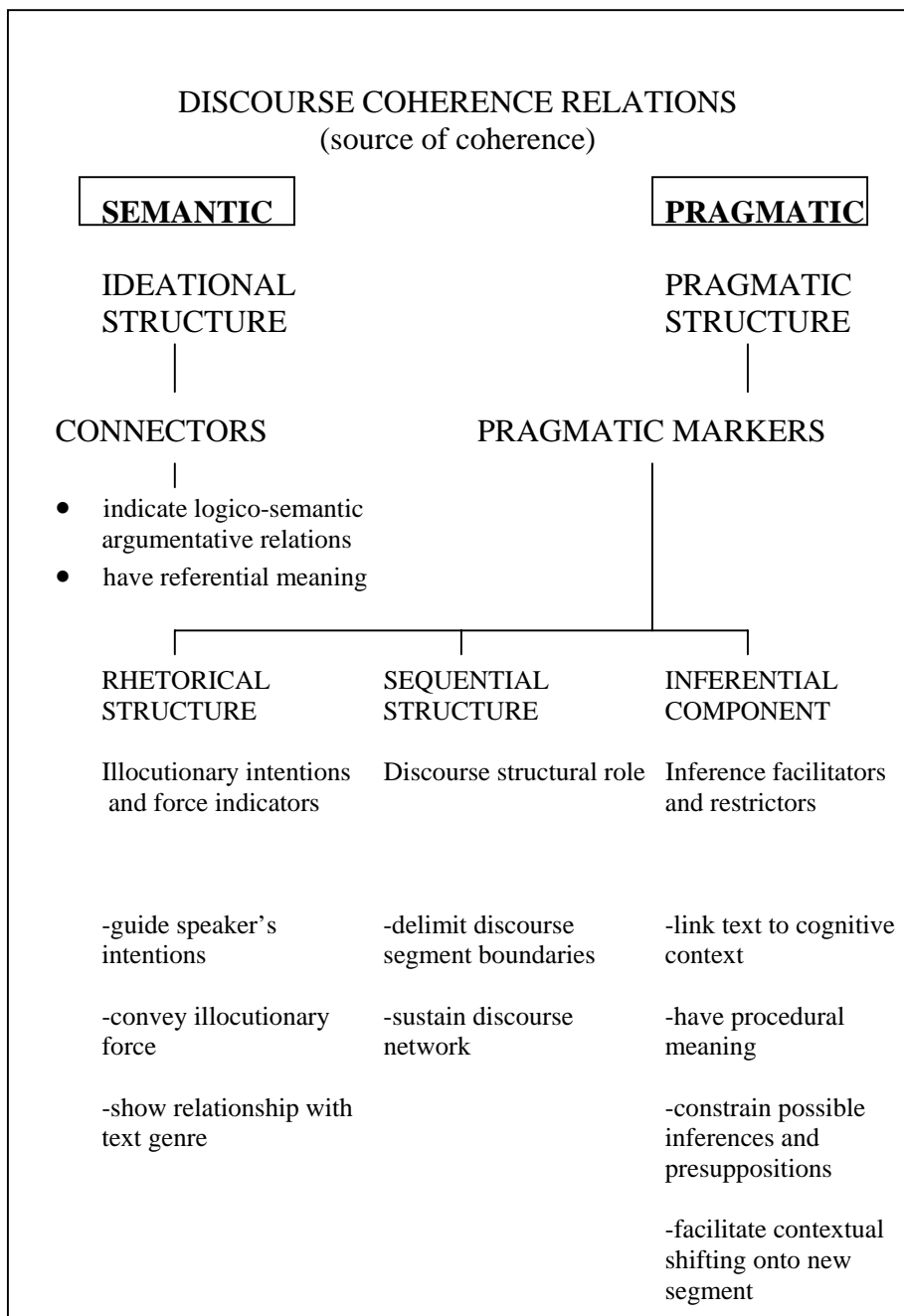


Figure 5. González's (2005) proposal of the distribution of markers in the discourse structure components.

From her study on pragmatic markers on oral narratives in Catalan and English, González (2005) concluded that both English and Catalan pragmatic markers have a predominant role in the rhetorical and sequential narrative structures and that English markers are more attached to the

ideational structure than Catalan markers. She also observed that the appearance of some markers is not arbitrary in the narrative since certain markers recurrently appear in certain discourse segments. González' study of pragmatic markers is important for our own research, not only because we share common traits in our study (contrastive analysis, use of a corpus, analysis of pragmatic marker use in oral discourse, etc.) but because of her proposal of the distribution of DMs, which in fact is an attempt for a DM categorization according to pragmatic and semantic discourse relations. However, González (2005) centers the scope of her study only on pragmatic markers, disregarding those markers used in the ideational structure and not providing then a complete classification of DMs. In our study we do not disregard any kind of DM relation aiming at fulfilling all domains or levels. It can also be observed from González' study that when contrasting two languages DMs counterparts are not always possible or do not express the same discourse relation.

We have considered Portolés' (1998) Spanish DM classification as being one of the most complete ones, although not based on the study of any specialised spoken or written corpora. He relies on his knowledge of DMs as a native speaker of Spanish and takes examples from Spanish literary works. He offers a classification based on his main distinction between *marcadores* and *conectores* (see section 4.2. Definitions of DMs, in this chapter). A closer look at this classification reveals that there are five basic categories: *estructuradores* (structural markers), *conectores per se* (connectors), *reformuladores* (reformulators), *operadores* (operators) and *marcadores* (markers). Four of these five broad categories are divided into other sub-

classes (except for *Marcadores de control de contacto* or *Proximity monitoring markers*) (see Table 7 below).

STRUCTURAL MARKERS	ORGANISERS	En primer lugar/ en segundo, por una parte/ por otra, de un lado/ de otro lado, asimismo, por lo demás, etc.
	TOPIC-SHIFTERS	Por cierto, a todo esto, a propósito, etc.
CONNECTORS	ADDITIONAL	Además, encima, a parte, incluso, etc.
	CONSECUTIVE	Por lo tanto, en consecuencia, de ahí, entonces, pues, así pues, etc.
	CONTRASTIVE	En cambio, por el contrario, antes bien, sin embargo, no obstante, con todo, ahora bien, etc.
REFORMULATORS	EXPLICATIVE REFORMULATORS	O sea, es decir, esto es, en otras palabras, etc.
	REPHRASING REFORMULATORS	Mejor dicho, más bien, etc.
	ATTITUDINAL REFORMULATORS	En cualquier caso, en todo caso, de todos modos, de cualquier manera, etc.
	CONCLUDING REFORMULATORS	En suma, en conclusión, en definitiva, en fin, al fin y al cabo, etc.
DISCURSIVE OPERATORS	BACK-CHANNELS	En realidad, de hecho, claro, desde luego, etc.
	CLARIFICATORS	Por ejemplo, en concreto, etc.
	FILLERS	Bueno
PROXIMITY MONITORING MARKERS		Hombre/ mujer, mira, oye, etc.

Table 7. Portolés' (1998) DMs classification.¹

¹ This is an adapted English translation from Portolés' Spanish DMs classification.

For the better understanding of Portolés' classification we have provided a *sui generis* translation faithfully maintaining meanings of categories. As can be observed, DMs labels seem to be recurrent in the literature, in this classification we find discursive **operator** (*operador discursivo*) (Llorente 1996), **reformulator** (*reformulador*) (Del Saz 2003), **connector** (*conector*) (Llorente 1996, González 2004), **markers** (*marcador*) (Murphy & Candlin 1976, Portolés 1998), etc. Although labels may coincide in the literature, variation of languages, disciplines and genres may affect DM meaning in the discourse and their categorization.

Let's look at Portolés's classification in more detail. If we take the sub-classes of **organizers** and **concluding reformulators** (highlighted light yellow) they appear under two different categories (**structural markers** and **reformulators**); however, I think that when we are concluding, this section of talk is part of the macro-structure organization of the discourse and therefore these two DMs would be under the same category of structural markers. Still within the category of **reformulators** we have **explicative reformulator** and **rephrasing reformulators** (highlighted light blue) as being two different sub-classes. To our understanding, the DM *más bien* (similar to *I mean*) (**rephrasing reformulator**) and *en otras palabras* (*in other words*) respond to the same purpose, that is, rephrasing or explaining the same concept in different words. Along with this meaning the DMs *en realidad* or *de hecho* (*in fact*) could also be considered **reformulators** in some contexts since they aid the listener's comprehension process, while in the other contexts they can function as plain emphasizees. In Portolés classification these last DMs have been included within a different category

I have translated as ‘back-channel’. By ‘back-channels’ we understand DMs which aim at prompting or agreeing with the main speaker or what has been said by him/ her. Then, DMs such as *claro* or *desde luego* (*that’s right, excellent*) clearly fit in this category but not *en realidad* or *de hecho* (*in fact*) from our point of view.

What has been considered by Portolés (1998) a **marker of proximity** (highlighted pink) is in fact a group vocative (Swales & Malczewski 2001) which for us does not have the status of DM and is probably more characteristic of colloquial conversation discourse, in which case it could also be considered as attention-getter.

In the next chapter (Method of study) one of my efforts is centered on the design of a DMs classification model valid both for English and Spanish DMs, taking into consideration the advantages and disadvantages of previous classifications and the peculiarities of our lecture genre corpus.

In this chapter I have tried to prove that there is not a unique definition of what a DM is or a single unique taxonomy; however, there are some shared fundamentals on the relations DMs convey among utterances and therefore this will be our starting point for providing a novel classification.

4.5. PURPOSE OF THE STUDY

The main purpose of this study is to analyze spoken academic discourse, and more concretely the genre of lecture as it is presented to undergraduate university students, by means of a contrastive analysis between Spanish and

English lectures. I analyze in particular the role and function of DMs in the lecture discourse.

Research on the effects of DMs in lectures has been approached regarding second language comprehension (Flowerdew & Taroza 1995). However, the study here presented concerns both the SLA approach as well as the effects of DMs for native speakers of English and Spanish. With the results of this study I intend to provide insights into how DMs are used and function in Spanish and English language for both native and non-native speakers in tertiary institutions. I intend to:

- a) help **native Spanish lecturers** to improve their lecture discourse both in Spanish and in English
- b) help **native Spanish/English lecturers** to improve their lecture discourse both in Spanish and in English.
- c) help and benefit **English/Spanish both L2 and native undergraduate** students for the comprehension of lecture discourse in their learning process.

Our corpus consists of twenty-four transcripts of university lectures within the academic division of Humanities and Social Sciences. On the one hand, I have the North-American corpus (NAC) made of twelve North-American English lecture transcripts from the University of Michigan (United States). On the other hand, the Spanish Corpus (SC) with twelve Spanish lectures recorded and transcribed at Universitat Jaume I in Castellón (Spain). Once

the corpus and the features to be analyzed chosen, I consider more concrete objectives departing from the following research questions:

- a) Is there any difference in the use of DMs between North-American and Spanish monologic lectures in the discourse of Humanities and Social Sciences?

- b) What is the relation between the several types of markers? Do some specific DMs usually collocate?

Taking into consideration previous results and premises in similar studies, I will try to shed some light on the differences and similarities between North-American and Spanish lecture discourse in the field of Humanities and Social Sciences regarding the use, role and function of DMs.

Chapter V: Method

5. METHOD

5.1. Corpus selection

In order to reach the objective of this study I have gathered a corpus with a total amount of twenty-four spoken lecture transcripts. As this is a contrastive analysis, half of the corpus (twelve lectures) consists of Spanish lectures; the other half contains North-American English lectures. The North-American English lecture transcripts, also called North-American corpus (NAC), have been taken from MICASE (Michigan Corpus of Academic Spoken English) (R.C. Simpson, S.L. Briggs, J. Ovens, & J.M. Swales, 2002) available on the Internet thanks to the English Language Institute at the University of Michigan (United States). MICASE is available through an on-line search engine containing a collection of transcripts of academic speech events recorded at the University of Michigan at Ann Arbor. The corpus consists of approximately 1.8 million words transcribed from a variety of speech events that goes from February 1998 up to 2003.

The Spanish part of the corpus (SC) consists of twelve lectures recorded at Universitat Jaume I, Castelló (Spain) and transcribed for the purpose of this dissertation. The SC compilation has been an arduous and time-consuming activity that took us a whole academic year. Firstly, lecturers were selected and contacted in order to get their approval to be recorded. It is worthwhile mentioning that most of them agreed and actively cooperated in the lecture recording. Anecdotes have also been part of the lecture recording process.

For the selection of lectures I chose some that were delivered during the first and second semester and that matched or shared traits with the NAC lectures in order to have both sub-corpora (NAC/SC) as homogeneous as possible. The process was the same for most Spanish lecture recording²; with a minidisk the whole lecture was recorded while I observed the lecture delivery. That is, I was present while the lecture was going on, so that I could do classroom observation and take notes about several things such as: classroom distribution, number of students, students' characteristics, number of students' interruptions, lecturers' use of resources (bb, OHP, Power Point Slides, handouts, etc.) or any other classroom event worthy to be noted and that could affect lecture delivery. At this point, I would like to mention that a Classroom Observation sheet was fully developed departing from the lecture observation, this Classroom Observation Guide served to take notes from the recordings that could be of great help not only for this dissertation but for any further research on the field. This Classroom Observation Guide is currently used by the GRAPE (Group of Research on Academic and Professional English) when gathering a corpus. The complete Classroom Observation guide with a full description of the Spanish lecture corpus gathering can be seen in Appendix C at the end of this dissertation. It was up to the lecturers whether to inform students the lecture was being recorded or not; although some would not tell students, others urged me to do so, therefore I left lecturers the decision. In the classroom, I tried to mix up with the audience and pass unnoticed.

² Lectures LE3, LE4, LE5 and LE11 were recorded by the lecturers themselves and there was no classroom observation as the Classroom Observation Guide shows in Appendix C.

Once the lectures were recorded, what I have called the computer technology process began. The lecture recordings were transferred to the computer and saved as sound files; this procedure was done in order to facilitate the transcription. Before beginning transcribing, I established the norms that had to be followed which are mostly detailed in Table 8 below.

<p>...: short pause, stammering [down tone] [up tone] [falling intonation]: after a rhetorical question, lecturer is not seeking for response but gives it him/herself [irony] [LECTURER] [LECTURER DOES NOT TAKE INTO ACCOUNT STUDENTS INTERVENTION.] [LECTURE WRITES ON THE BOARD] [LECTURER REFERS TO OHT/PPT] [INTERRUPTION, STUDENT COMES INTO THE CLASSROOM] [STUDENT INTERVENTION](male student/female student answers, makes a comment) [INTERRUPTION FOR LECTURERS' EVALUATION] [pause expecting an question or confirmation check] [pause, waiting for an answer] [pause]: when is long [QUESTION ADDRESSED TO OBSERVER] [slows down the discourse] [partial repetition] [repetition, adding info. [repetition, paraphrasing] [repetition] [murmur] [laugh] underline (shows <i>code-switching</i>) <i>italics</i> for anglicisms <unclear></p>
--

Table 8. SC Transcription norms³.

³ These norms were originally used in Spanish and have been translated for a better understanding.

The transcribing process has surely been the most arduous and time-consuming task I had to develop and which had to be done with the help of team work. Once the first transcript draft for each lecture was ready, the sound files saved in the PC were recorded in the form of CDs and both CD and lecture transcript draft were sent to each lecturer, who was asked to proof-read and correct any misunderstanding or omitted information. This first reviewing process was done by the lecturers; later, these corrections were again revised and proper modifications made until the final transcript version was ready for the analysis.

As long as I could, I tried to record lectures delivered during or in the middle of the academic year, disregarding first or last lectures and those which required other specific characteristics different from the normal lecture session delivery (video showing, group seminar work, laboratory, etc). As an example, LE12 (see Table 9 for lecture attributes) is slightly shorter than the other Spanish lectures due to its specific characteristics, it is the only one delivered in a computer laboratory (see Classroom Observation Guide LE12 in Appendix C), since this subject has theoretical as well as practical credits, the first 33m were devoted to a theoretical explanation fitting the monologic lecture discourse, what followed is the practical computer lab session where each student works on his/her own; therefore this part was disregarded.

As to participants, the students and teachers involved in the North-American English lectures are mostly native speakers of the language (NS). In the SC, participants are NSs of Spanish but in a language context where Spanish and Catalan are both official languages and where most students and teachers

are mother-tongue bilingual in these two languages, as some examples of code-switching have shown in the SC transcripts (see Appendices A and B). Lecturers' gender was also taken into account, consequently for the SC I tried to record both male and female lecturers having seven male and 5 female lecturers. Regarding the NAC the amount is reverse, five male and seven female lecturers. Although not equal in number, I wanted to get a balance between male-female lectures in both sub-corpora as an effort to have a large homogeneous corpus.

Both sub-corpora, the NAC and SC are then available on audio files. Regarding the Spanish corpus (SC), it was recorded following the procedure already explained above. As for the NAC, one of the restrictions I had when choosing the lectures was whether the lectures were available in on-line audio files or not. I considered strictly necessary being able to audio check lectures as to the use of some DMs and the lecturers' attitudes, tone, intonation or other traits that can be observed from a recording and that cannot be found in a transcript. Some of the NAC lectures are included in the On-line audio file on the MICASE website: (<http://www.lsa.umich.edu/eli/micase/Audio/index.htm>). However, not all lectures are available in audio format on-line, and some had to be ordered directly from the ELI (English Language Institute) in Ann Arbor (MI) which sent two CDs, and others were kindly provided by the ELI staff.

As mentioned elsewhere, most lectures taken as part of the corpus for this study belong to the division of Humanities and Social Sciences. This choice was not made arbitrarily. Social Sciences as an academic field of study has currently gained importance because of a rising need for communicative

competence in the field. The reason is that teachers have to be professionally trained for the market demand in all oral genres, whether professional or academic, as it is the case of lectures. On the other hand, as a linguist, a corpus on Humanities and Social Sciences makes me feel closer to the field of study, than other “harder” sciences, triggering those results that can be relevant for the study. Another reason for this choice deals with the notion of ‘neutrality’; some researchers have pointed out that the discourse of Social Sciences tends to be more neutral than that of Technical Sciences or the field of Humanities. Technical Sciences are believed to show an objective and positive discourse, whereas Humanities discourse tends to be more creative and subjective. Social Sciences, therefore, would be placed in between, with a more ‘neutral’ type of discourse (Giménez, 2000). In this sense, the lectures recorded were especially chosen disregarding others that could have their own specific characteristics, in an attempt to look for ‘neutrality’ and avoiding deviances. A lecture on *circuits* or *graphic design* for instance, would obviously have different traits, being more graphical and largely objective.

5.2. Corpus description

The description of both parts of the corpus or sub-corpora is done following the traits found in MICASE (MICASE, R.C. Simpson, S.L. Briggs, J. Ovens, and J.M. Swales, 2002). In order to facilitate the organization of the corpus to be analyzed, the main lecture (LE) attributes have been

categorized. Thus, important information about the lectures is supplied, e.g. the title, primary discourse mode, speech event, number of words, as well as recording duration. All the lectures gathered for the purpose of this study belong to the Academic division of Social Sciences. However, the scope of Social Sciences is extremely large including an extensive amount of areas such as Anthropology, Business Administration, Communication, Economics, Education, History, Public Policy, Political Science, Psychology, Social Work, Sociology or Urban and Regional Planning. Therefore, I consider relevant the fact of including the title of each lecture in order to narrow the scope of the field for a better positioning in the analysis. Nevertheless, this study does not aim at analyzing lectures according to discipline variation, but to unify the field of Social Sciences as a single entity in order to study how DMs are used within its context.

The attribute *Primary discourse mode* refers to “the predominant type of discourse characterizing the *speech event*” (MICASE, R.C. Simpson, S.L. Briggs, J. Ovens, and J.M. Swales, 2002). All lectures analyzed for this study are monologic lectures where “one speaker monopolizes the floor, sometimes followed by question and answer period” (MICASE, R.C. Simpson, S.L. Briggs, J. Ovens, and J.M. Swales, 2002). After a first search of MICASE, I realized that there was a larger amount of monologic lectures than of other discourse modes (e.g. interactive, panel, mixed), so according to availability it was going to be easier to make a selection among monologic lectures. Consequently, the SC happens to be also a compilation of monologic lectures and the final corpus was established.

Another feature included in the tables below is *Speech event*. In the MICASE corpus *Speech events* are classified according to *classroom events* and *non-class events*. As the corpus is a compilation of lectures, they are included within *classroom events*. According to the number of students in the audience, two groups can be distinguished: small lectures (LES) - a lecture class of 40 or fewer students, and large lectures (LEL)- a lecture class of more than 40 students. The corpus presented here includes both. Among the twelve North-American English lectures we find four LES and eight LEL. Note that in the MICASE browser we find a larger number of LELs than LESs in the field of Social Sciences, this is the reason why there are more LELs in the NAC. On the contrary, the number of students in the Spanish lecture corpus does not generally surpass 40, giving therefore more LESs than LELs. This could be seen a differentiating trait between North-American and Spanish lecturing settings and lecturing styles; North-American lecturing implying a larger number of students per room than Spanish lecturing. However, I cannot make extended generalizations about this differentiating trait between the two sub-corpora, since a larger corpus would be needed.

Taking into account features such as number of words and recording duration, I can observe that some lectures taken from the MICASE are slightly longer than the Spanish lectures, being the average number of words per lecture 10,452 and the average duration 72m for the North-American English lectures and 6,650 words and an average LE duration of 53.6m for the Spanish ones. Spanish lectures seem to be shorter than the North-American lectures maybe due to the different styles of lecturing in both

universities, the University of Michigan in the USA and Universitat Jaume I in Spain as seen in Chapter III sections 3.4 and 3.5 on this dissertation. However, this aspect does not seem to be significant for the aim of this study, whether the lecture has a larger number of words or a longer duration does not necessarily have to affect the use of DMs, neither seems to be an impediment for the analysis of the most recurrent DMs.

Tables 9 and 10 categorize the corpus with the attributes explained above. Table 9 displays the twelve North-American English lectures taken from the MICASE. Table 10 includes the twelve Spanish lectures recorded at Universitat Jaume I, Castellón (Spain) for the purpose of this PhD dissertation.

AMERICAN ENGLISH LECTURES

	Title	Primary discourse mode	Speech event	Number of words	Recording duration
LE1	Intro Anthropology Lecture	Monologic lecture	Large lecture	11,549	74m
LE2	Sports and Daily Life in Ancient Rome	Monologic lecture	Large lecture	12,912	71m
LE3	Intro Communication Lecture	Monologic lecture	Large lecture	9,696	76m.
LE4	Media Impact in Communication Lecture	Monologic lecture	Large lecture	9,684	72m
LE5	Intro Psychology Lecture	Monologic lecture	Large lecture	7,744	47m
LE6	Macroeconomics Lecture	Monologic lecture	Large lecture	8,373	76m
LE7	Labor economics Lecture	Monologic lecture	Small lecture	12,391	77m
LE8	Twentieth Century Arts	Monologic lecture	Large lecture	6,205	41m
LE9	Perspectives on the Holocaust lecture	Monologic lecture	Large Lecture	9,172	100m
LE10	Statistics in Social Sciences Lecture	Monologic lecture	Small lecture	16,438	109m
LE 11	Intro to Psychopathology Lecture	Monologic lecture	Small lecture	8,326	52m
LE 12	Historical Linguistics lecture	Monologic lecture	Small lecture	12,935	69m
Total				125,425w	864m
Average words & time x LE				10,452w	72m

Table 9. North-American English lectures description.

SPANISH LECTURES

	Title	Primary discourse mode	Speech event	Number of words	Recording duration
LE1	Historia de las primeras civilizaciones	Monologic lecture	Small lecture	5,652	59m
LE2	Arte del Renacimiento	Monologic lecture	Small lecture	6,404	43m
LE3	Psicoestadística I	Monologic lecture	Small lecture	8,319	75m
LE4	Auditoría contable	Monologic lecture	Small lecture	8,066	60m
LE5	Marketing por Internet	Monologic lecture	Small lecture	5,628	55m
LE6	Introducción a la economía	Monologic lecture	Large lecture	6,874	62m
LE7	Economía mundial	Monologic lecture	Small lecture	9,228	68m
LE8	Dirección comercial	Monologic lecture	Small lecture	5,213	51m
LE9	Introducción a la psicología social	Monologic lecture	Small lecture	5,671	49m
LE10	Lingüística	Monologic lecture	Small lecture	8,093	49m
LE11	Lenguaje publicitario	Monologic lecture	Large lecture	5,379	40m
LE12	Documentación informativa	Monologic lecture	Small lecture	5,284	33m
Total				79,811w	644m
Average words & time x LE				6,650w	53.6m

Table 10. Spanish lectures description.

5.3. Model and method of analysis

In the previous chapter I have presented some of the most cited classifications of DMs both in English and Spanish (Murphy and Candlin 1979, Chaudron and Richards 1986, Portolés , Morell 2001, González 2005 among others). A close review of these classifications does not clearly provide a unique valid classification for both languages Spanish and English. However, deviations take place among English and Spanish researchers. As far as I know, researchers studying DMs have tried to provide suitable taxonomies either for English or Spanish, but in no case a taxonomy that could be valid and used for both languages. Additionally, the taxonomies previously presented have proven to be inconsistent in some cases, as already commented. Therefore, the objective here goes a step further. I aim at providing a more consistent DM classification model as well as at making that classification valid for both languages, Spanish and English. The necessity of a single unique classification model that could be used for both sub-corpora here is obvious; otherwise the contrastive study would not have had logical sense. Clearly, I cannot compare disparate traits that do not share anything in common (or categories in this case) if I aim at getting a relevant and coherent study. Moreover, as far as I know, no contrastive study on the effects of DMs in lectures such as the one presented here has been carried out. Therefore, one of the aims is to develop a valid DMs categorization both for the SC and NAC.

Some Spanish authors have been influenced by the study of DMs in English; however, I have to consider to which extent generalizations made about English can be carried over to other languages such as Spanish. Already in 1990 Fraser raised the question: “To which extent do all languages share a basic set of DMs with the same core pragmatic meaning?” (1990: 395). It may occur that a DM exists and is often used in a language but we can find no traces of that DM in other languages, or even when the same DM can be found in two or more different languages the function performed may be different. With this, the idea of having English and Spanish DMs counterparts is disregarded, as it is expected some DMs will be the same for both languages whereas others will be different for the same categories.

Bearing this in mind and along with the fundamental theories of DMs, in the following section I will try to develop a taxonomy of DMs that could be applicable for both sub-corpora (SC and NAC) although, as I said before, I do not expect DMs to coincide among categories in both languages.

Categories will be filled in after submitting a search for the most recurrent DMs in both Spanish and English lectures. The aim is to see if there are any differences or similarities in the use of DM categories between the two languages and what kind of deviations and changes I might find. I have also focused the scope of the study on looking for those DMs that normally co-occur in both the SC and NAC and the reasons why this may happen.

The findings of this analysis will be objectively presented in the next chapter (Chapter VI), going on with a general interpretation and discussion in Chapter VII.

5.3.1. DM classification model.

The aim is to create a DM classification model which can be valid for both corpora under study (North-American and Spanish lectures). Previous classifications failed to explain some categories or these were only valid for English DMs and not for the Spanish ones. Moreover, categories under micro and macro- markers proved unsuccessful as they did not obey any firm linguistic rule: whereas some categories were miscellaneous, others were semantic, morpho-syntactic or even pragmatic. Under such a pandemonium some DMs struggled to fit into one of the categories mentioned (Chaudron & Richards 1986, Murphy & Candlin 1979, Morell 2001, Portolés 1998). Therefore, an improved and corpus-centered classification for the ongoing analysis was needed.

There seems to be an agreement in DM literature that coherence is constructed through relations that can sometimes be expressed by linguistic units I call here DMs along with inferences in discourse. The main function of DM results in creating explicit discourse relations to convey coherence through different meanings along the discourse utterances. Already Schiffrin (1987) argues in favor of three DM relational levels: DMs that relate the semantic reality (facts), DMs that relate clauses on a logical (epistemic)

level and those which relate clauses on speech acts or pragmatic level. Similarly Fraser (1990, 1999) distinguished three domains (or relations) to which DMs apply: semantic, epistemic and speech act. Generally, the semantic and logical domains melt together to form the logico-semantic level; reducing and distinguishing now two main levels, the logico-semantic and the pragmatic level (González 2005).

In order to develop my DM classification model, I have departed from the concepts of relational and attitudinal DMs meanings and functions among discourse, without forgetting Redeker's (1990) assumptions upon the discourse coherence model on the search for coherent discourse relations. Everyone can become acquainted with three different relational categories between discourse elements in the communicative act that can be detected easily, these are: i) relation part of discourse-part of discourse, ii) relation speaker- hearer or vice versa, and iii) relation speaker- speech. These three element relations can be conveyed in many different ways: kinesics, visuals or the most common, the use of linguistic units such as DMs. The main goal of these relations is to express meanings along the discourse utterances. In the 90s Halliday distinguished three functional components of meaning (Halliday 1994): ideational, interpersonal and textual (or discoursal in this particular case).

A closer look shows that the relational categories between discourse elements explained above can take place in any of the Hallidayan's functional meanings and thus, serve as the basis of the DM classification model. If I want to provide a valid taxonomy for both Spanish and English, the linguistic reasons and fundamentals have to be largely valid and

universal to any language and social culture. As a result, I propose a classification of DMs which is based on the three functional meanings mentioned and the relations they can convey along the discourse utterances. I can distinguish three different types of DMs according to the meaning these convey and the relational functions present. In Figure 6 below we can see how the three functional meanings are distributed according to the three relational functions.

MEANINGS	RELATIONAL FUNCTIONS
<ul style="list-style-type: none"> • Internal (ideational) 	<i>part of discourse-part of discourse</i>
<ul style="list-style-type: none"> • Structural meaning (global discourse structure relations) 	<i>part of discourse-part of discourse</i>
<ul style="list-style-type: none"> • Attitudinal (interpersonal) meaning 	<i>speaker-hearer and/or speaker-speech</i>

Figure 6. DM functional meanings and matching relational functions.

The first meaning refers to the logico-semantic relations (Schiffrin 1987, Fraser 1990, González 2005) DMs express in the discourse; these kinds of

relations indicate the links between *part of discourse-part of discourse* elements. DMs within this category happen to have lexical or descriptive meaning. According to this definition, categories such as **causal**, **contrastive**, **consecutive** or **additional** DMs would be included here; I have called them micro-markers, following Chaudron and Richards' (1986). or Morell's (2001) previous similar classifications.

The next proposed DM category would also express *part of discourse-part of discourse* relations. The overall structure of the discourse is signaled through structural relations by means of DMs such as *to begin with* and *vamos a comenzar*. The categories conveying structural relations are for example, **starter**, **organizer**, **topic shifter**, etc. These kinds of DMs will be here referred to as macro-markers borrowing Chaudron and Richards' (1986) terminology. Studies have shown that the presence of macro-markers improves retention and recall in post-lecture tests (Chaudron & Richards 1986, Jung 2003) and that it is generally beneficial for activating content schemata (DeCarrico & Nattinger 1988) and helping listeners to successfully follow the lecture (Khuwaileh 1999).

Those relations between *speaker-speech* and *speaker-hearer* (or vice versa) are conveyed through another type of DMs I have called here **operators**. These markers are more specifically related to conversational, spoken discourse rather than written discourse (Llorente 1996), these have been traditionally called in the literature 'pragmatic markers'. The categories I have included here are **attitudinal**, **pause filler**, **elicitation**, **acceptance** and **confirmation-check**. These DMs are those which rhetorically signal the speaker's intentions and goal (the illocutionary force) as long as they play a

dominant inferential role in the discourse, frequently monitoring proximity between speaker-hearer and speaker-speech. Some of these DMs already appeared in Portolés's (1998) classification (see previous chapter) under categories such as **discursive operators**, **proximity monitoring markers** and some **reformulators**.

The resulting scheme consists of five different categories keeping homogeneous grammatical categories for each classification of DMs: micro-markers, macro-markers and operators. Therefore, the DM classification model shows as follows.

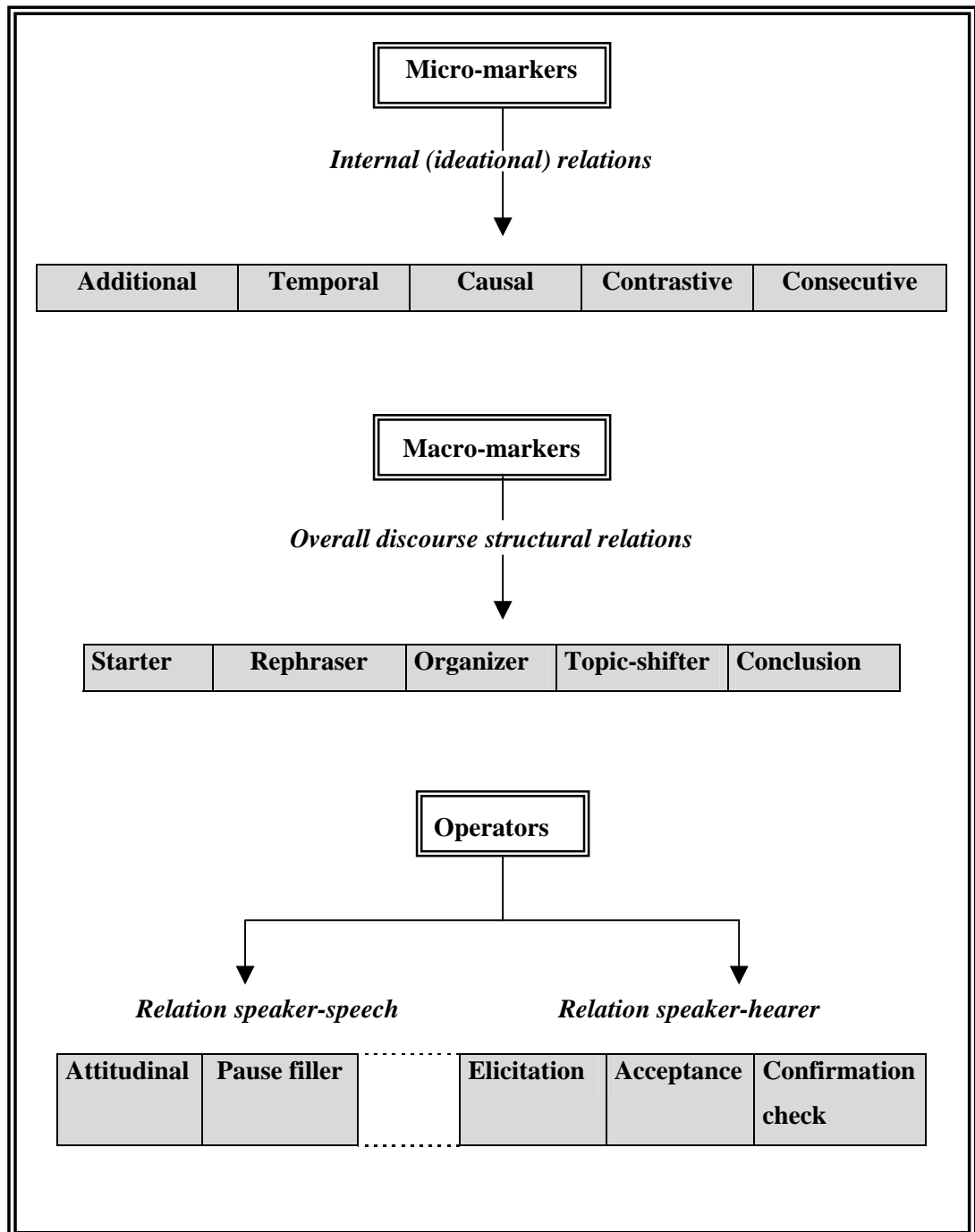


Figure 7. DM classification model for this study.

Categories had to be filled in, taking instances from the corpus. I took into consideration those DMs from each category with the highest number of

occurrences. Along this line, a minimum of three DMs were established for each category; the reason was that, given the complexity and variety of DMs, it seemed necessary and methodical to center the scope of the study and avoid getting lost in a myriad of DMs. An attempt to embrace too many DMs in a large corpus, like the one being used here, could have led us to a state of disorientation and could have interfered the aim of the current study, that is, to find out if there are any significant differences in the use of DMs between North-American and Spanish lectures and the factors that may cause these variations.

I first tried to go through the English DM classification departing from a preliminary study on English DMs (Bellés-Fortuño 2004). There, I carried out a contrastive study between British and North-American English lectures and provided a classification of the most recurrent English DMs in both North-American and British English lectures. I consider this English DM classification as a research-based point of departure, however and as mentioned in the previous chapter, the taxonomy followed then was based on previous classifications (Murphy & Candlin 1979, Chaudron & Richards 1986, Morell 2001) and therefore categories did not coincide with the new DM relational classification model (Figure 6). As a consequence, adaptations from the previous classification to the new proposed classification model had to be done. I aimed at fitting DMs within the new classification; once this was done I had to validate the resulting taxonomy to check that I was not wrong in the assumptions that these DMs were as relevant in the NAC as they were in my previous study (Bellés-Fortuño 2004). I submitted a search on the concordancer option in the software

Wordsmith Tools 4.0. The result matched the hypothesis; the English DMs validated in the 2004 classification were equally recurrent in the 2006 NAC with only some exceptions. However, a further step was taken and the NAC lectures were individually proof-read in search of some recurrent marker I could have obviated in my preliminary study in order to add it to the novel classification. A detailed and filled in classification of English DMs in the NAC is presented in the results chapter that follows.

The challenging part of this section was to develop the Spanish DM classification model with appropriate filled categories, since it was the first time I approached a Spanish lecture corpus. I have already mentioned and criticized previous classifications of Spanish DMs, such as the one presented by Portolés (1998) (see Table 7 in Chapter IV). Along with previous English DMs classifications, Portolés' (1998) taxonomy of Spanish DMs aroused inconveniences for the development of the Spanish DM classification model, that is to say, some Spanish DMs pointed out by Portolés (1998) were not found in the SC such as for example instances of *así las cosas*, *encima*, *con todo*, etc; since he used a different genre as a corpus. As a consequence, I decided to proof-read the twelve Spanish lectures that shaped the SC in search of the most recurrent DMs. Thus, I could make a miscellaneous list of the Spanish DMs used in the twelve lectures to later check their frequency of use and classify them according to categories. The resulting Spanish DM classification model is presented in the next chapter.

It is worthy to point out that the English DMs found and included in the classification model were not expected to be counterparts with the Spanish

DMs. Some can coincide while others are different. This is mainly due to cross-cultural differences between English and Spanish as well as disciplinary variations and language usage which signal DM types and occurrences. Note that both the English and the Spanish DM classification models are very closely restricted to the corpora under study as well as to the individual usage or idiolectal variation of lecturers' own preferences for DMs.

In the next chapter the detailed English and Spanish DM classifications are presented along with the results of the search in both sub-corpora according to variables of frequency rate and number of occurrences. Peculiarities arisen from the search are then commented on pointing out similarities or differences in the use of micro-markers, macro-markers and operators in both the SC and NAC. Moreover, I will particularly show how some specific DMs seem to co-occur, both in English and Spanish lectures. Comparison of both sub-corpora will be discussed in Chapter VI.

Chapter VI: Analysis of Results

6. ANALYSIS OF RESULTS: DMs Classification model application

6.1. Introduction

For the application of the analysis we used the DMs classification model proposed in the previous chapter. A maximum of three DMs were selected for each category, those DMs that proved to be the most recurrent ones in both sub-corpora (SC and NAC).

In this section, I present the results of analysing how micro-markers, macro-markers and operators are used in the two sub-corpora in order to detect any similarities or differences in the way DMs work and behave in both NAC and SC, and the reasons why these may occur. For this purpose, I have developed numeric tables including the variables to be analysed that will be of great help for the final analysis of results. The data in those tables will shed light on the overall study and allow concluding with some generalisations about the use of DMs in North-American and Spanish lectures.

This chapter is organised in five main sections, the first one presents an overall view of the general results obtained from the analysis of the NAC and SC, the following three sections are parallel and show DM results according to the proposed taxonomy (micro-markers, macro-markers and operators); and a final section focuses on DM collocations. The variables used for the analysis of DMs are the number of occurrences and the frequency rate of each pre-established marker in the two sub-corpora. Moreover, I highlight differences found in the

comparative analysis between the NAC and SC, which are illustrated with examples from the corpora.

Along the last section of this chapter, we demonstrate that there are some DMs that seem to co-occur with others in the discourse of academic lectures in English and Spanish. The most recurrent DMs collocations are detected and presented as well as compared in both sub-corpora.

6.2. DMs results in the NAC and SC: an overall view

A general overview of the analysis of DMs in both the NAC and SC gives the following results: DMs seem to be more often used in the NAC than in the SC as the resulting rate shows (see Figure 8). According to frequency rate 43 DMs are found in the NAC for every 1,000 words whereas in the SC the rate is slightly lower, 38 DMs for every 1,000 words.

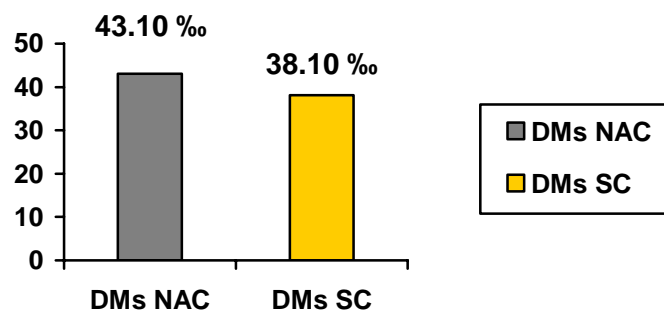


Figure 8. DM results in the NAC and SC.

Without distinguishing and comparing both sub-corpora, the total rate for each individual categories of DM, that is, micro-marker, macro-markers and operators, is as follows: micro-markers are the most often used type of DM, followed by operators and macro-markers in the last instance. To better illustrate these findings, the results are shown in Figure 9 below.

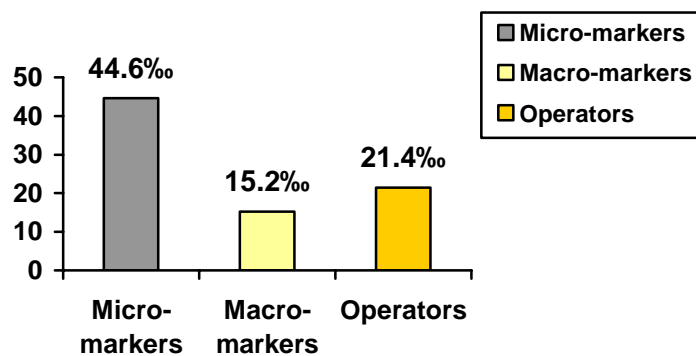


Figure 9. Total rate of micro-markers, macro-markers and operators in the NAC and SC.

A closer look at individual distinctions among DMs and more specifically the micro-markers have revealed that the NAC tends to use more micro-markers than the SC, where the resulting rate is much lower (see Figure 10 for frequency rates).

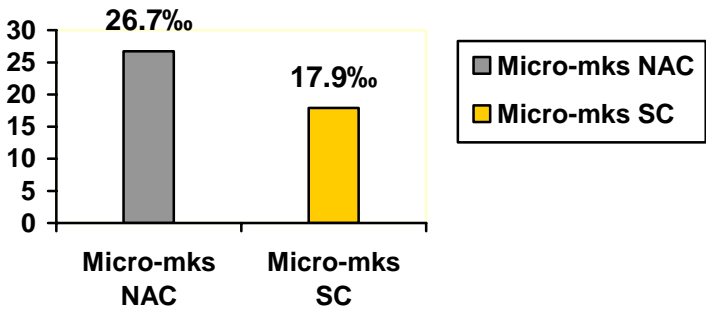


Figure 10. Total rate of Micro-markers in the NAC and SC.

Regarding macro-markers and again comparing general results according to frequency rate, these have a lower rate in both sub-corpora compared with the rates obtained for micro-markers. Macro-markers are less used than micro-markers in the NAC and SC, and the resulting rate for each sub-corpus very similar, although slightly higher in the NAC. Results are illustrated in Figure 11 below.

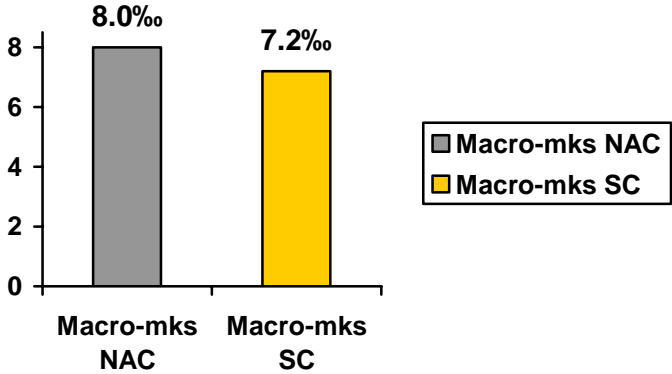


Figure 11. Total rate of Macro-markers in the NAC and SC.

The last type of DM under analysis according to the classification model is operators. The figures show that operators are not as largely used as micro-markers in the NAC and SC, but they have a slightly higher rate than macro-markers in both sub-corpora. Comparing the general results of operators we can observe that they are more often used in the SC than the NAC as the frequency rates show (see Figure 12 for operators' results).

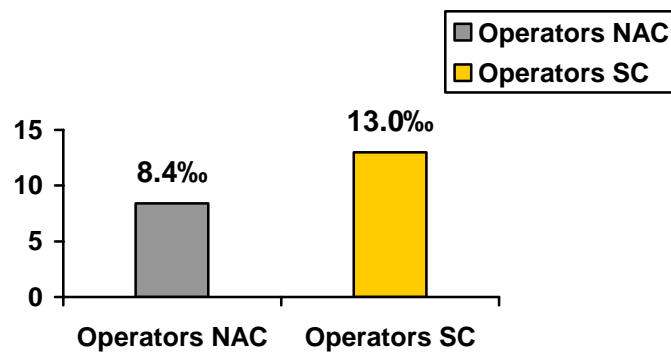


Figure 12. Total rate of Operators in the NAC and SC.

After this overall view at the results obtained after the analysis of DMs in the NAC and SC, in the next section we present the results obtained as to specific categories (micro-markers, macro-markers and operators) and individual cases of DMs according to the variables: frequency rate and number of occurrences.

6.3. Micro-markers

6.3.1. Micro-markers in the NAC

The first corpus under analysis is NAC, that is, the twelve North-American English lectures. The English DM classification model was implemented according to the following semantic categories: **Additional**, **Temporal**, **Causal**, **Contrastive**, and **Consecutive**.

An overall view of the NAC (see Table 11) gives a total of 3,355 micro-markers used in the twelve lectures, this represents a frequency rate of 26.7 micro-markers every 1,000 words in the whole NAC.

Table 11 below shows the most recurrent micro-markers in the NAC. The results obtained in the search of the three most relevant micro-markers in the NAC show that the categories having a higher frequency rate in use are **Additional** (12.8‰), **Contrastive** (5.6‰) and **Causal** (3.1‰) followed by **Temporal** (2.7‰) and then **Consecutive** with the lowest frequency rate (2.2‰).

a) Micro markers (*Internal (ideational) relations*)*NAC Results*

Additional	#	%	Temporal	#	%	Causal	#	%	Contrastive	#	%	Consecutive	#	%
	DMs			DMs			DMs			DMs			DMs	
<i>and</i>	1,255	10.0	<i>then</i>	222	1.8	<i>because</i> (<i>cuz</i>)	342	2.7	<i>but</i>	655	5.2	<i>so</i>	149	1.2
<i>or</i>	289	2.3	<i>after</i>	86	0.7	<i>since</i>	27	0.2	<i>although/</i> <i>though/even</i> <i>though</i>	71	0.6	<i>then</i>	65	0.5
<i>now</i>	63	0.5	<i>before</i>	31	0.2	<i>because</i> <i>of</i>	20	0.15	<i>however</i>	17	0.1	<i>so that</i>	63	0.5
TOTAL 1	1,607	12.8		339	2.7		389	3.1		743	5.9		277	2.2

TOTAL 2	# DMs	%
Micro-markers	3,355	26.7

Table 11. Micro-markers results in number of occurrences and frequency rate in NAC.

Taking a closer look at individual categories, I observed that the use of *and* within the **additional** category is the most relevant for the final amount of micro-markers used; *and* has given a frequency rate of 10‰ (1,255 occurrences), the highest rate compared to the rest of the micro-markers becoming the most frequently used micro-marker in the NAC when conveying additional meaning.

In a first general search, *and* came to have a much higher number of occurrences (3,381) in our NAC; however, not all these instances of *and* conveyed additional meaning or functioned as a DM. Thus, we consider *and* as a DM not when it is used to join two words together, to enumerate elements as it occurs in the sentence structure *noun + and + noun* or when showing that one thing happens after another. Consequently, instances such as *years and years*, *back and forth*, *water and mud*, etc...have been disregarded. We found that *and* as an additive marker normally occurs after or before a short or long pause in order to add something new to the ongoing topic joining two clauses together, being this the criteria followed to make a valid search. See examples below taken from the NAC in which *and* is considered an **additional** micro-marker:

AND

- (4) second, to show you what, the area that i'm talking about. Samburu area is here. **and** (Marawal) is the main center, where t- where S- Samburu come to trade (LE1/NAC)

- (5) is limiting and it's it's uh, it's bothersome... she's, she's become a saint. **and**, it's interesting that when he, he wrote The Ghost Writer um, he at first, m, is limiting and it's it's uh, it's bothersome... (LE9/NAC)

Moreover, *and* as a DM can also work as an operator conveying relations speaker-speech where it is used as a hesitator or pause filler rather than having a strict semantic meaning (see Table 15 for operators). We have also observed that *and* can appear in isolation or collocating with other DMs such as the case of *then*, as we will explain in section 6.6 on DM collocations.

Still with the additional category we find the DM *or*. Following the same criteria we used for the micro-marker *and*, only instances of *or* functioning as an additional marker have been taken into consideration. *Or* as an enumerator and with the syntactical structure *noun + or + noun* has been disregarded.

OR

- (6) between the distribution. how things are distributed may, increase consumption *or* decrease consumption, as opposed to, you know, there's that's much more... (LE1/NAC)

Only instances in which *or* is used with a semantic additional meaning are considered, as the examples below show. Notice that in example 7 the first *or* is one of the situations mentioned before and therefore not a DM, whereas “or, I mean...” is an instance of additional *or*.

- (7) with three it just means that, there're some large ones with maybe six or seven, *or*, i mean there's a school down, not very far from here, that only includes uh (LE10/NAC)
- (8) i want to encourage you again to stop and ask whenever, you don't understand, *or*, need some clarification *or* whatever of anything i'm saying or, this goes (LE10/NAC)

The third most recurrent additional micro-marker in the classification is *now* with 63 occurrences (0.5‰). *Now* can function as a temporal adverb but as such it has not been considered a DM in this study. In contrast, Swales and Malczewski consider *now* as a temporal marker when functioning as an adverb, they say: “it often serves its traditional function as a temporal marker” (2001: 159). However, when *now* introduces a clause, normally followed and in some cases also preceded by a pause (whether long or short), without changing the speaker’s topic but adding new information to previous clauses in the discourse, in this case, *now* has been considered here as an additional DM. In example 9 below we can see *now* functioning as a temporal adverb and therefore not a DM. On the contrary, example 10 shows *now* as an additional micro-marker, notice that there is no topic change but it adds new information for the ongoing discourse:

NOW

- (9) You can have more than one, in one community. so as i just said the foragers are **now** mixing foraging, hunting and gathering, with some, basic food production some (LE1/NAC)
- (10) the grave of Sir Isaac Newton, which in some ways is highly appropriate, because, **now**, Isaac Newton's impact on physical science, is of the same magnitude as Darwin (LE5/NAC)

Following in frequency is the **contrastive** category; there is also a micro-marker that stands out: *but*. It represents 5.2‰ out of the total frequency rate (5.9‰) for the **contrastive** category; but also becomes the second most frequently used micro-marker after *and*. Thus, *but* places **contrastive** second in the category

ranking. However, *but* does not always function as a contrastive DM, in some cases this DM gets released of its semantic meaning becoming a pragmatic operator working as a pause-filler, this occurs when *but* collocates with words such as *um*, *uh* and after *or* in many cases also before a very long pause giving the speaker time to give an explanation or think about what comes next. In any case, *but* as an operator is less used than the semantic contrastive micro-marker *but* (see example below).

BUT

- (11) those are the those are the, at times the physiological effects of anxiety, *but* that's not a really good anxiety test. a good anxiety test is gonna ask about
(LE11/NAC)

The other recurrent micro-markers in the NAC are *although* and its varieties *even though* and *though* together with *however*. Instances of *although* and *however* are not as relevant as the use of *but* in the NAC according to the number of occurrences. Here below are examples of the micro-markers *although* and *however* in the NAC.

ALTHOUGH

- (12) with you in a minute okay? for example with the verb oh you're perfectly right *although*, many Spaniards today, obviously any educated Spaniard writes like this <WRITING ON BOARD> and distinguishes between these two forms...(LE12/NAC)

HOWEVER

(13) /x/ was replaced the /f/ replaced by /x/ because they're structurally similar.

however he also was aware of one other factor which he didn't pass over he didn't (LE12/NAC)

As said above, the **causal** category comes third due to the high frequency rate of the marker *because*, this represents 2.7‰ (342 occurrences) out of the total causal category (3.1‰). Consequently, we looked for those instances of *because* in isolation and then considered *because of* another independent marker due to the relevant number of occurrences (the third most relevant one). Later, we realised that according to MICASE transcriptions norms the relaxed form of *because* was also transcribed as such, therefore instances of colloquial *cuZ* (79 occurrences) appeared in the transcripts and were also taken into account (see example from the NAC below).

CUZ

(14) settled. and if you were to visit a kindergarten around here, you better hurry *cuZ* i think school's gonna close before our class closes um, y- i think you would (LE11/NAC)

In addition to *because* and *because of* another relevant causal micro-marker according to frequency use is *since*. However, the occurrences of *since* and *because of* are few compared to the use of *because* as a causal DM. No doubt, *because* is the micro-marker par excellence used to convey causal meaning among discourse utterances in the NAC. When searching for *since* conveying causal meaning as a DM we had to obviate instances of *since* as a temporal preposition

or adverb (*since then, ever since, since May*, etc.), only those cases where *since* appeared as a subordinate conjunction conveying cause were taken into consideration. Compare the two examples below; in example 15 *since* functions as a temporal preposition but not as a DM, in the next example an illustration of *since* as a causal micro-marker is shown.

SINCE

(15) women, print ads. so your question would logically be, okay so what's gone on *since* nineteen seventy-nine? well, Mian Kang who used to be a grad student here (LE4/NAC)

(16) if you feel like, by the end of the lab that you're not kinda on top of this, *since* the assignment is on Monday and, David and Katy are not planning to spend (LE10/NAC)

In summary, I find three micro-markers that lead the ranking according to frequency rate and number of occurrences, these are, *and*, *but* and *because*. The number of occurrences of these micro-markers seems to be relevant for the evaluation of the most frequently used categories in the NAC, which are in order:

Additional, Contrastive and Causal.

On the other hand, the categories showing the least frequency rate of DMs are in order **Temporal** and **Consecutive**. Prominent as well as ambiguous for the **temporal** category is the DM *then*. Originally placed within the **temporal** category, the micro-marker *then*, although mostly functioning as a **temporal** DM (222 occurrences), can also convey cause-effect relations between parts of the discourse. When searching for *then*, we observed the contexts in which it

appeared and its different functions, all instances of *then* were proof-read to determine whether *then* was functioning as a temporal or a consecutive micro-marker showing the result of the events that had been described before; those instances of *then* working as an adjective were also disregarded. Curiously, *then* had a high number of occurrences when collocating with the also micro-marker *and*, in which cases *then* had mainly a temporal meaning. The first two examples below show the micro-marker *then* with a resultative consecutive function, the third and last example of *then* conveys temporal meaning.

THEN

- (17) not thought of as, as highly desirable, and it brings in a profit, you cannot *then* channel that money into things related to the regeneration of your family. (LE1/NAC)
- (18) would the argument go. and therefore third part of the argument, because of that, *then*, this ideological domination, comes to sort of surround us all. that we're (LE3/NAC)
- (19) this, Kirtland Air Force Base, is, a black-and-white photograph that he's taken, *then* he goes back to his studio, uh he blows the image up, usually it's a sort (LE8/NAC)

In example 18, we observe that the speaker (lecturer) is very concerned with the idea of reinforcing the causal-consecutive relation between the clauses; *then* comes to reinforce the consequence already conveyed by the causal DM *because*

of, any of the two micro-markers could be eliminated and the meaning conveyed would practically be the same.

Within the **temporal** category we have also included other recurrent DMs such as *after* and *before*. These function as temporal micro-markers when introducing a new clause normally following the syntactic structure *before/ after + subject + verb + complements* or the structure *before/ after + verb-**ing** form + complements*, where *before* and *after* are considered subordinate conjunctions and not prepositions. When followed by the determiners *this/ that*, we understand *this* and *that* as pronouns substituting a whole clause that has been introduced before, in that case, *before* and *after* are still considered conjunctions. See examples of *before* and *after* as temporal micro-markers.

BEFORE

(20) common, that children have experienced, some kind of, educational something, ***before*** they come to kindergarten. and that educational something_ well if we (LE10/NAC)

AFTER

(21) assassins of Caesar were finally defeated at the battle of Philippi. um, then um ***after*** maintaining himself in this wonderful style, um in the next few years he (LE2/NAC)

As a whole, we observe that **temporal** represents 2.7‰ (339 occurrences) closely followed by **consecutive** with a rate of 2.2‰ (277 occurrences), representing the least often used category of micro-markers in the NAC. Within the **consecutive** category there is a micro-marker that outnumbers the rest, *so*. *So* has largely been

studied as being one of the most ambiguous DMs as well as one of the most commonly used ones together with *and* and *but*. From the NAC analysis we have observed that *so* can have more than one function among discourse utterances as well as convey different meanings. We have considered here *so* as a semantic DM that affects ideational internal relations functioning as a micro-marker within the **consecutive** category and marking fact-based, knowledge-based or action-based consecutive relations, mostly exchangeable by *therefore*. These instances of *so* generally occur at the end of a section of speech. At this point we understand a section of speech as characterized with louder more prominent speech at the beginning of the section and quieter, faster speech at the end of it. The beginning of each section introduces a new topic or idea, thus resembling the paragraph-initial indentation of the written paragraph. (Brown & Yule 1983, Chafe 1979, Hinds 1979). However, for boundaries of sections to be identifiable other discourse and prosodic features have to coincide, such as the pause, shift of amplitude, shift in speed of delivery, pitch changes and even non-verbal actions (Rendle-Short 2003). Other uses of the DM *so* are taken into consideration later on in the chapter when *so* functions either as a macro-marker or as an operator. The following two examples taken from the NAC show the semantic consecutive meaning of *so* at the end of a speech section.

SO

- (22) probability. okay so the assumption is if you're not there you get no utility. and *so* your utility is discounted not only for, t- time preference, but also for t- time preference, but also for survival probabilities. uh combining those things (LE7/NAC)

- (23) use technologies of various kinds to make sure that you're continually rejuvenating the soil. *so* fertilizers would be one means. um, in Madagascar, the rice terraces are ways (LE1/NAC)

Apparently, as it happened with other English micro-markers *so* tends to co-occur with the additional DM *and*, when this happens *so* generally functions as a consecutive micro-marker and not as a macro-marker or operator. The other recurrent consecutive micro-markers in the classification are *so that* and as commented before, the ambiguous *then* (see examples 17, 18 and 19). Note that due to its number of occurrences *so that* has been considered as a single identifiable DM unit (see example 24 below).

SO THAT

- (24) their house. okay? so the idea that that people more or less consume their income *so that* their consumption is equal to their income so their utility of consumption (LE7/NAC)

6.3.2. Micro-markers in SC

Still on the first stage of the analysis, we aim now at examining micro-markers, but this time in the SC (Spanish Corpus).

On the whole, there are 1,433 micro-markers occurrences in the overall SC; this represents a frequency rate of 17.9‰ out of the total corpus. As can be seen in Table 12, the results obtained in the search of micro-markers within the SC show that the most frequent categories are **Contrastive** (5.3‰), **Causal** (3.8‰) and

Temporal (3.8‰), not coinciding with the NAC (see Table 12 at the end of this section for SC micro-marker results).

The **Contrastive** category has the highest frequency rate (5.3‰), where the micro-marker *pero* (365 occurrences) is the most frequently used marker not only within the category but in the overall SC followed by **consecutive** *luego* and **temporal** *entonces*. *Pero* represents a rate of 4.6‰ out of the total 17.9‰ of micro-markers in the SC, being the DM par excellence in Spanish lectures and relevant for the **consecutive** category.

PERO

- (25) el plátano le sugiere esto pues es un signo natural no es un signo arbitrario *pero* claro si coges ya dos tipos de serpientes que son muy parecidas y los gritos son muy diferentes hombre no podemos afirmarlo al cien por cien lo único que muchos lingüistas dirían se negarían en redondo a que los chimpancés (LE10/SC)

Within the consecutive category we also find *aunque* with 41 occurrences (0.5‰) and *sin embargo* with 17 and a frequency rate of 0.2‰, neither of these DMs being as relevant as the use of contrastive *pero*. (See examples 26, 27).

AUNQUE

- (26) no se cumple porque no se dan todas las condiciones que acabamos de ver aquí, *aunque* se dice [eleva el tono] que a muy y a largo plazo y en términos relativos (LE7/SC)

SIN EMBARGO

- (27) otra a lo largo del site, visitando aquellos lugares que nos interesa que vea, *sin embargo* una de las características de internet y aquello que buscan (LE5/SC)

Causal is the second category in the ranking very close to **Temporal**, both representing 3.8‰ but with a higher number of occurrences for **Causal** with 314 compared to 304 for **Temporal**. Within **Causal** the micro-marker that stands out is *porque* being the most representative in this category since other causal DMs are rarely found in the SC. The causal micro-marker *porque* stands out, compared to *por eso* with 27 occurrences (0.3‰) or *ya que* with only 3 instances (0.03‰). Examples of **Causal** micro-markers taken from the SC are found here below.

PORQUE

- (28) amigo eh... familiar directo, si no era hijo era sobrino y en todo caso era yerno, *porque* se casó con una de sus hijas con una hermana o hermanastra, esto era normal (LE1/SC)

POR ESO

- (29) de demostrar teóricamente, que b, que b es igual a la correlación entre x e y, *por eso* os decía, hay una redundancia entre la correlación y la regresión (LE3/SC)

YA QUE

- (30) veámoslo en este pequeño esquema. Insisto como os he dicho en alguna ocasión *ya que* [eleva el tono] en todos los periodos históricos cuando analicéis (LE1/SC)

Following on the **temporal** category, *cuando* is more common than the rest (*entonces* and *luego*), being used in 218 instances at a rate of 2.8%, *cuando* is undoubtedly key for the temporal category (see example below).

CUANDO

- (31) decir si es una conversación este amigos, la información presupuesta es la que yo *cuando* la introduzco en mis frases, como tú ya la sabes la presento de forma (LE11/SC)

Within **Temporal** we also find a key micro-marker, *entonces* (44 occurrences). This micro-marker is one of the most complex and recurrent ones along with *luego* (42 occurrences). Both *entonces* (0.55‰) and *luego* (0.5‰) have not only similar frequency rates but they can also fit in more than one category depending on the semantic meaning conveyed. *Entonces* can be used as a temporal DM when conveying temporal relations between discourse utterances joining syntactic clauses together, in some of these cases *entonces* has a somewhat similar meaning to the temporal conjunction *después* or a meaning similar to the expression *más tarde*, although not always. *Entonces* can also be a temporal DM when collocating with temporal adverbs such as *ahora* or *ya* emphasizing time meaning. Take for instance the following example from the SC.

AHORA ENTONCES

- (32) la fórmula rápida que os he dado está anteriormente especificada, pero ahora, *ahora entonces*, también se puede comprobar incluso sería más fácil de demostrar (LE3/SC)

Temporal instances of *entonces* are not as frequently used as *entonces* when it functions as a consecutive micro-marker, as can be seen in Table 12 at the end of this section. We have also noticed that *entonces* tends to collocate with the coordinate conjunction *y* and this normally occurs after a pause, however, in that case *entonces* seems to convey consecutive meaning rather than temporal. Read example 33.

Y ENTONCES

- (33) O sea, si un acontecimiento, es una situación incontrolable, es externa, **y entonces** hagamos lo que hagamos, siempre ocurrirá lo mismo. (LE3/SC)

Still with **Temporal** and also complex, as we said before, is the micro-marker *luego*. It can convey temporal, consecutive and also less commonly additional meaning. As a temporal DM, *luego* has a similar meaning to the adverb *después* as can be appreciated in the following examples.

LUEGO

- (34) simplemente dedicaremos unas [eleva el tono] pocas líneas a cada uno de ellos **y luego** ya iniciaremos en los que sea adecuado pues más adelante. Y empiezo por el... (LE2/SC)
- (35) lo que sea, un trabajador, eh... la mecánica por si no lo sabéis es la siguiente, **luego** hablaremos con más detalle el tema del control interno pero, el cheque... (LE3/SC)

Notice that in example 34 above the temporal use of *luego* co-occurs with the coordinate additional conjunction *y*; in any of the two examples above *luego* can be easily substituted by the adverb *después* with no change of meaning.

The two remaining categories with a lower frequency rate are in order **Consecutive** with 3.6‰ (293 occurrences) and **Additional** with 1.2‰ (99) in the whole corpus. Within **Consecutive** the micro-marker that stands out is *entonces*, a complex and polysemous DM as we commented above for the temporal category. *Entonces* is considered to be a consecutive micro-marker when functioning as an argumentative conjunction mostly interchangeable by Spanish expressions such as *en tal caso* (*in that case*), *siendo así* (*being so*), these uses of *entonces* are more recurrent than *entonces* as a temporal micro-marker (see Table 12). Instances of *entonces* had to be checked to find out the semantic meaning conveyed in each case. Example 35 illustrates the use of *entonces* as a consecutive micro-marker.

ENTONCES

- (36) por cada euro por lo tanto eso significa que la moneda cotiza ¿con qué?
[pausa, esperando respuesta] Con descuento y lo contrario si es menor que
cero **entonces** la moneda va a cotizar con premio [pausa]. Con estos
conocimientos vamos... (LE7/SC)

The other recurrent micro-markers within the consecutive category are *por lo tanto* and *luego*. *Por lo tanto* occurs 84 times with a frequency rate of 1.0‰. It has proven to be more recurrent in certain lectures than in others not being homogeneously used across the twelve lectures that form the SC. An example of *por lo tanto* is shown below.

POR LO TANTO

- (37) mayor bienestar [eleva tono] general gracias a esta política imperialista. *Por lo tanto* una consecuencia no directa es una [eleva el tono] satisfacción...
(LE1/SC)

With a lower frequency rate (0.2‰) and 13 occurrences comes the consecutive micro-marker *luego*, also commented as a temporal DM. When functioning as a consecutive DM, *luego* could be substituted by *por lo tanto/ therefore* having the same meaning. Instances of *luego* expressing the consequence happened to occur in some concrete lectures, especially when *luego* is used to explain mathematical formulas as in example 38 below.

LUEGO

- (38) ¿de acuerdo? Multiplicado por una cantidad menor de la unidad, en valor absoluto, *luego* nos tiene que dar una pendiente pequeñita, pero no de 0, no una paralela (LE3/SC)

Unlike the case in the NAC, the category with the lowest frequency rate in the SC is **Additional** with only 99 occurrences (1.2‰), including micro-markers such as *además*, *incluso* and the already discussed *luego*. The DM *además* is different with its 49 occurrences and a 0.6‰ rate. It is worth mentioning that in 18 of the total 49 occurrences, *además* collocates with the additional conjunction *y* (see example 39 below).

ADEMÁS

- (39) [pausa, esperando respuesta] Inflación y otro problema más para su economía y *además* si se eleva el tipo de interés ¿qué pasa con la inversión? (LE8/SC)

The next micro-marker in order of frequency within the **additional** category is *incluso*, instances of *incluso* as an additional DM have been taken into consideration when its meaning coincides with *además*, as can be read in the following example.

INCLUSO

- (40) Si... nuestro producto o servicio ha conseguido igualar sus expectativas o *incluso* superar las expectativas que él tenía nuestro cliente estará satisfecho (LE8/SC)

As said before in this section *luego* can also convey semantically an additional meaning and function as a micro-marker joining clauses within the discourse. However, such instances of *luego* are not very frequent and have shown to be specific to one or two lectures, so it could be due to disciplinary variations or idiosyncrasies in lecturer's discourse. Additional *luego* has 16 occurrences with a 0.2% frequency rate (see example below).

LUEGO

- (41) y este tema lo vamos a desarrollar en la parte de apoyo social e interés sociales, en otro de los temas de la asignatura. *Luego* está el voluntariado, que también lo vamos a desarrollar en otro de los temas (LE9/SC)

In brief, the most important micro-markers in the SC seem to be *pero*, *porque* and *cuando*. The number of occurrences of these micro-markers is significant for the three most frequently used categories in the SC, which are in order: **Contrastive**, **Causal** and **Temporal**.

On the contrary, the analysis of the SC has revealed that the category that shows the lowest frequency rate of micro-markers is **Additional**, representing only a 1.2‰, which corresponds to 99 occurrences from the total SC.

On the whole, there are 1,433 micro-marker occurrences in the overall SC; this represents a frequency rate of 17.9‰ in the total corpus.

a) Micro-markers

SC results

Additional	#	%	Temporal	#	%	Causal	#	%	Contrastive	#	%	Consecutive	#	%
	DMs			DMs			DMs			DMs			DMs	
<i>(y) además</i>	49	0.6	<i>cuando</i>	218	2.7	<i>porque</i>	284	3.5	<i>pero</i>	365	4.6	<i>entonces</i>	196	2.4
<i>incluso</i>	34	0.4	<i>entonces</i>	44	0.55	<i>por eso</i>	27	0.3	<i>aunque</i>	41	0.5	<i>por lo tanto</i>	84	1.0
<i>luego</i>	16	0.2	<i>luego</i>	42	0.5	<i>ya que</i>	3	0.03	<i>sin embargo</i>	17	0.2	<i>luego</i>	13	0.2
TOTAL 1	99	1.2		304	3.8		314	3.8		423	5.3		293	3.6

TOTAL 2	# DMs	%
Micro-markers	1,433	17.9

Table 12. Micro-marker results in number of occurrences and frequency rate in SC.

6.3.3. Comparison of Micro-markers in the NAC and SC

After an individual analysis of the two sub-corpora, I present here a closer comparison between the NAC and SC. An overall view of the results can be seen in **total 2** (see Tables 12 and 13 above). As can be observed, the total frequency rate of micro-markers when contrasting SC and NAC shows an important difference in the use of micro-makers, resulting in a larger number of occurrences in the NAC due to some particular categories such as **Additional** or **Contrastive**. The total frequency rate of micro-markers in the NAC is 27.2 ‰ (3,419 occurrences), higher than the frequency rate in the SC, that is, 17.9 ‰ (1,433 occurrences).

Tables 11 and 12 show the results obtained in both sub-corpora in the search of micro-markers. As said above, the analysis of micro-markers in the NAC revealed three semantic categories with the highest frequency rate, namely, **Additional**, **Contrastive** and **Causal**. In the analysis of the SC the three most frequently used categories are not fully identical to those in the NAC, which are, **Contrastive**, **Causal** and **Temporal**, **Causal** and **Contrastive** rank high in both sub-corpora. Within the NAC the three categories in order of frequency were: **Additional** (12.8‰), **Contrastive** (5.9‰) and **Causal** (3.1‰). In the SC the category leading the ranking is **Contrastive** (5.3‰), although the frequency rate is similar to the one in the NAC where **Contrastive** comes in a second place. After them and still in the SC comes **Causal** with 3.8 ‰ and **Temporal** also with 3.8‰. Unlike the SC, where temporal micro-markers end up third, in the NAC **Temporal** represents one of the lowest rates (2.7‰) only followed by **Consecutive** with 2.2‰ (277 occurrences).

The main difference observed in both sub-corpora is marked by the **additional** category, which has the highest frequency rate (12.8‰) in the NAC and the lowest (1.2‰) in the SC. Notice that the three most frequently used micro-markers in English are in order *and* (10.0‰), *but* (5.2‰) and *because* (2.7‰), whereas in the SC we find *pero* (4.6‰) *porque* (3.5‰) and *cuando* (2.7‰), *pero/ but* and *porque/ because* in both sub-corpora. We also have to point out that the number of micro-marker occurrences in the NAC is larger than in the SC, mainly due to the extended use of the **additional** DM *and*, which is not found in such a large number of instances in the Spanish lectures or at least not in isolation with an additional meaning as it occurs in the NAC; rather in the SC *and* accompanies other micro-markers, co-occurring with them.

6.4. Macro-markers

6.4.1. Macro-markers in NAC

In we start with a general view of the results we can see in Table 13 that the global use of macro-markers in the NAC reaches 1,007 of occurrences, representing a frequency rate of 8.0‰ in the overall NAC.

Taking the DMs classification model, the macro-markers were categorized in the following divisions: **Starter**, **Rephraser**, **Organizer**, **Topic-shifter** and **Conclusion**. The variables studied were number of occurrences and frequency rate, as we did with micro-markers. Results from the analysis of macro-markers in the NAC are first presented.

Table 13 below shows that those categories having the higher frequency rate in the use of macro-markers in the NAC are in order: **Topic-shifter** (4.6‰), **Organiser** (1.6‰), and **Rephraser** (1.2‰). The lowest rates are those of **Starter** (0.3‰) and **Conclusion** (0.2‰) with only thirty-nine and thirty-one occurrences respectively.

A more detailed analysis per category reveals that within **Topic-shifter**, there is one of the macro-markers that distinguishes itself from the rest; *so*, representing 3.5‰ out of the total with 439 occurrences. As explained in the previous section, *so* can function as several types of DM conveying different meanings. *So* as a macro-marker affects the overall discourse structural relations, in this case it normally occurs at the beginning of a section of talk, generally after a long pause, and its function is to introduce a new topic; in contrast to *so* as a micro-marker with a consecutive meaning. This function of *so* as a topic-shifter is similar to the notion of *so* as a ‘flag’ mentioned by Swales and Malczewski (2001), or the ‘global’ function of *so* that indicates relationships to a whole stretch of discourse pointed out by Schiffrin (1987). The macro-marker *so* not only stands out within the **topic-shifter** category, but in the whole macro-marker classification as the most recurrent one (see examples below).

SO (at the beginning of a section of speech)

(42) it is, of the resources needed, nutrients needed for that particular crop. *so*
 what will happen is that after a period of time, people will leave that plot
 (LE1/NAC)

- (43) day, but after that we'll sort a look at the parallel case, dealing with English.
so, the next thing that you should be reading in the coursepack are the series
(LE12/NAC)

The next topic-shifter is *now* with 127 occurrences and a rate of 1.0 ‰. It is also a polysemous DM that can function in different ways according to the meaning conveyed and the type of relations in the discourse. We have seen examples of *now* as an additional micro-marker, but as a macro-marker it serves to introduce a new topic and as such it generally occurs after a pause and at the beginning of a section of talk as the following examples show.

NOW

- (44) the economic relations that determine the nature of that society. <P :12> *now*,
what we're gonna do, today, this is the uh, lemme just get a sense (LE3/NAC)
- (45) cuz the page number. (in p-) i guess it's page four. SU-F: four S1: alright.
now, i am gonna write this again on the board (LE10/NAC)

The third macro-marker within the **topic-shifter** category is *actually* with quite few occurrences compared to *so* and *now*. When searching for *actually* we realized that this macro-marker does not always behave or function as a topic-shifter in the discourse, only instances of the adverb *actually* in initial clause position showed a shift in the topic. By contrast, middle or final clause positions clearly conveyed the reality of the situation not introducing a new topic (see examples below of *actually* as a topic-shifter).

ACTUALLY

(46) she could almost be like a typical American, girl at the time. S1: right. um *actually*, i think i've, was i was planning to (communicate) this message of some (LE9/NAC)

(47) are a lot of ornamentation going on? no. no, so very sort of streamlined, sparse_ *actually* if you look at it close up, Mies Van der Rohe the architect who coined (LE8/NAC)

The next category is **Organizer** with 1.6‰, not very close to the preceding category of **Topic-shifter** (4.6‰). Within **Organizer**, there are two macro-markers worth mentioning, these are *let's*, in the contracted form or *let us* in the full form with the first person plural object pronoun and *let me*, or the contracted form *lemme*, using the first person singular object pronoun. These macro-markers appeared in the NAC followed by verbs such as *go back*, *run through*, *focus*, *look*, etc. The frequency rate for *Let me* and *Let's* is very similar getting an 0.6‰ for *let's/ let us* and 0.5‰ for *let me/ lemme* with 75 and 65 occurrences respectively. We want to point out that instances of the relaxed and contracted form of *let* with the first person plural pronoun (*let's*) are more numerous (75 occurrences) than the full form *let us* (5 occurrences). By contrast, instances of the relaxed colloquial form *lemme* are less frequently used (20 occurrences) than the full form *let me* (45 occurrences) with the first person singular object pronoun. At this point we have to mention that those instances of *let's/ let us/ let me/ lemme* followed by the verb *begin* have been excluded from the organizer category and considered for the Starter category, we consider the DM *let us begin/ let me begin* or any of the

possible contracted forms as an starter and not as an organizer. Examples 48 and 49 below show *let's* and *lemme* as organisers.

LET'S

- (48) features. so this is a a point is a is a is deb- is much debated. but anyway *let's* look at the specific changes now here i'll have to give you some linguistics
(LE12/NAC)

LEMME

- (49) is a small country, in the world, the small country assumption... alright. well *lemme*, quickly *run through*, some uh comparative statics with this. let's uh...
(LE6/NAC)

It is the union of these two macro-marker occurrences (*let us/ let me*) what positions the **organizer** category as the second most frequently used category. There is a third DM in the **organizer** category, that is, *I wanna* (0.5‰) with the first person singular subject pronoun; no instances of *wanna* with the first person subject pronoun (*we*) were found. *I wanna* functions as an organizer DM when followed by mental and action verbs such as *discuss*, *do*, *emphasize*, *digress*, and usually accompanied by the time adverb *today* or any other temporal phrase. No instances of *I wanna* followed by action verbs such as *begin*, *start*, *finish up* or *end* have been considered as organizers, in fact these would be instances of starters or conclusion macro-markers, in any case, not recurrent enough to be considered in the classification. Note that the contracted colloquial form *I wanna* (40 occurrences) has a higher number of occurrences than the full form *I want to* (20 occurrences), although both are considered as a single organizer DM. In the

following examples we can observe the macro-marker *I wanna* as an organizer followed by a time phrase or adverb.

I WANNA

(50) in lecture and it will, uh enable me to, uh see how, things are going. um, what *i wanna do today*, as well, uh is, technology permitting, uh to have a look at
(LE2/NAC)

(51) *i wanna start* talking about the Static Neoclassical Model. but before i do that *i wanna digress* for a second and talk about aggregate production functions
(LE6/NAC)

The third category with a high frequency rate of macro-markers is **Rephraser** (1.2‰) very close to the previous organizer category (1.6‰). Within **Rephraser**, we observe that there is a relevant macro-marker with 96 occurrences (0.8‰), that is, *I mean*. This rephraser DM always occurs in the NAC with the first person singular subject pronoun and it is mainly used by the lecturer/ speaker in order to reword what he/ she has said immediately before, for the better understanding of the student/ hearer (see example below).

I MEAN

(52) culturalists, and shopkeepers? you can't say that, it's just for human survival, *i mean* if there was just one way of surviving, then we'd all be doing, pretty
(LE1/NAC)

With the same function as *I mean*, the second most frequently used rephraser macro-marker is *that is* with 58 occurrences (0.5‰). It appears in the discourse, as

expected, after a pause (normally short) and sometimes also before another pause, although this is not always so. Not so frequently used as *I mean* and *that is* is the rephraser *in other words*, contrary to what we would expect. Other instances of rephraser macro-markers were hard to find. Examples 53 and 54 below illustrate the rephraser macro-markers *that is* and *in other words* as they appear in the NAC.

THAT IS

- (53) her nonverbals you know how confident she seems. and they're blind to condition, *that is*, they don't know whether the woman speaking, saw the traditional ads or (LE4/NAC)

IN OTHER WORDS

- (54) that Basque did not have originally in its phonemic inventory the /f/ sound. *in other words* Basque lacked in its phonemic inventory, a voiceless labiodental (LE12/NAC)

Among macro-markers the categories that showed a lower frequency rate were **Starter** and **Conclusion** having very similar rates, in order, 0.31‰ (39 occurrences) and 0.2‰ (31 occurrences). The most recurrent macro-markers within the starter category are *first of all* or simply the adverb *first* following in number of occurrences, with expressions including action verbs such as *start*, *begin* and the verb *want* followed by to-clauses, the usual expressions found in the NAC were: *to begin with*, *we're gonna begin*, *let's begin*, *I want to/ wanna do today/ start with/ talk about*. Notice that, once more, relaxed colloquial forms of verbs such as *gonna* (*going to*) and *wanna* (*want to*) have been taken into

consideration as reflected by the transcripts. Also interesting to point out is the fact that there are no instances of any first person plural subject pronoun (*we*) accompanying the volition verb *want to/ wanna*, instead it only appears with the first person singular subject pronoun. Some examples of starter macro-markers are shown below, notice that in some cases and especially with expressions such as *I wanna start* or *we're going to begin* time adverbs can usually be found in the vicinity (see example 58 with *today*).

FIRST

(55) has made absolutely crucial contributions to analyses of popular culture. so, *first* i want to tell you who Karl Marx was and uh why those of us studying the (LE3/NAC)

(56) at the end he lost out to the emperor Augustus. now who was Augustus? *first of all*, he begins, with the name of Caesar and we'll look at a slide, uh (LE2/NAC)

BEGIN

(57) eks, is try to look at, different biological takes, on uh, on uh, behavior. and *we're gonna begin* by talking about evolution, and behavior, uh and that's what (LE5/NAC)

WANNA DO TODAY

(58) S1: okay, uh while we (xx) um recover the technology, um, what *i wanna do today*, uh is to continue on obviously through our rapid tour, of Roman (LE2/NAC)

Regarding the **conclusion** category with the least frequency rate (0.2‰) and only 31 occurrences, we find the macro-marker *finally* as the most frequently used in order to conclude and close the lecture discourse in the NAC. We have to point out that not all instances of *finally* in the NAC function as a conclusion macro-marker expressing that something is last in a series of actions or events or when introducing a final point in speech. The adverb *finally* also appears in the NAC with a resultative meaning expressing that something is done or happens after a long period of time, in this case interchangeable with *in the end* or *eventually*. These instances have not been considered as conclusion DMs. Example 59 below is an instance of *finally* with a resultative meaning and therefore not considered a conclusion DM in this study. The next example shows *finally* as a conclusion macro-marker in the NAC accompanied by the additional micro-marker *and*.

FINALLY

(59) problem is. and after doing this for about fifty or sixty years psychologists have *finally*, come to the conclusion that it doesn't work, and that people would be (LE11/NAC)

(60) don't have other meaning systems that would allow them to do that, McKay...? *and finally* the media themselves, according to a Marxist, relay interpretive framework (LE3/NAC)

The following conclusion macro-marker is formed by combinations such as *to end up/ with*, *to finish/ up* with very few occurrences (5 occurrences). As conclusion macro-markers were hard to be found, we decided to make closer reading of the NAC in order to find out how speakers ended their lectures. This reading revealed

that in some lectures the concluding remark came preceded by the speech act *I'll see you* followed by a day of the week or time adverb. (See examples below).

SEE YOU

(61) again comment on the politics of identity, but also do so particularly on, the role and the representation of women. so have an excellent weekend and *i'll see you* on Monday. {END OF TRANSCRIPT} (LE8/NAC)

(62) it's warm anyway, and uh so let's give the tape recorder a break too, and so thank you very much and *i'll see you* on Thursday. (LE12/NAC)

b) Macro-marker (*Overall discourse structural relations*)*NAC Results*

Starter	# DMs	‰	Rephraser	# DMs	‰	Organizer	# DMs	‰	Topic-shifter	# DMs	‰	Conclusion	# DMs	‰
<i>first (of all)</i>	20	0.15	<i>I mean</i>	96	0.8	<i>let's(let us try, go back/through focus, look</i>	80	0.6	<i>so</i>	439	3.5	<i>finally</i>	21	0.2
<i>to begin (with),we're gonna begin, let's begin</i>	10	0.07	<i>in other words</i>	58	0.5	<i>let me (lemme) go back/through focus, look</i>	65	0.5	<i>now</i>	127	1.0	<i>to end up/with, to finish/up</i>	5	0.03
<i>I want to/wanna do today/start with/talk about</i>	9	0.07	<i>that is</i>	4	0.03	<i>I wanna/want to discuss, do, emphasize...</i>	60	0.5	<i>actuallyt</i>	13	0.1	<i>I'll see you</i>	5	0.03
TOTAL 1	39	0.31		158	1.2		200	1.6		579	4.6		31	0.2

TOTAL 2	# DMs	‰
Macro-markers	<i>1,007</i>	8.0

Table 13. Macro-markers results in number of occurrences and frequency rate in the NAC.

6.4.2. Macro-markers in SC

Still in the second part of the analysis, we present now the results obtained from the study of macro-markers in the SC (Spanish Corpus).

Taking a general look at the results obtained in the SC, we observe that, on the whole, there are 577 macro-marker occurrences in the overall SC; this represents a frequency rate of 7.2‰ out of the total Spanish lectures.

As can be seen in Table 14 below, the results in the search of macro-markers in the SC revealed three most important categories according to frequency rate, they are: **Rephraser** (3.5‰), **Organizer** (1.9‰) and **Topic-shifter** (0.8‰), closely followed by Starter with the same frequency rate (0.8‰), but with a lower number of occurrences (62 occurrences).

The **Rephraser** category has the highest frequency rate due to the marker *es decir*, which comes to be the most recurrent macro-marker in the SC. The macro-marker *es decir* is surely relevant within the **Rephraser** category followed by *o sea* but with a lower number of occurrences (56) compared to 212 occurrences for *es decir*. There is a third rephraser macro-marker that happened to be used across the SC and which had no counterpart in the observed NAC, this is *insisto*. Spanish lectures seem to use *insisto* (possible translation: *I insist*, not used as a **rephraser** in the NAC) in order to paraphrase and reword what they have just said or explained, the frequency rate for *insisto* is 0.2‰ with 14 occurrences, a lower rate than *es decir* or *o sea* but still relevant. Some examples of these rephrasers from the SC are given below.

ES DECIR

- (63) para hablar y reflexionar [repetición] sobre el propio sistema de comunicación [pausa]. *Es decir*, dicho de otra manera, fijaros que las... las [repetición] primeras (LE10/SC)

O SEA

- (64) foros no es bien recibido que haya publicidad, que haya publicity encubierta, *o sea*, que se hable de marcas... o que descaradamente se esté utilizando el foro (LE5/SC)

INSISTO

- (65) escalonado de Deir el Bahari que es el templo funerario de la reina Hatshepsut, *insisto* es uno de los más hermosos, creo que lo visteis el otro día en un reportaje (LE2/SC)

The second most recurrent category is **Organizer** with 1.9‰ (156 occurrences). We have observed that within the **organizer** category, markers do not tend to be fixed linguistic items but a combination of words or expressions that serve to structure the ongoing discourse, usually formed by perception and/ or action verbs. Thus, expressions such as *(ahora) vamos a ver/ hacer, (hoy)(el próximo día) veremos* are the most frequently used in the **organizer** category with 113 occurrences (1.4‰). Notice that these macro-markers normally co-occur with time adverbs such as *ahora, hoy, el próximo día*, etc., placing the hearer in space (organizing the discourse sequences) and time. I have gathered some of these macro-markers in the examples below.

VAMOS A VER

- (66) en lugar de pasar ahora ya al tema 7.3, 7.4 que es lo que llevábamos en marcha *vamos a ver* hoy el tema 8, y ya mañana retomaremos el 7.2.2 ¿vale?
un grupo (LE11/SC)

VEREMOS

- (67) que son las conciliaciones de las cuentas corrientes bancarias, que también *veremos luego más adelante*, y que tienen como objetivo, el comprobar si,
entre (LE4/SC)

A quite literal organizer is *por un lado/ por otro (lado)* when organizing ideas or events in the discourse, this organizer macro-marker has 34 occurrences (0.4%) being the second most recurrent organizer in the SC, although with a lower number of occurrences compared with the previous macro-marker. Other **organizers** that come in the third place make reference to the resources and materials used to deliver the lecture, such as handouts or visual aids (OHPTs, PPTs Slides, etc.). These macro-markers (0.2% frequency rate) are: *en la siguiente diapo/ tema*, etc. Some examples of organiser macro-markers can be read below.

POR UN LADO

- (68) En realidad con dos objetivos, *por un lado* ver hasta que punto eh... los sistemas de comunicación de los otros primates son diferentes o son parecidos al sistema de comunicación de los humanos (LE10/SC)

EN LA SIGUIENTE DIAPOSITIVA

- (69) eso está en la diapositiva veinticinco de forma resumida a qué se dedican, *en la siguiente diapositiva* os he puesto otras dos [eleva el tono] eh... (LE12/SC)

Topic-shifter and **Starter** are the following categories in the ranking, both at the same level, since they share the same frequency rate (0.8‰), although **Starter** has a higher number of occurrences (62) compared to the 66 occurrences of the **topic-shifter** category. Again, within each of these two categories, there is a macro-marker that seems to be relevant for the final score. Within **Topic-shifter**, the most frequently used marker is *ahora*, usually co-occurring with the additional micro-marker *and* as well as with other markers such as *bien* or *bueno*. *Ahora* as a macro-marker is used to introduce a new topic or change it; those instances of *ahora* as a temporal adverb have not been regarded as a DM. In example 70 below *ahora* is functioning as a topic-shifter macro-marker whereas in the next example, (71) *ahora* functions as an adverb with temporal meaning.

AHORA

- (70) ¿de acuerdo? he hecho así como un panorama muy rápido para situarnos, pero, *ahora* vamos a por ello ¿qué es la pragmática? ¿m? en principio, para que veamos (LE11/SC)
- (71) 130 respecto al dólar [INTERVENCIÓN ESTUDIANTE] (alumno responde) las exportaciones *ahora* son más caras [ralentiza el discurso] por lo tanto vamos a exportar menos (LE7/SC)

The other two macro-markers within the **topic-shifter** category that have proven to be relevant among the SC are *en realidad* (0.2‰) with 14 occurrences and *de hecho* with a rate of 0.1‰ and 9 occurrences.

With the same frequency rate (0.8‰) but with fewer occurrences than **Topic-shifter**, the **starter** category is the fourth category in the ranking only followed by **Conclusion**, which is the least recurrent category with a total number of 11 occurrences which corresponds to 0.13‰ out of the total macro-markers in the SC (7.2‰).

Within **Starter** there is a macro-marker that stands out, expressions such as *en primer lugar*, *el/ lo primero*, *veamos primero* come first place 56 occurrences and a rate of 0.7‰ out of the total 0.8‰ for the **starter** category. Other instances of starter macro-markers were hard to find in the SC and only a few instances of *vamos a comenzar/ empezar* or *sin más preámbulo* have been found (see Table 14). In the examples below the most important starter macro-markers are shown:

EN PRIMER LUGAR

(72) de marketing relacional? (tono descendente) utilizando para ello internet.
 Bien *en primer lugar* deberíamos recoger información, habíamos dicho antes que los elementos... (LE5/NAC)

PRIMERO

(73) [P. LEE DE TRANSPARENCIA]. ¿Qué pasa?, ¿cuáles son las consecuencias? ¿eh? *primero* vamos a entrar en las expresiones de desencanto y luego vamos a ver cual (LE9/SC)

The category that had the lowest frequency rate of DMs in the SC was **Conclusion**, just an in the NAC. **Conclusion** represents only a 0.13‰ which corresponds with just 11 occurrences from the total SC. Within the **conclusion** category the three most recurrent macro-markers were *lo dejamos aqui, (y) finalmente, con esto terminamos.*

b) Macro-markers

SC results

Starter	#	%	Rephraser	#	%	Organiser	#	%	Topic-shifter	#	%	Conclusion	#	%
	DMs			DMs			DMs			DMs			DMs	
<i>en primer lugar, e/lo primero, veamos primero</i>	56	0.7	<i>es decir</i>	212	2.6	<i>(ahora)vamos a ver/hacer (hoy)(el próximo día) veremos</i>	113	1.4	<i>(y) ahora</i>	43	0.5	<i>lo dejamos aquí</i>	6	0.07
<i>vamos a comenzar/ empezar</i>	5	0.06	<i>o sea</i>	56	0.7	<i>por un lado/por otro</i>	32	0.4	<i>en realidad</i>	14	0.2	<i>(y) finalmente</i>	3	0.03
<i>sin más preámbulo</i>	1	0.01	<i>insisto</i>	14	0.2	<i>en la/el siguiente diapo/tema...</i>	11	0.13	<i>de hecho</i>	9	0.1	<i>(con esto/eso) terminamos</i>	2	0.02
TOTAL 1	62	0.8		282	3.5		156	1.9		66	0.8		11	0.13

TOTAL 2	# DMs	%
Macro-markers	577	7.2

Table 14. Macro-markers results in number of occurrences and frequency rate in SC.

6.4.3. Comparison of macro-markers in the NAC and SC

An overall view of the results can be seen in total 2 (see Tables 14 and 15 above). We observe that the total frequency rate of macro-markers when contrasting the SC and NAC shows similarities in the three most recurrent macro-marker categories although in a different order of frequency (**Topic-shifter**, **Organizer**, **Rephraser**) along with the least frequently used one (**Conclusion**). However, differences between the NAC and the SC can be seen as to the total amount of macro-markers used, being the total frequency rate 8.0‰ (1,007 occurrences) in the NAC, slightly higher than the frequency rate in the SC with 7.2‰ (577 occurrences).

After an individual analysis of both NAC and SC, we now present a comparison of the results obtained in both sub-corpora. The analysis of the macro-markers in the NAC has revealed three most frequent categories, namely, **Topic-shifter**, **Organizer** and **Rephraser**. Surprisingly, the three most recurrent categories in the SC coincide with the NAC categories although in a different frequency order. In the SC the three most outstanding categories are: **Rephraser** (3.5‰), **Organiser** (1.9‰) and **Topic-shifter** (0.8‰). A closer look at the frequency rates in the NAC macro-markers results has revealed that the **topic-shifter** category has a quite higher rate (4.6‰) compared to the SC due to the large number of occurrences of the macro-marker *so* (439 occurrences). A middle position in both sub-corpora has the **organiser** category with a quite similar rate in both sub-corpora; 1.6‰ for the NAC and slightly higher, 1.9‰ for the SC. The first category in the ranking in the SC, that is, **Rephraser** comes to have a third place in the NAC with a 1.2‰ frequency rate compared to the 3.5‰ in the SC.

More similarities between the NAC and the SC in the use of macro-markers are found in. In both sub-corpora, **Conclusion** is the least used category with close frequency rates, that is, 0.2‰ (31 occurrences) for the NAC and 0.13‰ for the SC (11 occurrences), in both sub-corpora instances of conclusion macro-markers were hard to find.

Taking into account the use of individual macro-markers, I observe that in the NAC the three most recurrent macro-markers are: *so* (439 occurrences), *I mean* (96 occurrences) and *let's/ let us try, go back, etc.* (80 occurrences), whereas in the SC we find *es decir* (212 occurrences) (*ahora*) *vamos a ver/ hacer...*, (113 occurrences) and *en primer lugar, el/ lo primero, etc.* (56 occurrences) having *I mean/ es decir* (rephraser category) and *let's/ let us try, go back/ (ahora) vamos a ver/ hacer* (organizer category) also in the NAC. It is worth pointing out that the number of macro-marker occurrences in the NAC is twice the number in the SC, mainly due to the extended use of some English macro-markers such as *so* or *now* in the **topic-shifter** category which have fewer occurrences in the SC, especially the Spanish macro-marker *ahora* as a topic-shifter.

6.5. Operators

6.5.1. Operators in the NAC

The third stage of the analysis is the application of the DMs I have called here Operators and which convey relations between speaker-speech and/ or speaker-hearer.

On the whole, as we can observe from Total 2 in Table 15, there is a total of 1,066 occurrences of operators in the NAC which gives a rate of 8.4‰ out of the whole corpus.

Under those operators which signal speaker-speech relations we have considered categories such as **Attitudinal** and **Pause filler**; moreover, those categories which convey speaker-hearer relations are **Elicitation**, **Acceptance** and **Confirmation-Check**. As I proceeded before with the micro and macro-markers, the English DM classification model was first applied to the NAC. The results of the search for Operators in the NAC can be seen in Table 15 at the end of this section.

The categories within operators showing a higher frequency rate are **Pause-filler** with 3.4‰ (436 occurrences) and conveying relations speaker-speech, followed by **Confirmation-check** with 2.1‰ (273 occurrences) and **Acceptance** with a slightly lower rate (1.6‰) and 207 occurrences, these two last categories conveying relations speaker-hearer. The categories within operators having the lowest frequency rate in the NAC are **Attitudinal** with 1.03‰ (136 occurrences) and **Elicitation** with the lowest rate (0.1‰) and very few occurrences (14 occurrences).

The operator *par excellence* in the NAC seems to be *okay* with a total of 415 occurrences only followed by *and* with 199 occurrences. The operator *okay* is quite broad since it can function in many different ways as an operator (functional categories illustrated in Table 15 below). It can be used as a response to questions when prompting or agreeing with the main speaker, what other authors have called *back-channel* instances of *okay* (Swales and Malczewski, 2001) and what we have categorized here as **Acceptance**; *okay* can also be used as a kind of question tag, that is, *okay?*, in which case it is a **confirmation-check** (both cases marking speaker-hearer relations) and shows a higher number of occurrences than *okay* as **Acceptance** or **Pause-filler**. However, and contrary to what could be expected from a direct question, most instances of *okay* as a **confirmation-check** in the NAC are not intended to elicit a direct verbal response. The third function of *okay* is as a pause-filler operator to maintain the floor in which case it is hard to interpret, barely holding a semantic meaning. In the following example we can see *okay* as a pause-filler operator, in this case co-occurring with another pause-filler operator, *well* (underlined). We have also observed that *okay* normally co-occurs with other DMs such as *so*, *now* or *and* with different meanings and functions:

OKAY (Pause-filler)

(74) of the total variance in your outcome, that lies systematically between groups.

okay. well, there's the variance between groups, and here's the total variance

(LE10/NAC)

Still within the **pause-filler** category we find, in order, *well* and *and* with 121 (1.0%) and 116 (0.9%) occurrences respectively. In order to obtain the desired results in the case of *well*, I had to take away expressions such as *as well*, *very*

well, *well done*, *well-known*, *quite well* and indirect speech segments such as *say well* or *said well*. Instances of *well* as a pause-filler usually came after pauses or what I call here hesitators *um*, *uh*. See example 75 for operator *well*.

WELL

(75) somebody gets steamed up about it. alright. i hope this has convinced you,
well... i don't know whether convinced is the word uh it slightly, has has has
(LE10/NAC)

However, the most numerous pause-filler in the NAC is *and* with 199 occurrences (1.5%). I already discussed *and* in section 6.3.1 of this chapter where *and* also functioned as a micro-marker. *And* as a pause-filler usually lacks semantic meaning and it is used as a hesitator giving the speaker time to think about what comes next or re-take the lecture discourse (see next example). In the NAC it is usually followed by hesitators *um* or *uh*.

AND

(76) you're eligible, we'll call you, um, and we'll set up a time for you to come in
and uh give you the (xx) so, it's it's a pretty simple process you come in
(LE3/NAC)

The following most recurrent category in number of occurrences (273) and a frequency rate of 2.1% is **Confirmation-check**. Here we again find the DM *okay*, this time used as a kind of question tag without the function of eliciting information or inviting someone to take the floor. *Okay?* has the highest number

of occurrences compared to the other categories of *okay* as an operator (**Acceptance** and **Pause-filler**).

OKAY?

(77) amount of it back to, th- to um, to somebody else. so they had to give tribute.

okay? another example would be chiefdoms, where you have that, where people give (LE1/NAC)

Right? (illustrated in example 78) is the second most recurrent **confirmation-check** with 73 occurrences (0.6‰), functioning in the same way as the operator *okay?*. It seems to be less often used than the previous one followed by the compound operator *alright* with 34 occurrences (0.2‰). The preferred **confirmation-check** in the NAC seems to be the operator *okay?*.

RIGHT?

(78) socialist at all. it has a few, tiny, elements, of sort of socialist practice, *right?*

like Medicare, Medicaid, food stamps, those sorts of things. but in this but in this country, we believe in the sanctity of the market and private enterprise (LE3/NAC)

The next operator category in number of occurrences is **Acceptance** (1.4‰) with three most recurrent operators in the NAC, namely, *okay*, *right* and *alright*; the same DMs I found in the previous category but with different functions and meanings. We have already mentioned the case of *okay* as being very broad. *Okay* as an acceptance operator appears in a large number of occurrences (133), more often than the operator *okay* as a pause-filler. *Okay* within the acceptance category

is used when prompting or agreeing with the main speaker whether lecturer/speaker or student/ hearer as the following example shows.

OKAY

(79) show you how to edit output, so that it will save you pages. so that you [S1: *okay*] don't use up your printing. we can, we can show you how to do that, using, Microsoft Word. (LE10/NAC)

Alright is the next operator in the Acceptance category with 38 occurrences (0.3‰) closely followed by the isolated form *right*, lower in number of occurrences (36) with a frequency rate of 0.3‰. Although *alright* and *right* can be used as acceptance operators in the same way as *okay*, they are not as widely used as *okay* in the NAC. In the case of *right* I excluded expressions which contained *right* such as *the right*, *the right time*, *right here* or *right there* among others; moreover instances of *right* as an acceptance operator were followed by a pause (whether long or short) in the NAC (see the following examples).

RIGHT

(80) be, transcribed like written down? or [S2: mhm] or it's gonna be played? S2: *right*. it's, play- well we hope to have it both available in the the form of a (LE12/NAC)

ALRIGHT

(81) is that? cuz the page number. (in p-) i guess it's page four. SU-F: four S1: *alright*. now, i am gonna write this again on the board (LE10/NAC)

The categories showing the lowest frequency rate in order were **Attitudinal** and **Elicitation**. **Attitudinal** showed 136 occurrences, that is, a rate of 1.04‰ with three most recurrent operators, namely, *I think/ we think, as you know, I believe/ we believe*. I have observed that the most frequently used attitudinal operator is *I/ we think* with 119 occurrences (0.9‰); instances of *we think* with the first person plural pronoun correspond only to 8 occurrences, the rest, 111 occurrences for the operator *I think* with the first person singular pronoun. *I think/ we think* is no doubt key in the Attitudinal category because of its recurrence since it is distantly followed by *as you know* with a rate of 0.1‰ and 12 occurrences. The last attitudinal operator is *I believe/ we believe* with only 5 occurrences, three occurrences for *I believe* and two for *we believe*. Apparently, in the NAC the attitudinal verb *think* is preferred over the stronger meaning of attitudinal *believe* (see some examples of attitudinal operators below).

I THINK

(82) and she peeks out from the top of it. Goffman's example, is also a cigarette ad. *i think* this woman looks like um... is it Malcolm McDowell who's in uh, A Clockwork Orange? you ever see this movie? he's got these false eyelashes that go like that. (LE4/NAC)

I BELIEVE

(83) people's expectations. okay? the story line that they're going to want you to accept *i believe*, um which *i think* is probably okay in this application and maybe not as okay in other applications (LE7/NAC)

AS YOU KNOW

- (84) you know Eastern European you know it wouldn't have had nearly the same impact **as you know**, pretty much, i mean sh- sh- could almost be like a typical American (LE9/NAC)

Elicitation is the last category within operators according to frequency rate (0.1‰) and number of occurrences, only 14. The three most recurrent operators within this category are in order of frequency: *any questions?*, *why is that?*, *anyone?* By means of these expressions the speaker tries to elicit information from the audience and invites the audience to take the floor, most of the time this does not happen because the speaker does not succeed in his/ her elicitation. Some examples of elicitation operators in the NAC are shown here.

ANY QUESTIONS

- (85) alright and that's the point that i actually want to try to get across. alright? ***any questions*** on that particular example? S4: you said the internal was the upper lower class? (LE12/NAC)

ANYONE...?

- (86) a little jaunt a little foray, into architecture. so on the left you have, ***does anyone recognize*** that building? any of you New York City, folks? S2: Seagram (LE8/NAC)

c) Operators

NAC results

Relation speaker-speech

Relation speaker-hearer

Attitudinal	# DMs	%o	Pause filler	# DMs	%o	Elicitation	# DMs	%o	Acceptance	# DMs	%o	Confirmation -check	# DMs	%o
<i>I think/we think</i>	119	0.9	<i>and</i>	199	1.5	<i>any questions(?)</i>	8	0.06	<i>okay</i>	133	1.1	<i>okay?</i>	166	1.3
<i>as you know</i>	12	0.1	<i>well</i>	121	1.0	<i>why is that?</i>	3	0.02	<i>alright</i>	38	0.3	<i>right?</i>	73	0.6
<i>I believe/we believe</i>	5	0.04	<i>okay</i>	116	0.9	<i>anyone?</i>	3	0.02	<i>right</i>	36	0.2	<i>alright?</i>	34	0.2
TOTAL 1	136	1.04		436	3.4		14	0.1		207	1.6		273	2.1

TOTAL 2	# DMs	%o
Operators	1,066	8.4

Table 15. Operator results in number of occurrences and frequency rate in NAC.

6.5.2. Operators in the SC

The findings from the analysis of operators in the SC can be seen in Table 16 at the end of this section.

An overall view of results in the SC can be seen in total 2, where the total number of operator occurrences is 1,039, which corresponds to a rate of 13%, slightly higher than in the NAC.

The results have revealed three most important categories, namely, **Confirmation-check** (5.7%), **Pause-filler** (4.8%) and **Acceptance** (1.06%), quite similar to the results obtained in the NAC. On the other hand, the two categories showing the lowest frequency rates are **Attitudinal** (1%) and **Elicitation** (0.4%), similar in order with the results for operators in the NAC (see Table 15 for results in the NAC).

Confirmation-check is the most recurrent category in the SC conveying relations between speaker-hearer with 457 occurrences (5.7%) mainly due to the frequently used confirmation-check operator *¿vale?* with 337 occurrences and being the most often used operator in the whole SC. *¿Vale?* as a confirmation check does not try to elicit information from the hearer but simply checks that what has been previously said has been understood or that the audience is following the ongoing discourse (see example below).

¿VALE?

(87) la sociedad y la lingüística en estudiar el lenguaje, dentro de esa sociedad
¿vale? la utilización del lenguaje dentro de esa sociedad, la pragmática lo que
hace es, en cierto modo (LE11/SC)

¿Vale? is distantly followed by the confirmation-check *¿de acuerdo?* with 84 occurrences and a rate of 1.0%. The last confirmation-check is *¿queda claro?* with very few occurrences, only 6, representing a low rate of 0.06%. The fact is that except for *¿vale?* and *¿de acuerdo?* other types of confirmation-check operators were hard to find. Examples of *¿de acuerdo?* and *¿queda claro?* from the SC are provided below. In example 94 two confirmation-checks come together *¿vale?* (underlined) and *¿de acuerdo?* reinforcing and strengthening the confirmation-check function. In the next example (ex. 95) we observe that *¿queda claro?* co-occurs with another operator functioning as an acceptance marker, this is the case of *vale* included within the acceptance category in the present operators taxonomy (see Table 16).

¿DE ACUERDO?

- (94) [pausa esperando respuesta] generalmente una línea con una gran pendiente
¿vale? *¿de acuerdo?* Y ahora pon supuesto, quiero decir que ahí te puedes encontrar con (LE3/SC)

¿QUEDA CLARO?

- (95) por lo tanto la media de puntuaciones diferenciales [P. ESCRIBE EN LA PIZARRA] *¿queda claro?* *Vale* ¿cuál será la desviación típica de esta variable? [ralentiza el discurso] (LE3/SC)

The second category in number of occurrences is **Pause-filler** with 377 occurrences (4.8%) distributed among the operators *pues*, *bueno* and *bien* and conveying relations between the speaker-speech, in contrast to the first category in the ranking, which conveyed relations between the speaker-hearer. It is worth

pointing out here that the three most recurrent pause-fillers have shown quite similar results with nearly the same frequency rate each (0.2%). Slightly higher in number of occurrences is the operator *pues*. It can have more than one function as a DM, as it could also be a consecutive micro-marker; although more relevant and recurrent as a pause-filler with only slight or no semantic meaning, which is why it has been placed in the operators classification. An example of *pues* as a pause-filler operator in the SC is provided here where we find an overuse of the operator *pues* (up to four times) giving the speaker time to think what he/ she is going to say.

PUES

- (96) en que para esta nueva serpiente pues el grito va a ser pues uh [eleva el tono]
pues no me gusta *pues* ah [eleva el tono] *pues* vale *pues* ah. Muchos
lingüistas dirían no eso no me lo creo (LE10/SC)

The other two pause-filler operators share the same number of occurrences (225) and rate (0.2%) and are used by the speaker/ lecturer to fill pauses or go on with the discourse. *Bueno* and *bien* would be equivalent to the English *well* and can be used indistinctly as a pause-filler; the similar results show no preferences. We have observed that within the pause-filler category these three operators seem to co-occur together in many contexts. (See examples below for *bueno* and *bien*).

BUENO

- (97) una pregunta de examen. De momento llevamos tres ¿no? Vale, nada, [baja el
tono] *bueno* experiencias de enseñanza del lenguaje verbal a primates,
(LE10/SC)

BIEN

- (98) puntuaciones diferenciales es igual a la varianza de las ys en puntuaciones directas, *bien* ¿a dónde voy a parar? [tono descendente] si nosotros en lugar de trabajar (LE3/SC)

The third category in the ranking is **Acceptance** (95 occurrences) with fewer occurrences compared to the previous pause-filler category but enough to be in the third place with a frequency rate of 1.06%. The most important operator within this category is *vale*, whose nature has already been commented on this section when functioning as a confirmation-check (¿*vale*?). *Vale* is here used to accept and agree with the main speaker but without intending to elicit information from the hearer. These instances of *vale* are not as numerous as ¿*vale*? as a confirmation-check but still relevant for the acceptance category because of the large number of occurrences compared to the other two recurrent acceptance operators: *desde luego* and *muy bien*. In the examples below instances of acceptance operators as they appear in the SC are given. Example 99 shows acceptance *vale* when the lecturer validates what the student has just said in her participation.

VALE

- (99) [pausa, esperando respuesta] grande ¿y el denominador? [INTERVENCIÓN ESTUDIANTE] (alumna responde) *vale*. Suponte tú que un numerador grande 100, lo dividimos entre una cantidad igual (LE3/SC)

MUY BIEN

(100) por lo polémico que es pues tiene publicidad gratuita [PROFESOR] aha
(asiente) *muy bien* eso es publicity de alguna forma sí. publicidad en los
medios de forma (LE5/SC)

DESDE LUEGO

(101) ¿qué ocurrirá si sale bien? [INTERVENCIÓN ESTUDIANTE] (alumno
responde) Efectivamente, *desde luego* se produce una retroalimentación, si se
consigue el éxito en esas campañas militares (LE1/SC)

The two least often used categories are **Attitudinal** and **Elicitation: Attitudinal** with 81 occurrences (1.0%) and operators such as *(yo) creo, parece que, fijaros*. *(Yo) creo* is the most recurrent one within **Attitudinal**, as expected according to Spanish subject pronoun grammar, only ten of the 39 occurrences of *(yo) creo* belong to *yo creo* where the first person subject pronoun is made explicit and not omitted. *(Me) parece que*, indicating hesitation or doubt about the accuracy of what is being said, has given 27 occurrences closely followed by *fijaros* with 25 instances. It is worth mentioning that no instances of *(yo) pienso* have been found in the SC, instead *(yo) creo* is the form used as equivalent to the English *I think*, a way to make a statement subjective or indicate some hesitation. *Fijaros* is a singular unusual form of an attitudinal operator, when using *fijaros*, the speaker/lecturer calls the attention of the hearer towards a topic. This marker is a popular form of *fijaos*, the standard form of the imperative plural for *vosotros* (see examples below).

CREO

- (102) [eleva el tono] con toda la denominación oficial excepto insisto todo potente,
creo por razones de estética seguramente [P. ACLARA LA VOZ] yo diría
 (LE1/SC)

ME PARECE QUE

- (103) los verdes en... en [repetición] una zona de... se les puso en una zona de... *me parece que* era de California o de Texas y... de repente se... digamos que en
 ese (LE10/SC)

FIJAROS

- (104) que se desarrolló Italia en España en el siglo dieciséis pero es un libro
 interesantísimo *fijaros* que el título es muy elocuente medidas del romano, el
 romano es el arte (LE2/SC)

The last operator category is **Elicitation** with 94 occurrences and a rate of 0.4%. **Elicitation** is the least used category among operators which conveys relation speaker/ hearer in the SC, as well as the least used in the NAC (see previous section 6.5.1). Within this category the three most recurrent operators in the SC are *¿alguna duda?*, *¿alguna pregunta?*, *¿recordáis?*, where the question with the word *duda* appears 12 times against 9 instances for *¿alguna pregunta?*; closely followed by *¿recordáis?* with 8 occurrences. In the case of *¿recordáis...?* it can appear followed by a Direct Object (DO), DO-clause or in isolation as a direct question. In example 105 below the lecturer is trying to elicit information from the students, waiting for an answer for a few seconds; however, no answer is given. These three elicitation markers try to prompt the hearer, although a hearer's

intervention does not always happen. Other elicitation markers were rare, being these three mentioned here the most frequently used ones. Examples of these elicitation operators as they occur in the SC are presented below.

¿RECORDÁIS?

(105) pueblos del desierto sirio-arábigo se producen cambios importantes en Mesopotamia *¿recordáis cuales?* [pausa esperando repuesta] *¿recordáis cuales fueron estos cambios?* [repetición, parafrasea] [pausa esperando repuesta] *¿quiénes fueron los pueblos que penetraron en Mesopotamia desde el desierto sirio-arábigo?* [repetición, parafrasea] [pausa esperando repuesta] Acabo de decir que hay dos corrientes... (LE1/SC)

¿ALGUNA PREGUNTA?

(106) no ha quedado claro el planteamiento de lo que hemos hecho en la clase anterior *¿alguna pregunta?* [INTERVENCIÓN ESTUDIANTE] (alumna pregunta) el sumatorio *¿de las?* [INTERVENCIÓN ESTUDIANTE] (alumna responde) (LE3/SC)

¿ALGUNA DUDA?

(107) el esperado son cálculos en función de la evolución de las variables macroeconómicas. *¿Tenéis alguna duda?* [pausa esperando pregunta o confirmación] Hacemos lo mismo que antes (LE7/SC)

c) Operators

SC results

Relation speaker-speech

Relation speaker-hearer

Attitudinal	# DMs	‰	Pause filler	# DMs	‰	Elicitation	# DMs	‰	Acceptance	# DMs	‰	Confirmation -check	# DMs	‰
<i>(yo) creo</i>	29	0.4	<i>pues</i>	127	1.6	<i>¿alguna duda?</i>	12	0.15	<i>vale</i>	70	0.9	<i>¿vale?</i>	367	4.5
<i>parece que...</i>	27	0.3	<i>bueno</i>	125	1.6	<i>¿alguna pregunta?</i>	9	0.11	<i>desde luego</i>	10	0.1	<i>¿de acuerdo?</i>	84	1.0
<i>fijaros...</i>	25	0.3	<i>bien</i>	125	1.6	<i>¿recordáis...?</i>	8	0.1	<i>muy bien</i>	5	0.06	<i>¿queda claro?</i>	6	0.07
TOTAL 1	81	1.0		377	4.8		29	0.4		95	1.06		457	5.7

TOTAL 2	# DMs	‰
Operators	1,039	13.0

Table 16. Operator results in number of occurrences and frequency rate in SC.

6.5.3. Comparison of operators in the NAC and SC

After analyzing both sub-corpora individually (NAC and SC), I now present a brief description of those relevant similarities and/ or differences that have resulted from comparing the NAC and SC.

In total (see total 2 in Tables 19 and 20) the results in the use of operators show that operators are slightly more recurrent in the SC with a frequency rate of 13.0‰ (1,039 occurrences) against the 8.4‰ (1,066 occurrences) rate in the NAC. Even so, according to the findings, the use of operators in the NAC and SC could be seen as similar.

As mentioned above the analysis of operators in the NAC has revealed **Pause-filler** (3.5‰), **Confirmation-check** (2.7‰) and **Acceptance** (1.4‰) in this order as the most important categories. Surprisingly, the same categories are the most frequently used in the SC with a slightly different order, that is, **Confirmation-check** (5.7‰) comes first in the SC followed by **Pause-filler** (4.8‰) and **Acceptance** (1.06‰). The total rate (see total 2 in Tables 19 and 20) for the use of operators is higher in the SC. Regarding the two different types of relations operators can convey we find that only one of the two categories conveying relation speaker-speech seems to be relevant in both sub-corpora, that is, **Pause-filler** with a high number of occurrences.

Another similarity between the NAC and SC is also observed with the least used categories among operators, in both sub-corpora the categories showing the lowest frequency rate are, in the same order, **Attitudinal** and **Elicitation**. The figures give a rate of 1.03 ‰ for the **attitudinal** category in the NAC and a close 1‰ in the SC. **Elicitation** is the least used category in both sub-corpora, slightly

lower than the results in the SC with 0.4‰ (29 occurrences) is the rate in the NAC with 0.1‰ (14 occurrences).

If we take a general view of the results obtained in the use of operators in both sub-corpora we can say the NAC and the SC operators in quite similarly ways in the lecture discourse. The most recurrent operator in the NAC is *okay* as it can have different functions as an operator, either as a pause-filler, an acceptance operator or a confirmation-check. Curiously, the most frequently used operator in the SC is its Spanish counterpart *vale*; however, in the Spanish operators classification *vale* can function as an acceptance operator or confirmation-check (*¿vale?*), but it is not recurrent enough in the SC to be included in the Spanish classification model as a pause-filler, in fact there are other more recurrent operators that function as pause-fillers in the SC (see Table 16).

6.6. DMs collocations

6.6.1. Introduction

As already described in the introduction, a fourth section dealing with the collocates of DMs closes this chapter. I have observed when analyzing and searching for results that some specific DMs seem to collocate in a quite high number of occurrences with others. The idea of DM collocations has already been introduced by authors such as Fraser (1990, 1999, 2004), Fraser and Malamud-Makowski (1996), and Swales and Malczewski (2001) who pointed out the idea

of looking at DMs as units that can co-occur together and the reasons why this might occur.

The aim here is to present the most recurrent DM collocations found in both sub-corpora, the NAC and the SC. Obviously and according to our classification not all DMs (micro-markers, macro-markers, operators) collocate all the time or the collocations may not appear in a large enough number of occurrences to be relevant for the study. First, the micro-marker co-occurrences are presented in the NAC and SC. I follow with macro-marker collocates and finally operator co-occurrences, first in the NAC and then in the SC to keep a parallel structure with previous sections.

6.6.2. Micro-marker collocations in the NAC

The first thing I observed when analyzing micro-markers in the NAC was the high number of occurrences of the DM *and*. Obviously, not all occurrences of this DM functioned as an additional micro-marker. A closer look at micro-marker *and* revealed that in most cases it appeared together with other DMs, which curiously enough were also included in other categories of the micro-marker classification. Thus, *and* as a micro-marker frequently co-occurred with *then* or *so*, also included in the micro-marker classification model under categories such as **Temporal** and **Consecutive**. One of the questions that came up after observing micro-markers collocations was whether they convey different meanings when in isolation or accompanied by other markers, as it is the case of *and*. I observed whether *and* kept the additional meaning when collocating with other micro-markers (*then, so*)

or adopted the co-occurring macro-marker meaning and function or vice versa, the co-occurring macro-marker inherited *and* additional meaning.

In Table 17 below the number of co-occurrences of *and* with *then* and *so* are shown, as well as their frequency rate. *Then* as a micro-marker is quite relevant as a **temporal** or **consecutive** marker, however, when co-occurring with additional *and* this function may vary, I will discuss this in the next chapter of this dissertation. *So* as a micro-marker has a main consecutive function in the NAC; however, it also tends to co-occur with *and* in many cases. Below we have included examples of collocations of *and then*, *and so* taken from the NAC.

AND THEN

(108) get 'em on, (xx) S1: (xx) just use, P- PowerPoint now. she scans pictures in
and then she has them in the computer, and and then projected them (xx)
 [S2: no, i don't know] i don't know how to do that yet. (LE1/NAC)

AND SO

(109) a fair amount of prior knowledge and skill and some, commitment and
 motivation to do work *and so*, people who don't have all that combination,
 sometimes drop away and, it's better for them (LE10/NAC)

So is by no means one of the markers that mostly appear in clusters. However, a quite high number of co-occurrences has the cluster *okay so*. Although *okay* has been included in the present study within the operators category with different functions (see Table 15 for NAC operators results), when *okay* accompanies *so*, it seems to adopt a meaning closer to the consecutive *so*. The fact is that in the

contexts where *okay so* appears, this cluster seems to express a summative, resultative evaluation as in the example below taken from the NAC.

OKAY SO

(110) his theory are those languages, that were there before, Latin, uh overlaid them ***okay so*** they formed a sort of linguistic substratum. and it was argued back then... (LE12/NAC)

Not as numerous as the three co-occurrences mentioned before (*and then*, *and so*, *okay so*) but still relevant is the collocation *but then*. In this case we find a contrastive marker *but* co-occurring with *then* as a **temporal** or **consecutive** micro-marker, since, as can be observed in examples 109 and 200 below, when *but* co-occurs with *then*, the latter micro-marker seems to take any of the two functional meanings temporal or consecutive, depending on the discourse context.

BUT THEN

(111) cuz it had the same numerator, and the denominator had the same first term ***but then*** it had another positive term. so in this Mundell-Flemming Model, monetary policy (LE6/NAC)-*then* as temporal.

(112) rooms nested in schools, families nested in communities definitely. [S6: okay] ***but then*** you don't want any measures on individual families you'd only have_ i mean on individual people in the families you'd have aggregates so that's fine yeah. (LE10/NAC)-*then* as consecutive.

Following with *but*, it has shown some interesting co-occurrences with this time operator *okay* resulting in *okay but*. The micro-marker *but* does not seem to lose

its contrastive meaning, although preceded by what it seems to be the use of an acceptance operator: *okay*, as the example below from the NAC shows:

OKAY BUT

(113) if I retire the utility if i retire depends on my benefits that all seems sensible.

okay but it's only when you sort of, put this paper next to the papers like Gustman and Steinmeier, where partial retirement is front and center, that you realize the partial retirement issue, is really missing. (LE7/NAC)

Although not collocating with another DM, *so what* (?) is still important in the number of occurrences and worth being taken into consideration. I found that the consecutive micro-marker *so* occurs in quite a large number of occurrences followed by a *wh*- pronoun whether as an interrogative question or as a consecutive explanation of what has been presented right before. The example presented below illustrates this use, this time not with a direct question but as the consequence of the situation explained:

SO WHAT...?

(114)ever it is, of the resources needed, nutrients needed for that particular crop. *so*

what will happen is that after a period of time, people will leave that plot alone, and move on, and start another, plot with the same crop. (LE1/NAC)

In general, we have found some relevant micro-marker collocations in the NAC, the most relevant ones according to number of occurrences and/ or frequency rate have been included in Table 17 below, highlighting *and then/ and so* as the most

recurrent ones. Other micro-marker collocations occur; however, these are not as frequent as to be mentioned here.

Micro-markers Collocations in the NAC	#DMs	‰
<i>and then</i>	163	1.3‰
<i>and so</i>	139	1.1‰
<i>okay so</i>	87	0.7‰
<i>so what</i>	42	0.3‰
<i>but then</i>	15	0.1‰
<i>okay but</i>	11	0.08‰
TOTAL	457	3.6‰

Table 17. Micro-marker collocations in the NAC.

6.6.3. Micro-marker collocations in the SC

The most recurrent micro-marker collocations in the SC and their figures can be seen in Table 18 below. A closer look at the SC collocations shows that the additional marker *y* usually collocates with other micro-markers. Moreover, micro-markers collocations seem to be as openly used as micro-markers collocations in the NAC.

A first search for micro-markers in the SC revealed some common collocations, especially with the conjunction *y*. This conjunction collocated in most of its occurrences with Spanish micro-markers such as additional *además* and temporal

or consecutive *entonces* and *luego*. In the SC *y* does not usually appear as an additive micro-marker, rather it functions as an enumerator or used to join two words together. By contrast, *y* with an additional meaning occurred when collocating with the micro-marker *además* (*y además*); eighteen of the total 49 occurrences of *además* (see Table 18 for SC micro-markers results) correspond to *y además* (see example 203).

Y ADEMÁS

(115) lo tanto los sorteos permiten conseguir bases de datos de clientes potenciales *y además* bases de datos depuradas, bases de datos creíbles y fiables, ¿por qué? (LE5/SC)

However, the collocation that showed the highest number of occurrences is *y luego*. As stated before, with the case of collocations in the NAC, one of the questions that came up is the function and role of *luego* when accompanied by the conjunction *y*, since as a micro-marker *luego* can function as an additional, temporal or consecutive micro-marker. Below are a couple of examples of *y luego* as they appear in the SC where *y luego* functions as a temporal micro-marker or with a consecutive meaning; depending on the discourse context *y luego* appears.

Y LUEGO

(116) que los artistas han de representar cualquier tipo de arte, eso está claro, *y luego* que [eleva el tono] esa tendencia esquemática [eleva el tono], sobre todo en la representación del... de la cabeza, del busto, de esta tendencia [P. ACLARA LA VOZ] a a esquematizar y a estirar el cráneo, se hace

representar también por todos los personajes de la corte de la época
(LE1/SC)

(117) es la doble articulación del lenguaje. Os lo dicto tal y como lo dice Hockett y *luego* lo comentamos, es una característica que siempre se cita como característica específica del lenguaje verbal que ya veremos si realmente lo es o no lo es. (LE10/SC)

Something similar can be observed with the collocation *y entonces*, since *entonces* can function as a quite recurrent temporal or consecutive Spanish micro-marker; however, when it co-occurs with the conjunction *y* (*y entonces*), the meaning may be affected, in all the *y entonces* instances found in the SC it seems to have a consecutive function. The co-occurrence *y entonces* together with the two previously mentioned collocations are the three most recurrent ones. The following examples show co-occurrences of *y entonces* as they appear in the SC with a consecutive meaning:

Y ENTONCES

(118) la cotiza con prima la expectativa sería expectativa de apreciación de la moneda *y entonces* el tipo de interés sería más bajo aquí que en el extranjero [pausa]. (LE7/SC)

(119) sea, si un acontecimiento, es una situación incontrolable, es externa, *y entonces* hagamos lo que hagamos, siempre ocurrirá lo mismo. Recordáis que ya os expliqué el... (LE9/SC)

Micro-markers Collocations in the SC	#DMs	‰
<i>y luego</i>	29	0.4‰
<i>y entonces</i>	19	0.2‰
<i>y además</i>	18	0.2‰
TOTAL	66	0.8‰

Table 18. Micro-marker collocations in the SC.

If we compare the cases in which micro-markers collocate in the NAC (see Table 17) we observe that there are more relevant micro-markers collocations in the NAC than in the SC, although the importance of the co-occurrence of the Spanish micro-markers is not less relevant if we look at the additional marker *and*, since it is the micro-marker that most often collocates rather than being used in isolation.

6.6.4. Macro-markers collocations in the NAC

When analyzing macro-markers in the NAC, I realized that there were also some repeated macro-markers clusters especially within the **topic-shifter** category. The co-occurrence *okay now* appeared in 19 occasions and mostly maintaining the topic-shifter function as the example 120 below illustrates. Other less numerous collocations with the macro-marker *now* but still relevant were *so now*, *and now* (see Table 19 below). In the case of *so now*, both topic-shifter macro-markers came together announcing very emphatically the topic change in the discourse. In

the case of *and now*, it seems to add and at the same time introduce a new topic in the ongoing discourse. Examples of these clusters can be read below.

OKAY NOW

(120) always useful if you really feel like, you can't do these things by hand. *okay now* that i've laid that bomb, let's go back to talking about the data for a for a little bit... (LE10/NAC)

SO NOW

(121) is positive i don't retire, if it's_ once it becomes zero i do retire. *so now* we're going to do a little notational thing we're gonna take the first (LE7/NAC)

AND NOW

(122) get the answer to be nothing. so from here on, we're leaving regression behind, *and now* we're gonna talk about H-L-M for the rest of the semester, so here it comes.(LE10/NAC)

Moving on to the **conclusion** category, I find the collocation *and finally*; eleven of the total 21 occurrences of *finally* co-occur with *and* (*and finally*). As happened before, *and finally* adds information to the ongoing lecture that is going to close the discourse or the section of talk. The following example illustrates this collocation.

AND FINALLY

(123) veil looking. it's like she's holding it up as a veil hiding behind it... *okay and finally*, Goffman says, when women and men are shown together, very frequently (LE4/NAC)

Macro-markers Collocations in the NAC	#DMs	‰
<i>okay now</i>	19	0.1‰
<i>and finally</i>	11	0.08‰
<i>so now</i>	8	0.06‰
<i>and now</i>	7	0.05‰
TOTAL	45	0.3‰

Table 19. Macro-marker collocations in the NAC.

Table 19 above shows the most relevant collocations in the NAC and their number of occurrences as well as the total rates. Compared to the micro-marker collocations in the NAC I observe that macro-markers clusters are not so numerous as micro-marker ones or that macro-markers do not seem to co-occur as much as micro-markers with other DMs.

6.6.5. Macro-marker collocations in the SC

The analysis of macro-markers in the SC also revealed some co-occurrences, especially those with *ahora* as a topic-shifter, as it happened with *now* in the NAC. The most recurrent collocation was *y ahora*, where the additional micro-marker *y*, as seen with the micro-markers, comes together with *ahora* functioning as a topic-shifter. Probably, that is why in some cases *y ahora* is followed by

volition verbs or a expression such as *vamos a ver* or *veremos* considered here as organizer macro-markers. Example 125 below shows a three-marker-collocation or cluster *ahora después veremos*. *Ahora después* can also form a collocation, in Spanish having a meaning altogether. It means ‘later’ but after a short period of time.

Y AHORA

(124) las tareas eran la administración, la creatividad, los medios y la investigación, y *ahora* se trataba de ir viendo... punto por punto estos cuatro elementos en qué (LE12/SC)

AHORA DESPUÉS VEREMOS

(125) su moneda en función de cuales sean los movimientos del mercado, *ahora después veremos* un ejemplo. ¿Cómo se interviene? [tono descendente]
Cuando la oferta de... (LE7/SC)

In the case of the collocation *ahora entonces*, we find the macro-marker *ahora* functioning as a topic-shifter but this time with the polysemous micro-marker *entonces*. Remember that *entonces* was relevant as a micro-marker when functioning as a temporal or consecutive micro-marker (see previous section on micro-marker results in this chapter); however, the question is which meaning or function *entonces* can adopt when co-occurring with the topic-shifter macro-marker *ahora*; it seems that *entonces* takes the temporal meaning of its partner *ahora*. Below there is an example of *ahora entonces* in a collocation taken from the SC.

AHORA ENTONCES

(126) la fórmula rápida que os he dado está anteriormente especificada, pero ahora, *ahora entonces*, también se puede comprobar incluso sería más fácil de demostrar (LE3/SC)

Within the **conclusion** category once more the additional micro-marker *y* co-occurs this time with a conclusion macro-marker (*finalmente*). We have to remember that in the search for conclusion macro-markers these were hard to find and very scarce. Thus, from the total number of occurrences (3 instances) of the conclusion macro-marker *finalmente* (see Table 20 for SC macro-markers results), two correspond to the collocation *y finalmente*. Although not very recurrent, we decided to include it here because once more *y* seems to generally accompany other markers, this time the scarcely used conclusion macro-marker *finalmente*.

Y FINALMENTE

(127) [P. HACE REFERENCIA A LA TRASPARENCIA] en tercer lugar seleccionaremos una estrategia comercial *y finalmente* realizaremos un análisis competitivo del mercado, un análisis que contemple la participación de la competencia (LE5/SC)

Table 20 below shows the SC most relevant macro-markers collocations as well as their number of occurrences and frequency rate. Compared to the relevant micro-markers collocations in the SC, the macro-markers seem not to have so many co-occurrences as the micro-markers collocations in the SC.

Macro-markers Collocations in the SC	#DMs	‰
<i>y ahora</i>	12	0.1‰
<i>y ahora después vamos a ver/ veremos</i>	4	0.05‰
<i>ahora entonces</i>	3	0.03‰
<i>y finalmente</i>	2	0.02‰
TOTAL	21	0.2‰

Table 20. Macro-marker collocations in the SC.

6.6.6. Operators collocations in the NAC

When analyzing the operators in the NAC I came across some collocations worthy of mention here, especially those with the DM *okay* such as *okay and* or *okay well*. *Okay* as an operator is quite a complex DM since it can convey two kinds of relation, relation speaker-speech as a **pause-filler** or relation speaker-hearer as an **acceptance** operator or **confirmation-check** (*okay?*), as I commented before in section 6.5. The most recurrent collocation is observed when the confirmation-check *okay?* co-occurs with the DM *and* (*okay? and*) (see Table 21). According to the analysis of the NAC, it seems that in those contexts where the cluster *okay? and* appears the operator *okay?* is used as a confirmation-check followed by an additional macro-maker (*and*) whose function is adding new information to the ongoing discourse. Both DMs maintain their initial functions and roles without interfering on each other as the example below illustrates.

OKAY? AND

(128) arm around you and say well now think about this a little bit more carefully.

okay? and the bottom line is gonna be in this situation you should keep working. (LE7/NAC)

In the case of *okay and*, which is more recurrent than *okay well*, the function of *okay* seems that of a **pause-filler** operator co-occurring with the additional micro-marker *and* as can be interpreted from the examples found. In such a case we get an operator and a macro-marker co-occurring and maintaining their initial meaning and functions as it occurred before with the cluster *okay? and* (see example below).

OKAY AND

(129) information that you'll gain... the option value model basically ignores that.

okay and the stochastic dynamic programming model says in a year you'll know more (LE//NAC)

There is something that drew our attention when observing the clusters *okay and/okay? and*; the fact is that except for 9 instances of these collocations, the rest were all used in Lecture 7 of the NAC, that is, Labor Economics lecture (see Table 9 in Chapter V). Idiolectal variations as well as individual lecturing styles may have affected spoken lecture discourse in this case. Only the study of a larger corpus can corroborate the use of this cluster in spoken English academic discourse in lectures.

Less used but still relevant is the collocation *okay well*, with instances that normally co-occur at the beginning of a speaker's speech section and therefore function as pause-fillers, giving the speaker time to start or re-start his/ her speech

as can be observed in example 130 below, where *well* is repeated to lengthen the time to re-start the ongoing speech.

OKAY WELL

(130)so we should not expect **the uh rationally not expect that, anything would happen. S1: *okay well well*** so far we're still talking about the, kind of normal policy. (LE7/NAC)

In Table 21 we can observe the three most important collocations for the operator taxonomy according to number of occurrences and frequency rate. Apart from the three collocation results shown here, other collocations were scarce or not very important.

Operators Collocations in the NAC	#DMs	‰
<i>okay? and</i>	21	0.2‰
<i>okay and</i>	14	0.1‰
<i>okay well</i>	9	0.07‰
TOTAL	44	0.3‰

Table 21.Operator collocations in the NAC.

6.6.7. Operators collocations in the SC

When we looked at operators in the SC we realized that there were some concrete operators that co-occurred in quite a high number of occurrences; these were the clusters *bueno pues*, *pues bien* and *¿vale? bien*. If we look back at the operator classification (see Table 16) we can observe that the first collocation is a cluster of two pause-fillers that co-occur quite often in the SC as Table 22 below shows. The joining of these two pause-fillers (*bueno pues*) makes the pause even longer, giving the speaker more time to think about what he/ she is going to say next, as example 131 illustrates; in this example the pause gets even longer since we add the hesitator *eh...* (underlined) to the pause-fillers *bueno pues*.

BUENO PUES

(131) Un sacerdote egipcio de la época románica [eleva el tono] que publica este libro Hieroglyphica que [eleva el tono] **bueno pues** eh... lo tenemos también en la biblioteca y... ilustrado lo podéis ver (LE2/SC)

Also relevant, although with fewer occurrences, is the collocation *pues bien*. This cluster seems a bit more complicated to interpret than the previous one. Although at first sight the two DMs in the collocation can be **pause-filler** operators and function as such, another interpretation could be closer to a consecutive function of the cluster *pues bien* and therefore behaves as a macro-marker. However, this could not always be the right interpretation since *pues bien* with a consecutive meaning normally occurs when *pues bien* comes after a pause or at the beginning of a speech section; this position is not as frequently used as *pues bien* in a middle position where they are clearly **pause-filler** operators. The following example

(example 132) illustrates this ambiguity, if I may call it so; however, in the next example (133) *pues bien* is clearly a pause-filler collocation.

PUES BIEN

(132) [eleva el tono] de oro del periodo están en Nubia en la actual Sudán [pausa].

Pues bien, cuando Egipto se enfrenta a los hititas eh perdón se enfrenta a... primero a los mitanios con las 16 campañas de Tutmosis III contra Mitanio (LE1/SC)

PUES BIEN

(133) porque ahora vie ¿yo sigo, no? [PREGUNTA DIRIGIDA A LA RESPONSABLE DE LA GRABACIÓN] nada *pues bien*, *pues muy bien*, pues nada eh... la paramos aquí un momento hoy es un día de todo a la vez (LE12/SC)

Coinciding with *pues bien* in number of occurrences, we have the collocation *¿vale? Bien*, where *¿vale?* seems to serve the function of a **confirmation-check**, whereas *Bien* can be interpreted as a **pause-filler** operator which, apart from giving the speaker time to think, links the discourse and keeps it going, avoiding a sudden discourse break. An example from the SC is shown below.

¿VALE? BIEN

(134) Entonces el flujo de comercio entre los dos países es este B exporta x y A exporta y, que son los bienes en los que se tiene una mayor eh... ventaja relativa *¿vale? Bien*, la siguen la siguiente expresión importante, la segunda idea importante (LE6/SC)

Table 22 below includes the frequency rate and the number of occurrences of these three most recurrent DMs collocations in the SC (*bueno pues*, *pues bien*, *¿vale? Bien*). Although with the same number of relevant collocations in the NAC (three instances), the SC has a higher frequency rate especially due to the cluster *bueno pues* with a rate of 0.4‰.

Operators Collocations in the SC	#DMs	‰
<i>bueno pues</i>	35	0.4‰
<i>pues bien</i>	12	0.15‰
<i>¿vale? Bien</i>	12	0.15‰
TOTAL	59	0.7‰

Table 22. Operator collocations in the SC

Chapter VII: Discussion

7. DISCUSSION

7.1. Introduction

In this section I am going to discuss those relevant aspects derived from the analysis of results presented in the previous chapter. The scope of this study embraces the analysis of discourse features, namely DMs, and the way they are used in the genre of lecture within the spoken discourse of the Humanities and Social Sciences.

The objective of this research project was to analyze the use of DMs in two different sub-corpora, the NAC and SC. I analyzed both the NAC, consisting of twelve North-American lectures recorded and transcribed at the University of Michigan at Ann Arbor (Michigan, USA), and then the SC consisting of twelve Spanish lectures recorded and transcribed at Universitat Jaume I of Castellón (Spain) for the purpose of this PhD dissertation. The primary research question was: Is there any difference in the use of DMs between North-American and Spanish monologic lectures in the discourse of Social Sciences?

After a review and discussion of previous literature on DMs (Murphy & Candlin 1979, Chaudron & Richards 1986, Llorente 1996, Schiffrin 1987, Portolés 1998, Fraser 1999, Morell 2001 among others), we provided our own DM classification model aimed at being valid for both English and Spanish DMs. As the DM classification model proposed in this dissertation shows (see Figure 7), DMs have been classified into micro-markers, macro-markers and operators.

Once the classification model was established, the initial research question was split up in three:

- Is there any difference in the use of **micro-markers** between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences?
- Is there any difference in the use of **macro-markers** between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences?
- Is there any difference in the use of **operators** between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences?

Secondly, I wanted to corroborate what some authors studying DMs have previously suggested, that some specific DMs tend to collocate with others in the discourse (Fraser & Malamud-Makowski 1996, Fraser 2004, Shourup 1999, Swales & Malczewski 2001). The resulting research question was: What is the relation between the different types of markers? Do some specific DMs tend to collocate? If any, which are the main differences and/or similarities of these collocations between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences? Taking these four research questions as a starting point and with the help of the results from Chapter VI, I attempt to give an interpretation of those results and discuss them. Firstly, I present general observations from the analysis in order to afterwards discuss the most relevant similarities and

differences between the NAC and SC in the use of micro-markers. Two more parallel sections comment on the results obtained in both sub-corpora with respect to macro-markers and operators. An overall synthesis of micro and macro markers in NAC and SC summarises the most important results. The last section within the discussion chapter deals with the results obtained in the search for collocations in both sub-corpora and our interpretations.

7.2. Comments on the overall results

An overall view to the results obtained from the study of DMs in the NAC and SC revealed that these linguistic units or signalling cues are more often used in the NAC than in the SC (see Figure 8). One might interpret this as the NAC tends to segment and use DMs more often than the SC; however, if we look at the average length and word per lecture results, we can see that the average number of words per lecture is higher in the NAC than in the SC as well as the average length per lecture. However, I cannot state that North-American lectures tend to structure, segment and produce a more coherent and cohesive discourse than the Spanish lecture discourse, rather I could say, and according to the results obtained, that the use of DMs in general is equivalent to the average length and words per lecture in both sub-corpora. Moreover, not all potential occurrences of DMs have been retrieved in this study. It goes without saying that both sub-corpora would have had to be completely identical in number of words and duration as to make statements

and generalizations about which type of lectures (North-American or Spanish) tend to use DMs more often.

A closer look at the three types of DMs I have distinguished here (micro-markers, macro-markers and operators) revealed that in general micro-markers are the type of DMs most widely used in both sub-corpora over the other two types. This may be due to the type of genre under study, monologic lectures. The fact that micro-markers are more often used than macro-markers and operators could be due to the peculiarities of the genre, that is, academic lectures, since this is a formal type of genre. Taking into account that micro-markers express logico-semantic relations in the discourse, it could be said that in the North-American and Spanish lecture discourses there is a need to convey lexical and descriptive meaning along the discourse of lectures as opposed to macro-markers, the least frequently used category, which convey an overall structure of the ongoing discourse and aim at segmenting and structuring utterances; although both micro and macro-markers affect and reinforce *part-of-discourse/ part-of discourse relations*. It is worth pointing out that those instances of micro-markers tend to be more fixed and less variable linguistic units, different from macro-markers, which tend to be longer chunks of language, more unsteady and sensitive to changes and prone to form language expressions, which may vary from one language to another, not having clear corresponding counterparts.

According to the present findings, operators are second in use in both sub-corpora, leaving macro-markers the last position. This is not surprising taking into consideration the specificities of a corpus like ours. I have

analyzed a spoken academic corpus and not a written one, the characteristics of an oral corpus differ from those of a written corpus. In spoken discourse *speaker-speech* or, even more so, *speaker-hearer* relations are important in the discourse community which has to share a common communicative purpose and use similar rhetorical devices in the communication process (Swales 1990). Operators are those DMs which rhetorically signal the speakers' intentions and affect the illocutionary force. Probably, the study of a written corpus would have given operators a minor and less significant position; however, the results obtained from our spoken academic discourse corpus give operators a quite important position. This fact could be linked to the trend towards a more 'open' lecture style (Swales 2002) which tends to be more participatory as has been already been mentioned in this dissertation. Relations *speaker-speech* or *speaker-hearer* gain importance resulting in a relevant use of operators in both the NAC and SC.

The least frequently used DMs in both sub-corpora (NAC and SC) were macro-markers. Although also expressing relations *part-of discourse/ part-of discourse* (as micro-markers do), these markers are characterized by conveying global discourse structural meanings and aim at segmenting and structuring utterances affecting discourse relations at a more segmental level of analysis. The fact that macro-markers are the least frequently used markers in both sub-corpora makes us think of macro-markers as having less relevance in the spoken academic discourse of North-American and Spanish lectures. As said above, it may happen that the specificities and peculiarities of spoken academic discourse do not aid and promote the use of macro-markers which would possibly be more relevant in the study of

written discourse. The relation *part-of-discourse/ part-of-discourse* seems to be primarily expressed by means of micro-markers rather than macro-markers in our NAC and SC. Previous studies regarding DMs in lecture discourse have pointed out the importance of macro-markers, which are said to be “more conducive to successfully recall of the lecture than micro-markers” (Chaudron & Richards 1986: 122). This idea is also supported by DeCarrico & Nattinger 1988, who postulated that macro-markers are beneficial for activating content schemata and more recently followed by Jung (2003). Although these authors have highlighted the importance of macro-markers rather than that of micro-markers, the findings from this study have proven that micro-markers have a prominent place in lectures, at least micro-markers have proven to be more often used and preferred by North-American and Spanish lecturers in the field of Social Sciences.

Moreover, the results have shown that operators are quite relevant in the North-American and Spanish lecture discourses provided that they can be said to be pragmatic devices specific to oral genres which strongly convey the manifestation of the illocutionary force, interaction with external elements, non-verbal information or prosody among other features. All these aspects are more closely related to oral language behaviour and not very specific to written genres.

After discussing general results, in the next section I discuss the three types of DMs proposed (micro-markers, macro-markers and operators) in more detail, comparing both sub-corpora (NAC and SC) and discussing relevant cases of DMs.

7.2.1 Comments on the micro-markers results

As I commented in the previous chapter, some differences between the NAC and SC regarding the use of micro-markers were found. In the NAC the three most frequently occurring categories are in order **Additional**, **Contrastive** and **Causal**. However, in the SC the **additional** category has the lowest rate and comes last; in the SC the three most important categories are **Contrastive**, **Causal** and **Temporal**. Thus, the main difference between both sub-corpora is in the **additional** category, undoubtedly due to the frequency of the additional micro-marker *and*.

The additional micro-marker *and* has proven to be the most frequently used marker in the NAC functioning in the majority of its uses as an **additional** micro-marker and giving this category a central place. However, its counterpart in Spanish, *y*, seems not to have the same importance in the Spanish lectures. The Spanish marker *y* as an **additional** micro-marker hardly ever occurs in isolation as the English marker *and* normally does; rather it tends to collocate with other additional expressions or markers such as *y además* or *y luego*. In our understanding this may occur because of the difference in length between the English marker *and* and the Spanish *y*. *And* as an additional marker seems to have stronger force in the discourse compared to the Spanish *y*; therefore results in a quite explicit illocutionary force. By contrast, *y* does not convey such an influential illocutionary force in the Spanish language unless it co-occurs with other expressions or markers.

The other main difference between the NAC and the SC is found in the **temporal** category. This category is very relevant in the Spanish lecture corpus, whereas it is one of the least frequently used ones in the NAC. We could say that temporal internal relations in Spanish lectures seem to be rather important, the lecturer is prone to place students and discourse in time much more than North-American lecturers which make a more extensive use of additional and contrastive micro-markers.

We have talked about differences, but similarities between the NAC and the SC are also worth mentioning. As I have already said in the previous chapter, the most relevant micro-markers in the NAC are *and*, *but* and *because*. In the SC these are *pero*, *porque* and *cuando*. Notice that except for the additional marker *and* which is very relevant in the NAC and a special case as we have already commented, and the marker *cuando*, also exclusively relevant in the SC; the other two micro-markers are counterparts in English and Spanish. *But/ pero* as a **contrastive** marker and *because/ porque* as a **causal** marker show that causal and contrastive internal discourse relations are very important in both the NAC and SC. We may say that the Spanish and North-American English lectures in the field of the Humanities and Social Sciences tend to mostly express internal (ideational) relations by means of cause and contrast DMs in most cases especially with the micro-markers *but/ pero* and *because/ porque*.

Special attention has also been drawn to the micro-marker *then*, but not because of its higher frequency rate in the NAC (although it is the most often used **temporal** micro-marker in the NAC), but because of its possible translations into the Spanish language. *Then* as a micro-marker can function

with a temporal or consecutive meaning, instances of *then* were checked in order to see this distinction. Nevertheless, we observed that *then* as a temporal micro-marker does not have a unique single counterpart in Spanish, rather *then* can be translated into different types of micro-markers. In Spanish, *then* can have two different translations *entonces* or *luego*, either with a temporal meaning or also with a consecutive meaning and function. However, if we closely observe *luego* in Spanish lectures, we will see that *luego* also occupies a relevant place as an additional micro-marker in the Spanish lectures while *then* as an additional marker can occur in the NAC but is not as relevant and recurrent as *luego* in the SC. *Luego* with a contrastive meaning in the Spanish corpus has proven to be more relevant in one specific lecture (LE4) as stated in the previous chapter. This is only an example of the complexities that can arise when comparing and contrasting DMs in two languages. We have said somewhere before in this dissertation that in order to make a comparative analysis of DMs in two languages, choosing counterparts is not the best linguistic option, take the case of *then* as a good example. Along this line, the DM classification model proposed tries to reflect the reality of the NAC and SC regardless of the symmetries of both languages. These considerations about DM counterparts should be taken into account and of great help in the field of translation studies.

As a conclusion, we can say that there are similarities in the use of micro-markers that convey internal ideational relations and affect *part-of-discourse/ part-of discourse* elements between North-American and Spanish lectures especially in the use of some specific categories such as **Contrastive** and **Causal**. However, important differences also appear in the

case of the additional category and some concrete micro-markers such as *and* with an additional meaning. Other special cases occur when trying to find counterparts or equivalent translations English/Spanish and/ or Spanish/English with some particular micro-markers as in the case of *then/ entonces/ luego*.

7.2.2. Comments on the macro-markers results

In general, the use of macro-markers in the NAC and SC seems to be quite similar if we take the three most common categories in both sub-corpora. These are **Topic-shifter**, **Organizer** and **Rephraser**, although in the SC **Rephraser** has the first position and **Topic-shifter** the last. A closer look at individual categories reveals that in the NAC the **topic-shifter** category owes its high frequency to the macro-marker *so*. *So* is the next most frequently used macro-marker in the NAC; however *so* is one of those polysemous markers that can have different meanings (we have already seen *so* as a consecutive marker in the micro-marker classification). In this case, *so* also shares the ambiguities of finding counterparts or equivalent translations in Spanish along with the micro-marker *then* mentioned before. The most frequently used macro-marker in the SC is the rephraser *es decir*, and this is the reason why **Rephraser** is the most often used category in the SC.

According to the findings, North-American lectures in the discourse of the Social Sciences tend to generally change the topic much more often than the

SC, whereas the Spanish lecturers are likely to rephrase much more than North-American lecturers when giving lecture explanations. This may result in a higher number of introductions of new topics in the North-American lectures whereas in the Spanish lectures fewer new topics are introduced, and the ones discussed are given different interpretations and/ or explanations for the better understanding of the Spanish students, as the rephraser category results would show. This does not necessarily mean that the Spanish lecture discourse is less informative than the North-American lecture discourse. We could rather say that the Spanish lecturers find the need to reword and rephrase for the students, since they may feel that students do not easily get the point of what has been explained.

It is worth mentioning some peculiarities found with respect to the Spanish rephraser *insisto*. This macro-marker is exclusive to Spanish and has no counterpart in English, since no instances or uses of *I insist* appear in the NAC. I am not saying that *I insist* cannot be used in the English language, but it is not used in the North-American lecture discourse as a rephraser macro-marker. When the lecturer uses *insisto* in the Spanish lecture discourse he/ she goes on to reword and remark information already said. It is quite curious that this rephraser use of *insisto* is widely used in and characteristic of two lectures (LE1/ LE2) within the field of Humanities and Art. We are prone to believe that once again idiolectal factors as well as individual lecturing styles can have a strong influence on the use of DMs.

Equally relevant in both sub-corpora is the **organizer** category. Lecture discourse needs to get structured for the students' retention of lecture information as the relevance of the organizer category in both sub-corpora

reveals. The most recurrent organizer macro-marker in the NAC is *let* followed by the object pronouns *me* or *us* and the colloquial relaxed varieties *lemme/let's*.

We can observe that both macro-markers (*let me, let's*) show a high frequency rate in the NAC, especially the contracted relaxed form of *let's* as opposed to the few instances of *let us*. Appealing to a dictionary *let's* is defined as follows: “You say *let us* or, in less formal English, *let's* to direct the attention of the people you are talking to towards the subject that you want to consider next.” (Collins Cobuild. English Language Dictionary). As to the definition of *let me/lemme* we find: “you can use *let me* in conversation and discussions before you ask a question, express an opinion or give an instruction” (Collins Cobuild. English Language Dictionary). According to these definitions both markers, *Let me* and *let's* are prone to be used in spoken rather than written discourse. It is interesting to mention here that in the case of *let me/lemme* with the first person singular pronoun the full form *let me* is slightly more used than the relaxed form *lemme*. Nevertheless, this difference in use is not really significant here and could be due to some transcription preferences rather than to linguistic factors. The same preference for the relaxed colloquial form *wanna* over *want to* occurs with the rephraser *I want*, following the peculiarities of spoken relaxed forms of the language.

Within the organizer category in the SC we find organizing expressions such as *ahora vamos a ver/hacer, hoy veremos*, etc. It is important to point out that time adverbs are of vital importance here; the organizer macro-marker strengthens its arranging meaning by means of time adverbs such as

hoy, ahora, el próximo día, etc. not giving a fixed or predetermined unique type of organizer macro-marker but creating new organizing linguistic units as the lecture discourse goes on. I have also noticed that the organizer *por un lado* does not necessarily imply the use of *por otro lado* in Spanish lectures; the first element appears in the discourse but not making the use of *por otro lado* compulsory. In the case of *en la/ el siguiente diapo/ tema* we have to say that the use of these organizers strictly depends on the visuals and/ or materials available for the students during the lecture delivery. These markers came to be used in only three of the total twelve Spanish lectures (LE4, LE6, and LE12), depending on the visuals and materials used or due to the lecturers' individual lecturing style.

Less important in both sub-corpora are the categories of **Starter** and **Conclusion** with very few occurrences. Even less often used than the **starter** category is the **conclusion** one. Both instances of conclusion and starter macro-markers, were hard to find. As a consequence, we could say that North-American and Spanish lectures do not tend to clearly start up or conclude the lecture speech act; rather the lecture discourse is organized in different moves by means of organizer markers, although without clear general starting or ending linguistic boundaries. The lack of starter and conclusion macro-markers could be linked to the fact that a lecture time session is not necessarily and strictly related to a lecture lesson time; that is to say, one content lesson or syllabus unit may need more than one session (lecture) to be fully and completely explained; therefore there are no clear beginning or concluding linguistic signals within a single lecture, but a need

to use organizers or topic-shifters to create an ongoing discourse structure, as the results in the NAC and SC have shown.

As a conclusion, it could be said that although sharing the same three most recurrent macro-markers categories and also the least frequently used ones, the NAC and SC differ in the most recurrent macro-marker as well as in the frequency of the categories. I have also commented on some peculiarities of individual macro-markers such as *insisto* with no identical use or counterpart in the English language.

7.2.3. Comments on the operator results

The operators results described in the previous chapter revealed at a first glance similarities between the NAC and SC, since the three most frequent categories were the same: **Pause-filler**, **Confirmation-check** and **Acceptance**. However, the most relevant category in the SC, unexpectedly, is **Confirmation-check** as opposed to **Pause-filler** in the NAC. Taking into account the two kinds of relations conveyed by operators (relations speaker-speech and/ or speaker-hearer), we observe that the most relevant category for the NAC (**Pause-filler**) conveys relations speaker-speech, whereas the most frequently used category in the SC (**Confirmation-check**) conveys relations speaker/ hearer. Presumably and according to authors such as Waggoner (1984), Benson (1994) and Mason (1994), among others, North-American lectures tend to move towards a more participatory and interactive lecturing style, if this is so, a larger number of confirmation-

checks would have been expected in the NAC than in the SC, but the results show the opposite. With a more participatory, interactive type of lecture and according to our understanding, it is expected to have more instances of those operators that convey and guide relations speaker/ hearer rather than those guiding speaker-speech relations. However, operators conveying relations speaker-hearer are more numerous in the SC (**Elicitation**, **Acceptance**, and **Confirmation-check**) than in the NAC. This could be due to the fact that Spanish lecturers need to continuously check that the audience has understood what they are explaining; either because the lecturer has the feeling the students are not having an effective lecture learning, because the topic under discussion is of a difficult nature, or due to the age of the audience (junior or senior), a junior student might feel shy to interrupt the lecturer, therefore not participating as much as a senior tertiary student, which would explain why **confirmation-check** is the most used category in the SC but still not explaining why North-American lecturers use fewer confirmation checks. One possible interpretation is that North-American University students are more used to participating in class, thus providing feedback to the lecturers in ways that do not occur in Spanish lectures. To strengthen and support my interpretation information about the audience age or seniority would be of great help; unfortunately, this kind of information is not available in the MICASE corpus.

Nevertheless, I also think that a high frequency of use of confirmation-checks does not always and strictly imply interaction between speaker-hearer, as could be expected; rather, a sign of a more interactional type of lecture can also be given by the use of acceptance operators. In this sense,

the NAC uses acceptance operators more often than the SC. When the lecturer accepts or agrees, he/ she does so because of someone's previous intervention. In order to analyze interaction and participation in the lecture discourse, the acceptance operators are a good guidance, much better guidance than the **elicitation** category. Although the main effort of elicitation markers is to invite the audience to participate in the lecture discourse, there is not always a successful outcome, and sometimes there is no answer producing default interaction. However, the use of an acceptance operator involves the existence of an earlier intervention with the result of successful interaction. A wider use of acceptance operators in the NAC would result in a higher number of students' interventions and therefore a more participatory lecture; North-American students intervene whenever they feel like or need it, without the help of the lecturer inviting them to do so, which would result in a more interactive lecturing style. As a consequence, I can still say that the North-American lecture discourse is more participatory as some authors have pointed out (Waggoner 1984, Benson 1994, Mason 1994).

Both sub-corpora have a high frequency of use of pause-fillers, which is the result of spoken language and is characteristic of an oral genre such as the lecture. Both Spanish or North-American lecturers need time to think and pause, even to breathe for a while and go ahead with the lecture delivery, this is when pause-fillers are used as lecturers' resources. This could be proof that the Spanish lecture discourse is also changing towards a more 'open' type of lecture (Swales 2002) along with the North-American lecturing style within the fields of Humanities and Social Sciences. Still

within the pause-filler category and in the SC we find *pues* as the most frequently used operator followed by *bueno* and *bien*. These two operators, although considered two separate entities, are both used for the same purpose. However, when we want to find an English counterpart, we realise that both fall under the meaning and use of the English pause-filler *well*; once more an example of the difficulties we may encounter when looking for counterpart DMs for a comparative study between two languages.

The most recurrent operator in the NAC is *okay*, but mainly due to its polysemous nature as an operator, since it can function as a pause filler, acceptance operator or confirmation-check. Also polysemous and the most recurrent is its Spanish counterpart *vale*. Although used in the Spanish lectures also as a confirmation-check and as an acceptance operator, instances of *vale* as a pause filler are not found in the SC.

The two least frequently used categories in both sub-corpora are **Elicitation** and **Attitudinal**. In the case of **Elicitation** we are not surprised, since the corpus under study is composed of monologic lectures where one speaker monopolizes the floor with scarce or little interventions; a wide use of elicitation operators was not expected in any of the two sub-corpora although one could be less monologic than the other, as explained above.

As for the **attitudinal** category we have to point out the case of the operator *I think/ we think*. *I think* has proven to be the most recurrent attitudinal operator in North-American lectures within the discourse of Social Sciences; however most of these instances implied the use of *I* and only a few appeared with the first person plural pronoun *we*. Since the specific use of personal pronouns determines the distance between speakers and

listeners, the association of the verb *think* with the pronoun *I* cannot be interpreted as a mere coincidence. In fact, some research on how personal pronouns behave in spoken academic discourse carried out recently (Morell 2001, Fortanet 2003) shows that the use of the first person singular pronoun (*I*) excludes the audience and creates a distance between speaker and hearer, as opposed to the most common meaning of the first person plural (*we*), in which speaker and hearer are usually included (Fortanet 2003). On this issue Blas Arroyo points out the existing relation between different pronoun uses depending on speech acts and communities: “las formas pronominales de tratamiento pueden adoptar diferentes valores en el discurso, y que estos se hallen íntimamente determinados por las situaciones comunicativas en cada comunidad de habla” (2005: 316).

One of the functions of the verb *think* is “to report the speakers’ own personal thoughts” (Biber et al., 1999: 669); in the same line Fortanet (2004) points out that by means of *I think*, “instead of asserting, the speaker modalizes the utterance by introducing it as an opinion”. This could also be interpreted as some kind of hedging as opposed to *I believe* as we are going to see, in any case the lecturers are giving their own personal opinion excluding the audience with the use of *I* rather than *we*.

Continuing with the **attitudinal** category, another marker analysed was *I believe/ we believe*; compared to the wide use of *I think*, the macro-marker *I believe/ we believe* has few instances, being the use of *I* or *we* equal in number and therefore not relevant as it was the case of *I think/ we think*. Our attention was drawn to the definition of *believe*. This verb, as opposed to *think*, implies a degree of commitment on behalf of the speaker. The low

frequency of this operator may be interpreted as the speaker's intentional wish to be detached from the lecture content. We could conclude that there is a North-American lecturers' intention to be detached from the delivery of a strong commitment lecture discourse.

The Spanish operator counterpart for *I think* is *(yo) creo* although not as important within the attitudinal category as the English operator *I think*. Note that no instances of *(nosotros) creemos* with the first person plural pronoun, also called royal/ academic *we* are found in the Spanish lectures. In this sense, Spanish lecturers are also detached from the audience not including it in their remarks and creating distance between the speaker and audience. The third attitudinal operator included in the taxonomy is *fijaros*, which deserves some attention because of its originality as an Spanish operator without known counterparts in English. As well as with instances of macro-markers such as *lemme* or *wanna*, which appeared in the relaxed colloquial form, the standard correct form of the plural imperative in Spanish is *fijaos*, though the colloquial *fijaros* is the one most commonly used by Spanish researchers. We have observed that when Spanish lecturers use *fijaros/ fijaos* they call the attention of the hearer towards a topic showing their attitudinal disposition. As the standard form of the imperative plural for *vosotros* with *fijaros/ fijaos* the speaker approaches the audience by addressing it as we illustrated in example 104 of the previous chapter.

Broadly speaking, the operators in the NAC and SC seem to be quite homogenous in their functions and meanings, no major differences in the set of specific operators have been found. However, I have discussed which sub-corpora could be more interactive or participatory according to the use

of some concrete categories as in the case of the acceptance category. On the other hand, there are operators that can be counterparts in only some special categories such as *okay/ vale*; by contrast, there are other operators that are unusual and distinctive of a language as it is the Spanish marker *fijaros/ fijaos*.

7.3. Comments on the overall collocations results

Through the analysis of DMs in both sub-corpora we realised that there were instances of some specific DM co-occurrences. This fact encouraged us to develop a section devoted to the study of collocations in the NAC and SC; moreover, this idea had already been pointed out by previous authors in the study of DMs who mentioned their interest in finding out how some DMs collocate and behave in clusters but not delving deeper into this (Fraser 1990, 2004, Shourup 1999, Swales and Malczewski 2001).

Thus, I gathered the most important collocations in both sub-corpora and analyzed the most frequently co-occurring cases more individually. As we have seen in the previous chapter, I found that micro-markers tend to collocate with other markers more often than macro-markers or operators and mainly in the NAC. This could be due to the fact that micro-markers are more fixed linguistic expressions and involve little or no variations, as opposed to macro-markers which are less permanent linguistic signals, closely related to context, and language dependent. To better illustrate this, we can take the case of Spanish/ English or English/ Spanish DM

counterparts; it is easier to find DM counterparts in the case of micro-markers than macro-markers, probably because macro-markers are sensitive to language and cultural variations, whereas micro-markers are types of more literal uses of linguistic devices and therefore somehow universal.

However, Spanish DMs seem to generally collocate less often than English DMs, this could be due to the fact that the SC uses fewer micro-markers than the NAC and, as we said above, micro-markers, because of their invariable nature, are prone to form collocations. Nevertheless, in general the SC tends to form fewer collocations than the NAC; this fact could also be linked to different lecturing styles and/ or discourse varieties. We have said before that according to some authors, North-American lectures tend to be more interactive and participatory and therefore somehow different from the specificities of highly monologic lectures. If the NAC tends to form collocations more often than the SC, it might happen that DM collocations are instances of more improvised, less fixed discourse elements, presumably more often used in a more interactive and less formal type of lecture.

I have found important the case of the English operator *okay*, which has shown to be prone to collocate with other markers much more than any other DM in the NAC. Probably, the operator *okay* neither conveys a very clear and strong semantic meaning nor has a significant illocutionary force when used in isolation, therefore *okay* needs to be accompanied by other markers, whichever their category (micro-markers, macro-markers or operators) to be relevant in the discourse. In the SC the DM that most frequently collocates is the additional micro-marker *y*. I have already mentioned that additive *y* in Spanish lacks strength in the ongoing discourse

because of its shortness and low significant illocutionary force, therefore it usually collocates with other markers; contrary to the English counterpart *and*, which conveys stronger meaning than *y* in isolation.

In the next section I will discuss how collocations behave regarding the proposed taxonomy (micro-markers, macro-markers and operators) in both sub-corpora.

7.3.1. Comments on the micro-markers collocations results

First I observed how micro-markers generally collocated in the NAC and SC. One of the initial questions I had in mind was whether DM clusters changed in meaning and function when co-occurring, or behaved differently when they appeared in isolation. The results regarding micro-marker collocations revealed that the micro-marker *and* tends to collocate with other markers such as *then* or *so* (*and then/ and so*). In the case of *and then*, a thorough analysis of the collocation examples found showed that when forming the cluster *and then*, *and* always behaves as an additional micro-marker whereas the second marker in the cluster, *then*, always has a temporal meaning rather than a consecutive one. Therefore, we could say that with the cluster *and then*, *then* adopts its matching part (*and*) additional meaning and functions as a temporal rather than a consecutive micro-marker. Another very recurrent collocation is *and so*. What we find here is the additional micro-marker *and* collocating with the form of consecutive

so. A closer look at *and so* instances makes us think that additional *and* is in this case reinforcing the consecutive meaning of *so* since, as we tested, if the additional micro-marker *and* is taken away, the sentence meaning does not vary (see example 109). Therefore, *and* comes to reinforce the illocutionary force of *so* when co-occurring but not necessarily transferring any additional meaning to the consecutive *so*.

I have mentioned before the cluster *and then* but there is also another cluster where *then* is present, this is *but then*. Contrary to what happened with *and then*, where *then* tends to function as a temporal marker, when co-occurring with *but*, the functional meaning of *then* is not restricted to a single choice, its semantic meaning may vary according to the discourse context in which it appears, sometimes it gets a temporal meaning, at others a consecutive one.

The cluster *okay so* also deserves our attention not only because of its wide use but also due to the co-occurrence of a micro-marker with another DM category. *Okay* is one of the most complex and polysemous DMs in the NAC; we have classified it in the operators' taxonomy within different categories: **Pause-filler**, **Acceptance** or **Confirmation-check**. Consequently, the case of *okay so* is different since two different categories of DMs collocate, one operator (*okay*) with a micro-marker (*so*). Nevertheless, observations of the *okay so* examples found in the NAC revealed that when accompanying the consecutive micro-marker *so*, *okay* does not seem to behave as an operator, it rather expresses a summative, resultative meaning closer to the consecutive meaning *so* conveys (see example 110).

As to micro-marker collocations in the SC, these are less frequently used than in the North-American lectures. This may be due to the lower rate of micro-markers in the SC. Among the three most often used collocations in the SC there is one that catches our attention and it is also the most recurrent one, *y luego*. This cluster has an identical English counterpart collocation we have discussed above: *and then*; moreover the function and meaning conveyed by *y luego* resembles the way *and then* behaves in the English lecture discourse. The marker *luego* when accompanied by *y* in the SC has a temporal meaning, as it occurred with *then* in the NAC instances co-occurring with *and*. The same occurs with the cluster *y entonces*, where *entonces* functions with a temporal rather than a consecutive meaning.

7.3.2. Comments on the macro-marker collocations results

In general macro-marker collocations did not occur as frequently as the micro-markers, which is perhaps due, as we said before, to the less linguistically consistent and more context-dependent nature of macro-markers. Still, we have found some relevant clusters in both sub-corpora.

Collocations in the NAC usually occur with DMs from the **topic-shifter** category, the most recurrent collocations functioned as topic-shifter clusters as the case of *okay now*, *so now* or *and now* express. The case of *okay now* or *and now* is different from *so now*, since in the last instance the cluster is formed by two identical categories, two macro-markers. In the case of *okay*

now an operator comes to collocate with a macro-marker, however, the meaning conveyed is that of the topic-shifter macro-marker. With *and now* the same occurs; however, here we are aware of an additional meaning conveyed by *and now*, in this case *and* maintains its additional meaning and reinforces the topic-shifter *now*. These examples take us back to one of the initial questions about collocates: Do DMs change their meaning when collocating? It seems that some markers keep their original meaning according to the taxonomy; however, there are other markers that lose or change their meaning and instead get the one of the marker they collocate with.

Along with the macro-marker collocations results in the NAC, in the SC, the macro-marker which collocated most often was also the topic-shifter *ahora* (*now* as the English counterpart). The case of *y ahora* is identical to *and now* where the additional micro-marker *y* collocates with the topic-shifter *ahora* and both markers apparently keep their original meanings. Fuzzier and more special is the case of *ahora entonces*, where two different DM categories collocate, *ahora* as a topic-shifter macro-marker and the polysemous micro-marker *entonces*. As illustrated in example 126, when the cluster *ahora entonces* occurs, the micro-marker *entonces* seems to convey a consecutive meaning rather than temporal, whereas *ahora* apparently keeps its topic-shifter meaning, however this meaning is weak and seems to fade to such an extent that if we take the macro-marker *ahora* away, the utterance is equally meaningful with no change at all. Once more we have instances of two different categories of DMs collocating and one marker

losing its meaning, in this particular case a topic-shifter macro-marker (*ahora*) gives way to a consecutive micro-marker (*entonces*).

7.3.3. Comments on the operator collocations results

Regarding operators collocations, we discovered some differences between the NAC and the SC. Firstly, that the SC tends to use operators collocations more often than the NAC, probably due to the higher rate of operators in the SC. Another difference arose with the operators that most frequently collocate: in the NAC the polysemous operator *okay* is clearly prone to collocate with other markers, whereas in the SC *pues* is more recurrent in clusters. One observation made is that operator collocations in both the NAC and the SC normally tend to collocate with other operators rather than DMs from other previous subcategories.

In the NAC the most recurrent collocation was *okay? and, okay and* and *okay well*, where the marker *okay* goes through all the meanings attributed to an operator: Confirmation-check, acceptance *okay* and pause-filler. In the SC the most frequently used collocation is *bueno pues* along with *pues bien* where we find two pause-fillers collocating. Nevertheless, this last collocation does not always behave in the SC as two pause-filler operators, examples found in the SC have shown that *pues bien* can sometimes function as a micro-marker with a consecutive meaning, which especially occurs when the cluster *pues bien* appears at the beginning of a speech section (see examples 132 and 133 in the previous chapter).

The third most recurrent collocation (*¿vale? pues*) in the SC is once more a cluster of two operators which, in this case, seem to keep their original meaning rather than one lending meaning to the other. *¿Vale?* functions as a confirmation-check, and its partner *bien* also seems to keep its function as a pause-filler, even though, they tend to form a Spanish collocation.

To conclude, we could say that when operators collocate they tend to do it with other operators and mainly within the same category, with this, what we have is an emphasis of the linguistic signal which far from having or expressing two different meanings strengthens the one conveyed, although there are some exceptions to this as we have discussed.

Chapter VIII: Conclusion

8. CONCLUSION

8.1. Concluding remarks

The current PhD dissertation took as its starting point a preliminary project (Bellés-Fortuño 2004) where I analyzed the use and function of DMs within two varieties of the English language, British and American. In that study I detected a number of important differences related to some specific DMs such as the categories of segmentation of elicitation, as well as a trend towards a more interactive kind of lecture in North-America. However, as a native Spanish speaker working in a tertiary education Spanish setting, there seemed to be a need to study Spanish lecture discourse provided that, to my knowledge, scarce research has recently been carried out on this topic. Moreover, greater interest in Spanish has lately arisen in the USA where it is the first choice for students who want to learn a foreign language, which led me to develop a contrastive analysis between Spanish and North-American English lecture discourses in order to address the findings of the study not only to tertiary education students but also to lecturers within the academic world. Therefore the aim of the current study could be summarized in two main ideas:

- a) help **native Spanish/ English lecturers** to improve their lecture discourse both in Spanish and in English.benefit and help **English/ Spanish both L2 and native undergraduate students** for the comprehension of lecture discourse in their learning process.Taking

into consideration the main beneficiaries of the study we departed from two broad research questions (RQs):

- a) Is there any difference in the use of DMs between North-American and Spanish monologic lectures in the discourse of Humanities and Social Sciences?
- b) What is the relation between the several types of markers? Do some specific DMs usually co-occur?

As the method and analysis steps were concluded, and once the classification model established, the first broad research question split up into three questions taking into consideration the three levels of analysis: micro-markers, macro-markers and operators and maintaining the fourth RQ related to DM collocations. The three-unfolded RQs are:

1. Is there any difference in the use of **micro-markers** between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences?
2. Is there any difference in the use of **macro-markers** between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences?
3. Is there any difference in the use of **operators** between North-American and Spanish monologic lectures in the discourse of the Humanities and Social Sciences?

Regarding the first broad RQ and observing the general differences between the NAC and the SC we found that, as a whole, the NAC seems to use DMs more often than the SC. One might interpret that the NAC tends to segment and use DMs more often than the SC; however, if we look at the average length and word per lecture results, we can see that both the average length in time and the number of words per lecture is higher in the NAC than in the SC.

A closer look at the three different types of markers (micro-markers, macro-markers and operators) showed that micro-markers are the type of DMs most widely used in both sub-corpora compared to macro-markers and operators. Therefore, there seems to be a greater need to express logico-semantic relations in the discourse, I could say that in the North-American and Spanish lecture discourses there is a trend to convey lexical and descriptive meaning in the discourse utterances of lectures and that this is done through DMs that affect internal and ideational relations among the discourse utterances, mainly conveying additional, temporal, causal, contrastive and/ or consecutive meanings. We can conclude that micro-markers have an important place in both the NAC and SC.

After micro-markers, operators take second place according to the frequency rate in both the NAC and SC. Operators are those DMs which rhetorically signal the speakers' intentions and affect the illocutionary force. Probably, the study of a written corpus would have given operators a minor and less significant position; however, in the study of an oral genre the reinforcement of relations *speaker-speech* and *speaker-hearer*, where the

audience takes a more participatory role, are prominent, these relations are usually conveyed and expressed through operators.

Macro-markers are the least used in both the NAC and SC. These DMs express relations *part-of discourse/ part-of discourse* (as micro-markers do). However they are characterized by conveying global discourse structural meanings and aim at segmenting and structuring utterances affecting discourse relations at a more segmental level of analysis. The fact that macro-markers are the least used markers in both sub-corpora makes us think of macro-markers as having less relevance in the spoken academic discourse of North-American and Spanish lectures. As said above, it may happen that the specificities and peculiarities of spoken academic discourse do not aid and promote the use of macro-markers which would possibly be more relevant in the study of written discourse.

Analyzing these findings, it can be said that the fact that micro-markers are the most frequently used DMs is because they in large part fill the gap of macro-makers in lectures. An explanation for this could be that lecturers and students have a well-developed knowledge of the structure and framing of a lecture and therefore macro-markers are less needed. Moreover, the lectures analyzed are part of a series of lectures that run throughout an academic term or year. None of the lectures analyzed (North-American or Spanish) belonged to an opening or closing lecture, instead of being isolated spoken events they are part of a series of lectures. Therefore, this may explain the scarce need for macro-markers compared to the extended use of micro-markers. It could happen that in an isolated spoken event such as for example an invited plenary lecture outlining and giving a global structure of

the speech event through macro-markers would likely be more necessary than in the series of lectures analyzed, since in an isolated spoken event the audience and also the lecturer probably need some aid for the framing of the spoken event.

In relation to the first more concrete RQ, this is, is there any difference in the use of **micro-markers** between North-American and Spanish monologic lectures in the discourse of Social Sciences? I can conclude that there are similarities in the use of micro-markers that convey internal ideational relations and affect *part-of-discourse/ part-of discourse* elements between North-American and Spanish lectures especially in the use of specific categories such as **Contrastive** and **Causal**. However, important differences also appear in the case of the **additional** category and some concrete micro-markers such as *and* with an additional meaning. Other special cases occur when trying to find counterparts or equivalent translations between English/Spanish and/ or Spanish/ English with some particular micro-markers as it is the case of *then/ entonces/ luego*.

When observing differences and similarities in the use of macro-markers in both the NAC and SC we concluded that although sharing the same three most recurrent macro-markers categories (**Topic-shifter**, **Organizer** and **Rephraser**) and also the least often used ones (**Starter** and **Conclusion**), the NAC and SC differ in the most recurrent macro-marker; *so* for the NAC as a topic-shifter and the rephraser *es decir* for the SC. Therefore, **Topic-shifter** for the NAC and **Rephraser** for the SC are the most often used categories. I have also commented on some peculiarities of individual

macro-markers such as *insisto* with no identical use or counterpart in the English language.

Regarding the third specific RQ: Is there any difference in the use of **operators** between North-American and Spanish monologic lectures in the discourse of Social Sciences? I can conclude that homogeneity was somehow found between the NAC and SC, since the three most outstanding categories coincided: **Pause-filler**, **Confirmation-check** and **Acceptance**. However, the most relevant category in the SC, unexpectedly, is **Confirmation-check** as opposed to **Pause-filler** in the NAC. Taking into account the two kinds of relations conveyed by operators (relations speaker-speech and/ or speaker-hearer), we observe that the most relevant category for the NAC (**Pause-filler**) conveys relations speaker-speech, whereas the most frequently used category in the SC (**Confirmation-check**) conveys relations speaker-hearer. Both categories imply a more 'open' type of lecture (Swales 2002) where lecturers do not read anymore from a lecture written text, rather, they speak from notes or an outline. Therefore, we could conclude that both sub-corpora, the SC and NAC tend to use a more interactive, 'open'-like type of lecture. We could also say that the use of a high number of confirmation-checks in the SC occurs because the speaker needs feedback of comprehension by students.

The two least used operators categories in both sub-corpora are **Elicitation** and **Attitudinal**. In the case of **Elicitation** we are not surprised, since the corpus under study is composed of monologic lectures where one speaker monopolizes the floor with scarce or little interventions; a wide use of

elicitation operators was not expected in any of the two sub-corpora although one could be less monologic than the other, as explained above.

As a conclusion, it can be said that operators seem to be quite homogenous in their functions and meanings in the NAC and SC, no relevant differences in the set of specific operators have been found. However, we have undergone a discussion about which sub-corpus could be more interactive or participatory according to the use of some concrete categories as it is the acceptance category. The use of an acceptance operator involves the existence of an earlier intervention with the result of successful interaction. A wider use of acceptance operators in the NAC would result in a higher number of students' interventions and therefore a more participatory lecture; North-American students intervene whenever they feel like or need it, without the help of the lecturer inviting them to do so, which would result in a more interactive lecturing style. As a consequence, we can go ahead saying that the North-American lecture discourse is more participatory as some authors have pointed out (Waggoner 1984, Benson 1994, Mason 1994).

On the other hand, there are operators that can be counterparts in only some special categories such as *okay/ vale*; contrarily, there are other operators that are unusual and distinctive of a language as it is the Spanish marker *fijaros/ fijaos*.

Going back to the two broad RQs earlier established, question b) referred to DMs collocations, this is: What is the relation established between the several types of markers? Do some specific DMs usually collocate? Considering the three levels of analysis (micro-markers, macro-markers and

operators) we aimed at displaying the most relevant and frequently used collocations for each type of DM. This idea had already been pointed out by previous authors in the study of DMs who mentioned their interest in finding out how some DMs collocate and behave in clusters but not going deeper into the issue (Fraser 1990, 2004; Shourup 1999; Swales & Malczewski 2001).

In general, Spanish DMs seem to collocate less than English DMs. This could be due to the fact that the SC uses fewer micro-markers than the NAC and micro-markers, because of their invariable nature, are prone to form collocations. Nevertheless, in general the SC tends to form fewer collocations than the NAC; this fact could also be linked to different lecturing styles and/ or discourse varieties. If the NAC presumably forms collocations often than the SC, it might happen that DMs collocations are instances of more improvised, less fixed discourse elements, presumably more often used in a less formal type of lecture as the ones in the NAC.

Regarding the micro-markers collocations in the NAC the most outstanding cases were instances of *and then*, *and so*, *but then* and *okay so*. One of the initial questions we had in mind was whether DMs clusters changed in meaning and function when co-occurring or behaved differently when they appeared in isolation. In the case of *and then*, *and* always behaves as an additional micro-marker whereas the second marker in the cluster, *then*, always has a temporal meaning rather than a consecutive one. A look at *and so* instances makes us think that the additional marker *and* is in this case reinforcing the consecutive meaning of *so* since if the additional micro-marker *and* is taken away, the sentence meaning does not vary. Therefore,

and comes to reinforce *so* illocutionary force when co-occurring but not necessarily transferring any additional meaning to the consecutive *so*. Contrarily to what happened with *and then*, where *then* tends to usually function as a temporal marker, with the cluster *but then*, the functional meaning of *then* is not restricted to a single choice, its semantic meaning may vary according to the discourse context in which it appears, some times it gets a temporal meaning, others a consecutive one. The cluster *okay so* has deserved special attention because in this case two different categories of DMs collocate, one operator (*okay*) with a micro-marker (*so*). Nevertheless, observations of the *okay so* examples found in the NAC revealed that when accompanying the consecutive micro-marker *so*, *okay* does not seem to behave as an operator, it rather expresses a summative, resultative meaning closer to the consecutive meaning *so* conveys.

As said above micro-markers collocations in the SC were less often used; however there are some outstanding clusters such as *y luego*, *y entonces* and *y además*. *Y luego* is the most recurrent collocation in the SC, this cluster has the identical English counterpart collocation *and then*; moreover the function and meaning conveyed by *y luego* resembles the way *and then* behaves in the English lecture discourse. The marker *luego* when accompanied by *y* in the SC has a temporal meaning, as it occurred with *then* in the NAC. The same occurs with the cluster *y entonces*, where *entonces* functions with a temporal rather than a consecutive meaning. *Y además* has fewer occurrences and it is used to express additional meaning in Spanish rather than using the single form *y*.

The most recurrent macro-markers collocations in the NAC worked as topic-shifter clusters as the case of *okay now*, *so now* or *and now* express. With *so now* the cluster is formed by two identical categories, two macro-markers. In the case of *okay now* or *and now* what we have is two different categories of markers collocating; with *okay now* an operator comes to collocate with a macro-marker, however, the meaning conveyed is that of the topic-shifter macro-marker. With *and now* the same occurs; however, in this case *and* maintains its additional meaning and reinforces the topic-shifter *now*. It seems that some markers when collocating keep their original meaning; however, there are other markers that lose or fade their meaning and get the one of the marker they collocate with.

In the SC the macro-marker that collocated most often was also the topic-shifter *ahora* (being *now* the English counterpart). The case of *y ahora* is identical to *and now* where the additional micro-marker *y* collocates with the topic-shifter *ahora* and both markers apparently keep their original meanings. We also find instances of two different categories of Spanish macro-markers co-occurring as it happens with *ahora entonces*, the micro-marker *entonces* seems to convey a consecutive meaning rather than temporal, whereas *ahora* apparently keeps its topic-shifter meaning but keeping a weak meaning and force being *entonces* the marker that conveys a stronger meaning.

Differences arose between the NAC and the SC regarding the use of operators' collocations. The SC tends to use operators collocations more often than the NAC, probably due to the higher rate of operators in the SC which results in a wider use of operators collocations. Moreover, in the

NAC the polysemous operator *okay* is clearly prone to collocate with other markers forming clusters such as *okay? and*, *okay and*, *okay well*; whereas in the SC *pues* is more recurrent in clusters such as *bueno pues*, *pues bien*. One observation made is that operators collocations in both the NAC and the SC do normally collocate with other operators rather than DMs from previous taxonomies.

To conclude, we can say that some concrete DMs collocations frequently occur within the spoken lecture genre. DMs can collocate not only with markers within the same category but with other DMs categories. Normally, the markers forming the cluster keep their original and individual meaning and functions. However, we have also seen that some markers within the same collocation seem to convey stronger illocutionary force and meaning than others, lending some of their meaning to the markers next to them; in some occasions some markers can even lose their meaning and/ or illocutionary force adopting the functions and meanings of their partner clusters.

8.2. Pedagogical implications

As said in this dissertation the lecture is the most extended practice among tertiary education institutions throughout the world (Dunkel & Davy 1989) and remains the central instructional activity (Flowerdew 1994). Classified as classroom genre within spoken discourse (Fortanet 2005), lectures are not homogeneous and static and the lecture class seems to be changing and the traditional methods are giving way to more interactive methods. Distances

between lecturers and the audience are getting narrower and formalisms avoided, at least in the U.S. and Spanish settings.

Due to the importance of the lecture genre among higher education institutions we believe that the analysis and study of lecture discourse can be beneficial both for university students and lecturers. Lecturers could definitely benefit from the findings of this study. Take for instance both L2 English and/or Spanish lecturers whose English or Spanish language level of proficiency is of different nature and background. Teacher training courses either addressed to L1 or L2 future teachers focus on two main concepts: teaching and research; however, they are basically concerned on the learning of pedagogical tools and teaching methodologies rather than on the improvement of the language level proficiency. Sometimes linguistic and grammatical aspects are taken into consideration and included as part of the syllabus; still, the main concern is that of the teaching and learning of linguistic features from written texts such as the study of DMs, meanwhile spoken discourse analysis is disregarded and neglected. In this written context, DMs are studied from an overall structural perspective, dealing with the teaching and learning of global structural markers as cohesive devices, which have been classified in this study as macro-markers. The current study is based on the analysis of spoken discourse and the findings have shed light on the importance of logico-semantic internal relations among discourse utterances in the lecture genre that are mainly conveyed by micro-markers and operators rather than by macro-markers, which are currently the most commonly features taught and learned. Therefore, we consider that teacher training courses should include in their syllabi the

teaching and learning of micro-markers and operators to future teachers and lecturers since this would result in an improvement of their general discourse and more concretely a stimulus of improvement of the quality of their lectures.

On behalf of undergraduate students both native or L2 English and Spanish, the results here obtained can be of help for the improvement of their learning practices. In general there seems to be a wide gap and a lack of bridging between high school and university, although secondary school teaching practices differ considerably from tertiary education teaching practices, where the lecture is the central instructional activity. The findings here obtained could be addressed towards the design of bridging courses for students in native or non-native contexts before or immediately after entering university. The teaching and learning of the characteristics of the lecture and features such as the DMs, as well as the differences between the lecturing practices in the USA and Spain would help undergraduate students to settle down university contexts which would in the last instance improve their learning comprehension process. On the other hand, L2 Spanish/English language courses commonly focus on the teaching and learning of general language concerns based on the four main skills, reading, writing, listening, and speaking. However, no attention is paid to the language of lectures and its peculiarities, although this would help learners in their learning process since the level of students' proficiency is not strictly related to the knowledge of lecture discourse.

As a whole, this comparative study between Spanish and English lectures makes the findings of this study useful in both the Spanish and English

tertiary school classroom settings. The results are in a first instance addressed to native Spanish and English lecturers within the field of Social Sciences that want to improve their lecture whether native or non-native discourse by using linguistic signalling devices in the correct way according to use, function and cross-cultural differences. We do not have to forget that mobility among faculty is becoming popular due to the European tertiary education programmes and the more numerous agreements and partnerships between Spanish and North-American universities. Also beneficiaries of the results obtained are native Spanish and English undergraduate students in their native or non-native (L2 context) classroom contexts that want to improve their lecture comprehension process for their successful academic training and results.

The DMs findings here presented could be included as part of the whole bunch of learning strategies and cross-cultural differences that should be taught or autonomously learned by tertiary education lecturers and students within the field of Social Sciences. In the long run the results obtained in this study could also be taken into consideration and applied in the postgraduate courses offered by universities throughout the world.

8.3. Limitations

The limitations of this research are clearly linked to the specificities of the corpus under study. We have studied twelve Spanish lectures and twelve North-American English lectures within the field of Social Sciences.

Presumably, in order to make wider and more general remarks on the use, role, function and cross-cultural differences of DMs between Spanish and North-American English lectures a broader and more numerous corpus of lectures would be needed; however, we consider that the findings obtained from the study of the two sub-corpora (NAC and SC) are valid enough to be taken as the starting point since, to our knowledge, there is not previous research such as the one hereby presented. We are also limited to two universities, the Universitat Jaume I in Castellón (Spain) and the University of Michigan in Ann Arbor (USA), which have been taken as the Spanish and North-American tertiary education institutional models. However, we are aware that lecturing styles may also vary according to institutions within the same country, even though we think these varieties and/ or differences should not be very relevant.

To conclude, a study with a broader corpus would help to make extended generalisations about the use, role, functions and cross-cultural differences between Spanish and North-American English lectures in the fields of Social Sciences. Moreover, other tertiary education institutions different from Universitat Jaume I or the University of Michigan could be used as institutional models to check if our findings can be extended to other Spanish and North-American institutions.

8.4. Further research

The results obtained in the present study attempted to shed some light on the spoken academic discourse field and also provide useful pedagogical

applications in the LAP courses. We consider that further research could be undergone regarding the study of disciplinary lecture variations in fields such as Engineering or Law to name some examples. Differences could be also observed when comparing and contrasting the use of DMs in Spanish and North-American English at a postgraduate lecture level and even PhD courses delivery modes rather than undergraduate lectures as we have done here.

The more and more widely extended incorporation of the Information and Computer Technologies (ICTs) into the university settings have resulted in the emergence of new modes and types of lecturing, what has been coined as *blended learning* (Thorne 2003, French et al. 2003). The incorporation of electronic visuals or the access to virtual classrooms, where you can find materials for the lecture course on the internet as well as the possibility to be involved in a classroom forum or submitting exercises for correction without the presence of a lecturer, (e.g. the so-called Virtual Learning Environment -Aula Virtual- at Universitat Jaume I) are really changing former classroom settings. Take for instance the on-line education (strictly related to Open University institutions) where classroom attendance is not necessary and therefore new modes of interaction and participation are used (on-line office hours and/ or tutorials, etc.). And of course, one of the possibilities that technology has brought into the academic lecture setting is the video-conferencing or teleconferencing, this latter is a live genre, where the presence of the lecturer changes the original conception, with this method a student can 'attend' a lecture, in live or one that has been previously recorded, from his/ her home. These forms of lectures are already

used by some lecturers as resources for their teaching, recorded lectures are uploaded onto a Website where students have access. Undoubtedly, these are emerging lecture genres that could be sensitive to the use, role and function of DMs in terms of participation and interaction, and therefore they should be taken into consideration for future research.

We also believe that broader studies need to be carried out regarding the use of DMs in other spoken academic genres such as seminars, conferences, workshops, PhD courses, etc. and even in teaching practices within long-live learning. Other contrastive analyses regarding the use of DMs between English and non-English speaking countries such as Sweden ,Finland or The Netherlands could be taken into consideration, provided that in these countries teaching in English practices in tertiary education are becoming very popular through the ICLHE (Integrating Content and Language in Higher Education).

Research on any of the aspects above cited could be really beneficial for both faculty and students around the educational world.

REFERENCES

REFERENCES

- Aijmer, K. & B. Altenberg (Eds.). 1991. *English corpus linguistics. Studies in honour of Jan Svartvik*. England: Longman group.
- Aldana, S. 2003. La enseñanza universitaria. Psicología del aprendizaje. Métodos de enseñanza. In López, M. (Coord.) *Problemas de la universidad: Pasado y futuro*. Valencia: Cátedra de emeritus. Fundación valenciana de estudios avanzados: 41-51.
- Atkinson, D. 2003. L2 writing in the post-process era: Introduction. *Journal of Second Language Writing* 12/ 1: 3-15.
- Barber, C. L. 1962. Some measurable characteristics of modern scientific prose. In *Contributions of English syntax and philology*. Reprinted in J. M. Swales (Ed.). 1988. *Episodes in ESP*. Hemel Hempstead: Prentice Hall International: 1-16.
- Bathia, V. K. 1993. *Analysing genre: language use in professional settings*. London: Longman.
- Bathia, V. K. 1999. Disciplinary variation in business English. In Hewings, M. & C. Nickerson (Eds.). *Business English: research into practice*. London: Longman and The British Council: 129-143.
- Bathia, V. K. 2002. A generic view of academic discourse. In Flowerdew, J. (Ed.). *Academic discourse*. London: Longman Pearson Education: 1-39.

- Bazerman, C. 1988. *Shaping written knowledge: the genre and activity of the experimental article in Science*. Madison, Wisconsin: University of Wisconsin Press.
- Bellés-Fortuño, B. & I. Fortanet. 2004. Handouts in conference presentations. In Fortanet, I, J.C. Palmer and S. Posteguillo (Eds.). *Linguistic studies in academic and professional English*. Castelló: Publicacions de la Universitat Jaume I: 63-76.
- Bellés-Fortuño, B. 2004. *The spoken academic discourse of the social sciences. Discourse Markers within the university lecture genre*. Unpublished MA Thesis. Castelló: Universitat Jaume I.
- Benson, M. J. 1989. The academic listening task: a case study. *TESOL Quarterly* 23/ 3: 421-445.
- Benson, M. J. 1994. Lecture listening in an ethnographic perspective. In Flowerdew, J. (Ed.). *Academic English: research perspectives*. Cambridge University Press: Cambridge: 181-198.
- Berkenkotter, C. & T. N. Huckin. 1995. *Genre knowledge in disciplinary communication. Cognition/ culture/ power*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bernhardt, E. B. 1983. Three approaches to reading comprehension in intermediate German. *Modern Language Journal* 67/ 2: 111-115.
- Biber, D., R. Reppen, V. Clark & J. Walter. 2001. In Simpson, R. & J. M. Swales (Eds.). *Corpus linguistics in North America*. Ann Arbor, MI: The University of Michigan Press: 48-57.

- Biber, D., S. Johanson, G. Leech, S. Conrad, & Finegan, E. (Eds.). 1999. *Longman grammar of spoken and written English*. London: Longman.
- Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. 2003. Variation among university spoken and written registers: A new multi-dimensional analysis. In Leistyna, P. & C. F. Meyer (Eds.). *Corpus analysis. language and structure and language use*. Amsterdam: Editions Rodopi B. V.: 47-67.
- Blakemore, D. 1987. *Semantic constraints on relevance*. Oxford: Blackwell.
- Blakemore, D. 1992. *Understanding utterances*. Oxford: Blackwell.
- Blakemore, D. 1995. *On so-called 'discourse connectives'*. Unpublished lecture.
- Blakemore, D. 1996. Are apposition markers discourse markers? *Journal of Linguistics* 32: 325-347.
- Blas Arroyo, J. L. 2005. *Sociolingüística del Español. Desarrollos y perspectivas en el estudio de la lengua española en contexto social*. Madrid: Cátedra.
- Bloomfield, L. 1933. *Language*. New York: Henry Holt and Co.
- Branigin, W. 1999. Spread of Spanish greeted by unwelcome signs: businesses facing language restrictions. *The Washington Post* February 6.
- Brooks, C. & R. P. Warren. 1958. *Modern rhetoric*. New York: Harcourt Brace.

- Brown, P. & G. Yule. 1983. *Discourse analysis*. Cambridge: Cambridge University Press.
- Brown, P. & S. Levinson. 1987. *Politeness: some universals in language usage*. Cambridge: Cambridge University Press.
- Cambridge International Corpus*. 2006. Cambridge University Press. URL: <http://www.cambridge.org/elt/corpus/resources.htm>
- Candlin, C. & D. Murphy. 1979. Engineering lecture discourse and listening comprehension on. *Practical Papers in English Language Education* 2: 1-79.
- Carranza, I. E. 2004. Discourse markers in the construction of the text, the activity, and the social relations. Evidence from courtroom discourse. In Marquéz-Reiter, R. & M. E. Placencia. (Eds.). *Current trends in the pragmatics of Spanish*. Amsterdam/ Philadelphia: John Benjamins: 203-227.
- Casado, M. 1991. Los operadores discursivos *es decir, esto es, o sea y a saber* en español actual: valores de lengua y funciones textuales. *LEA XIII*.
- Chafe, W. L. 1979. The flow of thought and the flow of language. In Givón, T. (Ed.). *Syntax and semantics: discourse and syntax*. New York: Academic Press: 159-181.
- Chaudron, C. & J. C. Richards. 1986. The effect of discourse markers on the comprehension of lectures. *Applied Linguistics* 7/ 2: 113-127.

- Chaudron, C. 1983. Simplification of input: topic reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly* 17/ 3: 437-458.
- Chew, P. G. 1999. Linguistic imperialism, globalism, and the English language. *The AILA Review* 13: 37-47.
- Chomsky, N. 1970. Remarks on normalizations. In Jacobs, R. & P. Rosenbaum. (Eds.). *Readings in English transformational grammar*. Waltham, MA: Blaisdell: 184-221.
- Chua, S. H. & P. G. Chew. 1993. English international -a future theme in native and non-native classrooms. *English International* 1/ 1: 49-55.
- Clyne, M. 1987. Cultural differences in the organisation of academic texts. *Journal of Pragmatics* 11: 211-247.
- Connor U. & R. B. Kaplan. (Eds.). 1987. *Writing across languages: analysis of L2 text*. Reading, MA: Addison-Wesley.
- Connor, U. 1996. *Contrastive Rhetoric: cross-cultural aspects of second-language writing*. Cambridge: Cambridge University Press.
- Connor, U. 2002. New directions in contrastive rhetoric. *TESOL Quarterly* 36/ 4: 493-510.
- Connor, U. 2003. Introduction. Contrastive rhetoric: recent developments and relevance for English for academic purposes. *Journal of English for Academic Purposes* 3/ 4: 271-276.
- Cook, J. R. S. 1975. *A communicative approach to the analysis of extended monologue discourse and its relevance to the development of teaching*

- materials for ESP*. Unpublished M. Litt. Thesis. Edinburgh: University of Edinburgh.
- Cortés de los Ríos, M. E. & M. Cruz. 2004. Análisis contrastivo inglés-español del folleto publicitario de hotel. In Sanz, I. & A. Felices. (Eds.). *Las nuevas tendencias de las lenguas de especialidad en un contexto internacional y multicultural*. Granada: Universidad de Granada: 87-95.
- Coulthard, M. & M. Montgomery. 1981. The structure of monologue. In Coulthard, M. & M. Montgomery. (Eds.). *Studies in discourse analysis*. London: Routledge and Kegan Paul: 31-39.
- Crawford, B. 2004. Interactive discourse structuring in L2 guest lectures: some insights from a comparative corpus-based study. *Journal of English for Academic Purposes* 3: 39-54.
- Crystal, D. 1997. *English as a global language*. Cambridge: Cambridge University Press.
- Curcó, C. 2004. Procedural constraints on context selection. *Siempre* as a discourse marker. In Marquéz-Reiter, R. & M. E. Placencia. (Eds.). *Current trends in the pragmatics of Spanish*. Amsterdam/Philadelphia: John Benjamins: 179-201.
- De la Cruz, M. A. 1996. Lección magistral y aprendizaje activo y cooperativo. In Blazquez, F., T. González & J. Terrón. (Eds.). *Materiales para la enseñanza universitaria*. Extremadura: I.C.E. (Instituto de Ciencias de la Educación). Universidad de Extremadura: 157-225.

- DeCarrico, J. & J. R. Nattinger. 1988. Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes* 7: 91-102.
- Décsy, G. 1986-1988. *Statistical report on the languages of the world*. Bloomington, IN: Eurolingua.
- Del Castillo, P. 2001. Riqueza y patrimonio: el español en la sociedad del conocimiento. *Nueva Revista* 74. URL: <http://www.nuevarevista.net>
- Del Saz, M. & B. Pennock. 2005. *Discourse markers of reformulation from the perspective of grammaticalization*. In Carrió, M. L. (Ed.). *Perspectivas interdisciplinarias de la lingüística aplicada*. Tomo II. València: Universitat Polotècnica de València: 89-100.
- Del Saz, M. 2003. *An analysis of English discourse markers of reformulation*. PhD dissertation. València: Universitat de València.
- Delibes, M., J. Jiménez, G. Martín & E. Santiago. 2001. Por una lengua y una cultura universales. *Nueva Revista* 77. URL: <http://www.nuevarevista.net>
- Dijk, T. A. van & W. Kintsch. 1978. Cognitive psychology and discourse recalling and summarizing stories. In Dressler, W. V. (Ed.). *Current trends in text linguistics*. Berlin/ New York: Walter de Gruyter: 61-81.
- Dijk, T. A. van & W. Kintsch. (Eds.). 1983. *Strategies of discourse comprehension*. New York and London: Academic Press.
- Dijk, T. A. van. 1972. *Some aspects of text grammars: a study in theoretical linguistics and poetics*. La Haya: Mouton.

- Dijk, T. A. van. 1977. Semantic macro-structures and knowledge frames in discourse comprehension. In P. Carpenter & M. Just. (Eds.). *Cognitive processes in comprehension proceedings of the twelfth Carnegie-Mellon symposium on cognition*. Hilldale, NJ: Lawrence Erlbaum: 3-32.
- Dijk, T. A. van. (Ed.). 1997. *Discourse as structure and process*. Discourse Studies 1. A multidisciplinary introduction. London: SAGE.
- Doménech, F. 1999. *Proceso de enseñanza/ aprendizaje universitario*. Castelló: Publicacions de la Universitat Jaume I.
- Doménech, F. 2004. *Psicología de la educación e instrucción: su aplicación al contexto de la clase*. Castellón: Publicacions de la Universitat Jaume I.
- Dudley-Evans, A. & T. Johns. 1981. A team teaching approach to lecture comprehension for overseas students. In *The teaching of listening comprehension*. ELT Documents Special. London: The British Council: 30-46.
- Dudley-Evans, A. & W. Henderson (Eds.). 1990a. *Language of economics: the analysis of the economics discourse*. ELT Documents 134. London: Modern English Publications and The British Council.
- Dudley-Evans, A. & W. Henderson. 1990b. The organisation of article introductions: evidence of change in Economics writing. In Dudley-Evans, A. & W. Henderson (Eds.). *Language of economics: the*

analysis of the economics discourse. ELT documents 134. London: Modern English Publications and The British Council: 67-78.

Dudley-Evans, A. 1994. Variations in the discourse patterns favoured by different disciplines and their pedagogical implications. In Flowerdew, J. (Ed.). *Academic listening: research perspectives*. Cambridge: Cambridge University Press: 146-158.

Dufrenne, M. 1963. *Language and philosophy*. Bloomington, IN: Indiana University Press.

Dulay, H. & M. Burt. 1974. Natural sequences in child second language acquisition. *Language Learning* 24: 37-53.

Dunkel, P. & S. Davy. 1989. The heuristic of lecture note taking. Perceptions of American and international students regarding the value and practice of note taking. *English for Specific Purposes* 8: 33-50.

Dunkel, P. A. & J. N. Davis. 1994. The effects of rhetorical signalling cues on the recall of English lecture information by speakers of English as a native or second language. In Flowerdew, J. (Ed.). *Academic listening: research perspectives*. Cambridge: Cambridge University Press: 55-74.

Enkvist, N. E. 1997. Why we need contrastive rhetoric. *Alternation* 4/ 1: 108-246.

Enkvist, N. E. 1984. Contrastive linguistics in text linguistics. In Fisiak, J. (Ed.). *Contrastive linguistics: prospects and problems*. The Hague: Mouton: 45-67.

- Escandell-Vidal, M. V. 1993. *Introducción a la pragmática*. Barcelona: Antropos and UNED.
- Fishman, J. 1977. The sociology of language: yesterday, today and tomorrow. In Cole, R. W. (Ed.). *Current issues in linguistic theory*. Bloomington, IN: Indiana University Press: 51-75.
- Flowerdew, J. & S. Tauroza, 1995. The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition* 17: 435-458.
- Flowerdew, J. (Ed.). 1994. *Academic listening: research perspectives*. Cambridge: Cambridge University Press.
- Flowerdew, J. (Ed.). 2002. *Academic discourse*. London: Longman Pearson Education.
- Flowerdew, J. 2003. Register-specificity of signalling nouns in discourse. In Leistyna, P. & F. Meyer. (Eds.). *Corpus analysis. Language and structure and language use*. Amsterdam: Editions Rodopi B.V.: 35-46.
- Fortanet, I., S. Posteguillo, J. C. Palmer & J. F. Coll. (Eds.). 1998. *Genre studies in English for academic purposes*. Castelló: Publicacions de la Universitat Jaume I.
- Fortanet, I. 2003. Verbal stance in spoken academic discourse. Paper presented at *Evaluation in academic discourse conference*. Siena (Italy). June 2003.
- Fortanet, I. 2004. The use of 'we' in university lectures: reference and function. *English for Specific Purposes* 23: 45-66.

- Fortanet, I. 2005. Honoris causa speeches: an approach to structure. *Discourse Processes* 7/1: 31-51
- Foucault, M. 1970. *The archaeology of knowledge and the discourse on language*. New York: Pantheon books.
- Foucault, M. 1980. Interviews. In Gordon, C. (Ed.). *Power/ Knowledge - Selected Interviews and Other Writings 1972-1977*. Brighton: Harvester Press.
- Fraser, B. & M. Malamud-Makowski. 1996. English and Spanish contrastive discourse markers. *Language Sciences* 18/ 3-4: 863-881.
- Fraser, B. 1987. Pragmatic Formatives. In Verschueren, J. & M. Bertucelli (Ed.). *The pragmatic perspective*. Amsterdam: John Benjamins.
- Fraser, B. 1988. Typed of English discourse markers. *Acta Linguistica Hungarica* 38/ 1-4: 19-33.
- Fraser, B. 1990. An approach to discourse markers. *Journal of Pragmatics* 14: 383-395.
- Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics* 31: 931-952.
- Fraser, B. 2004. An account of discourse markers. In Garcés, P., R. Gómez, L. Fernández, & M. Padilla. (Eds.). *Current trends in intercultural, cognitive and social pragmatics*. Sevilla: Universidad de Sevilla: 13-34.
- Frederick, P. J. 1986. The lively lecture -8 variations. *College Teaching* 34/ 2: 43-50.

- Freedman, A. & P. Medway. (Eds.). 1994a. *Genre and the new rhetoric*. London: Taylor & Francis.
- Freedman, A. & P. Medway. 1994b. Locating genre studies: antecedents and prospects. In Freedman, A. & P. Medway. (Eds.). *Genre and the new rhetoric*. London: Taylor & Francis: 1-22.
- French, D., N. Olrech, & C. Hale (Eds.). 2003. *Blended learning: an ongoing process for Internet integration*. Canada: Trafford Publishing.
- Fries, C. 1952. *The structure of English*. New York: Harcourt, Brace, and Company.
- Fuentes, C. 1987. *Enlaces extraoracionales*. Sevilla: Alfar.
- Giménez, R. 1998. The ancestors of an oral field of discourse. In I. Fortanet, S. Posteguillo, J. C. Palmer & J. F. Coll (Eds.). *Genre studies in English for academic purposes*. Castelló: Publicacions de la Universitat Jaume I: 297-237.
- Giménez, R. 2000. *La repetición lingüística en el género de la clase magistral: El inglés académico oral en el ámbito de las Ciencias Sociales*. Unpublished PhD Dissertation. València: Universitat de València.
- González, M. 2004. *Pragmatic markers in oral narrative: the case of English and Catalan*. Amsterdam: John Benjamins.
- González, M. 2005. Pragmatic markers and discourse coherence relations in English and Catalan oral narrative. *Discourse Studies* 7/ 1: 1461-4456.

- Goffman, E. 1981. The lecture. In Goffman, E. (Ed.). *Forms of talk*. Philadelphia, PA: University of Philadelphia Press: 162-195.
- Gracián, B. (1651-1655). El Criticón. In Cantarino, E. (Ed.). *Ética y política en las primeras obras de Baltasar Gracián*. Madrid: Espasa-Calpe, 1998.
- Graddol, D. 1997. *The future of English?* Great Britain: The British Council.
- Graddol, D. 1999. The decline of the native speaker. *The AILA Review* 13: 57-68.
- Greenbaum, S. 1991. ICE: The International corpus of English. *English Today* 28: 3-7.
- Grice, H. P. 1975. The logic of conversation. *Speech acts. Syntax and semantics* 3: 40-58.
- Grice, H. P. 1989. *Studies in the way of words*. Cambridge, Mass: Harvard University Press.
- Gutiérrez, S. 1993. Hacia dónde va el funcionalismo sintáctico. *Español Actual* 60: 13-34
- Gutiérrez, S. 2003. La universidad. esencias y fines. In Lopez, M. (Coord.). *Problemas de la universidad: Pasado y futuro*. Valencia: Cátedra de emeritus. Fundación valenciana de estudios avanzados: 69-81.
- Halliday, M. A. K., P. Strevens. & A. McIntosh. 1964. *The linguistic sciences and language teaching*. Oxford: Oxford University Press.

- Halliday, M. A. K. & R. Hasan, 1985. *Language, context and text: aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.
- Halliday, M. A. K. 1985. *Spoken and Written Language*. Geelong, Vic: Deakin University Press.
- Halliday, M. A. K. 1989. *Spoken and Written Language*. Oxford: Oxford University Press.
- Halliday, M. A. K. 1994. *Functional Grammar*. London: Edward Arnold.
- Hamel, R. E. 2002. *El español como lengua de las ciencias frente a la globalización del inglés. Diagnóstico y propuestas de acción para una política iberoamericana del lenguaje en las ciencias*. Paper presented at Congreso internacional sobre las lenguas neolatinas en la comunicación especializada. México.
- Henderson, W. & A. Hewings. 1990. A language of model building? In Dudley-Evans, T. & W. Henderson (Eds.). *Language of economics: the analysis of economics discourse*. ELT Documents 134. London: Modern English Publications and the British Council: 43-54.
- Hiller, J. H., G. A. Fisher & W. Kaess. 1969. A computer investigation of verbal characteristics of effective classroom lecturing. *American Educational Research Journal* 6: 661-675.
- Hinds, J. 1979. Organizational patterns in discourse. In Givón, T. (Ed.). *Syntax and semantics: discourse and syntax*. New York: Academic Press: 159-181.

- Hinds, J. 1983. *Contrastive rhetoric: Japanese and English*. *Text* 3/ 2: 183-196.
- Hirvela, A. 1997. Disciplinary portfolios and EAP writing instruction. *English for Specific Purposes* 16/ 2: 83-100.
- Hobbs, J. 1985. On the coherence and structure of discourse. *Technical report CSLI -85-37*. Center for the Study of Language and Information, Stanford: Stanford University.
- Holmes, R. 1997. Genre analysis and the social sciences: an investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes* 16/ 4: 321-337.
- Hovy, E. 1995. The multifunctionality of DMs. *English for Specific Purposes* 13/ 3: 239-256.
- Hyland, K. 1998. *Hedging in scientific research articles*. Amsterdam: John Benjamins.
- Hyland, K. 2000. *Disciplinary discourses: social interaction in academic writing*. London: Longman Pearson Education.
- Hymes, D. 1981. *"In vain I tried to Tell You": essays in native American ethnopoetics*. Philadelphia: University of Pennsylvania Press.
- Hyon, S. 1996. Genre in the three traditions: implications for ESL. *TESOL Quarterly* 30/ 4: 693-722.
- Ibáñez, A. & A. Ortigosa. 2004. A semantic typology of causative accomplishment movement verbs and their argument-adjuncts in role and reference grammar. *Atlantín* 26/ 2: 35-49

- Jakobson, R. 1960. Linguistics and poetics. In Sebeok, T. A. (Ed.). *Style in language*. New York. Trad. cast. de A. M. Gutiérrez y estudio preliminar de F. Abad: *Lingüística y poética*. Madrid: Cátedra, 1981, 1988 4a ed.:39-78.
- James, A. 2000. English as a European lingua franca. Current realities and existing dichotomies. In Cenoz, J. & U. Jessner. (Eds.). *English in Europe. The acquisition of a third language*. Clevedon, UK: Multilingual Matters: 22-38.
- James, A. Forthcoming. Die 'neue' Mehrsprachigkeit: Englisch als globale und regionale lingua franca. In James, A. (Ed.). *Vielerlei Zungen. Mehrsprachigkeit Plus*. Klagenfurt: Celovec.
- Jenkins, J. 2004. ELF at the gate: The position on English as a lingua franca. *The European English Messenger* 13/ 2:63-69.
- Jenkins, J. 2006. Global intelligibility a local diversity: possibility or paradox? In Rubdi, R. & M. Saraceni. (Eds.). *English in the world: global rules, global roles*. London: Continuum: 32-39.
- Jiménez, J. 2001. Por una lengua y una cultura universales. *Nueva Revista* 77. URL: <http://www.nuevarevista.net>
- Johns, A. M. 1981. Necessary English: a faculty survey. *TESOL Quarterly* 15: 51-57.
- Johnson, K. J. 1996. Political transitions and the internationalisation of English: implications for language planning, policy making, and pedagogy. *Asia Pacific Journal of Education* 16: 20-40.
- Jones, J. F. 1999. From silence to talk: cross-cultural ideas on students

participation in academic group discussion. *English for Specific Purposes* 18/ 3:243-259.

Jordan, R. R. 1977. Study skills and pre-sessional courses. In Holden, S. (Ed.). *English for Specific Purposes*. Oxford: Modern English Publications: 24-26

Jordan, R. R. 1989. English for academic purposes (EAP). *Language Teaching* 22/ 3: 151-164.

Jung, E.H. 2003 The role of discourse signalling cues in second language listening comprehension, *The Modern Language Journal* 87 (iv): 62-577.

Kachru, B. B. 1984. *English as a world Language*. Cambridge: Cambridge University Press.

Kaplan, R. B. 1966. Cultural thought patterns in intercultural education. *Language Learning* 16/ 1: 1-20.

Kaplan, R. B. 1983. Contrastive rhetorics: some implications for the written process. In Freedman, A., I. Pringle & J. Yalden. (Eds.). *Learning to write: first language/ second language*. New York: Longman: 139-161.

Kaplan, R. B. (Ed.). 1984. *Annual review of applied linguistics, 1983*. Rowley, MA: Newbury House.

Kaplan, R. B. 1987. Cultural thought patterns revisited. In Connor, U. & R. B. Kaplan. (Eds.). *Writing across languages: analysis of L2 text*. Newbury Park, CA: Sage: 9-22.

- Kaplan, R. B. 1988. Contrastive rhetoric and second language learning: notes towards a theory of contrastive rhetoric. In Purves, A. (Ed.). *Writing across languages and cultures: issues in contrastive rhetoric*. Newbury Park, CA: Sage: 275-304.
- Kerans, M. H. 2001. Live delivery of lectures in a discipline-relevant, sustained-content practice. *TESOL Journal* 10/ 2-3: 13-17.
- Khuwaileh, A. A. 1999. The role of chunks, phrases and body language in understanding co-ordinated academic lectures. *System* 27: 249-260.
- Kintsch, W. & J. C. Yarbrough. 1982. Role of rhetorical structure in text comprehension. *Journal of Educational Psychology* 74/ 6: 828-834.
- Knapp, P. & C. Meierkord. (Eds.). 2002. *Lingua Franca Communication*. Frankfurt am Main: Peter Lang.
- Knott, A. & R. Dale. 1994. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18/ 1: 35-62.
- Lahuerta, A. C. & M. F. Pelayo. 2003. Usos marginales de los marcadores del discurso. Su efecto en la comprensión lectora en español como lengua extranjera. *Ibérica* 2: 49-67.
- Lahuerta, A. C. 2004. Discourse markers in the expository writing of Spanish university students. *Ibérica* 8: 63-80.
- Leech, G. 1991. The state on the art in corpus linguistics. In Aijmer, K. & B. Altenberg. (Eds.). *English corpus linguistics. Studies in honour of Jan Svartvik*. England: Longman group: 8-29.

- Leistyna, P. & C. F. Meyer. (Eds.). 2003. *Corpus analysis. Language and structure and language use*. Amsterdam: Editions Rodopi B. V.
- Levinson, S. C. 1979. Pragmatics and social deixis. In Chiarello, C. (Ed.). *Proceedings of the fifth annual meetings of the Berkeley Linguistic Society*. Berkeley, CA: Berkeley Linguistic Society.
- Liebman, J. D. 1992. Toward a new contrastive rhetoric: differences between Arabic and Japanese rhetorical instruction. *Journal of Second Language Writing* 1/ 2: 141-166.
- Llorente, M. T. 1996. *Organizadores de la conversación. Operadores discursivos en español*. Salamanca: Publicaciones Universidad Pontificia de Salamanca.
- López, A. M. 2003. La lengua internacional de los medios de comunicación: Una convergencia de modelos lingüísticos. In Perez, M. & J. Coloma (Eds.). *Actas de XIII Congreso internacional de la Asociación ASELE*. Murcia: 522-532.
- Malamud-Makowski, M. 1997. *Discourse markers in Spanish*. PhD Dissertation. Boston: Boston University.
- Malatene, C. 1985. Contrastive rhetoric: an American writing teacher in China. *College English* 47: 789-808.
- Mann, W. & S. Thompson. 1987. *Rhetorical structure theory: a theory of text organization. Technical Report RR/ 87/ 190*. Marina del Rey, CA: Information Sciences Institute.
- Mann, W. & S. Thompson. 1988. Rhetorical structure theory: toward a functional theory of text organization. *Text* 8: 243-281.

- Marcos-Marín, F. 2001. El español, segunda lengua internacional. *La Razón* January 12.
- Marques de Tamarón et al. (Dir.). 1995. *El peso de la lengua española en el mundo*. Valladolid: Publicaciones de la Universidad de Valladolid, Fundación Duques de Soria and INCIPE.
- Marques de Tamarón. 2006. *El guirigay nacional. Ensayos sobre el habla de hoy*. Áltera: Barcelona.
- Marsà, I. 1992. *Estudio contrastivo de los marcadores de discurso en inglés y español: marcadores de transición*. PhD Dissertation. Barcelona: Publicacions de la Universidad de Barcelona.
- Martín, M. A. & J. Portolés. 1999. Los marcadores del discurso. In RAE (Ed.). *Gramática descriptiva de la lengua española 3*. Madrid: Espasa Calpe: 4051-4213.
- Mason, A. 1994. By dint of: Student and lecturer perceptions of lecture comprehension strategies in first-term graduate study. In Flowerdew, J. (Ed.). *Academic English: research perspectives*. Cambridge: Cambridge University Press: 199-218.
- Mauranen, A. 1993. *Cultural differences in academic rhetoric*. Frankfurt am Main: Peter Lang.
- Mauranen, A. 1998. Another look at genre: corpus linguistics vs. genre analysis. *Studia Anglica Posnaniensia* 2: 303-315.
- Mauranen, A. 2001. Reflexive Academic Talk: Observations from MICASE. In Simpson, R. C. & J. M. Swales. (Eds.). *Corpus*

linguistics in North America. Ann Arbor, MI: The University of Michigan Press: 165-178.

Mauranen, A. 2003a. The corpus of English as a lingua franca in international settings. *TESOL Quarterly* 37/ 3: 513-527.

Mauranen, A. 2003b. "But here's a flawed argument": socialisation into and through metadisocurse. In Leistyna, P. & C. F. Meyer. (Eds.). *Corpus analysis. Language and structure and language use*. Amsterdam: Editions Rodopi B. V.: 19-34.

McDonough, J. 1978. *Listening to lectures*. Oxford: Oxford University Press.

McNanaway, L. A. 1979. Teaching methods in higher-education-innovation and research. *Universities Quarterly* 24: 321-29.

Miller, C. R. 1984. Genre as social action. *Quarterly Journal of Speech* 70/ 2: 151-167.

Mohan B. A. & A-Y. Lo. 1985. Academic writing in Chinese students: transfer and developmental factors. *TESOL Quarterly* 19/ 3: 515-534.

Montaño-Harmon, M. 1988. *Discourse features in the compositions of Mexican, English-as-a-second language, Mexican/ American Chicano, and Anglo high school students: considerations of the formulation of educational policies*. PhD dissertation. California: University of Southern California.

- Montaño-Harmon, M. 1991. Discourse features of written Mexican Spanish: current research in contrastive rhetoric and its implications. *Spania* 74: 417-425.
- Montgomery, M. 1977. *Some aspects of discourse structure and cohesion in selected science lectures*. Unpublished MA thesis. Birmingham: University of Birmingham.
- Morell, T. 2000. *EFL Content Lectures: a discourse analysis of an interactive and a non-interactive style*. Alicante: Imprenta Gamma.
- Morell, T. 2001. The role of discourse markers and personal pronouns in lecture discourse. In Moreno, A. I. & V. Cowell. (Eds.). *Perspectivas recientes sobre el discurso*. León: Universidad de León.
- Moreno, A. I. 1996. *Estudio contrastivo inglés-español de la expresión de las relaciones de coherencia causal interoracional: El artículo académico sobre economía y empresa*. León: Publicaciones de la Universidad de León. (CD electronic format: *Tesis doctorales* 1995, vol 26).
- Moreno, A. I. 1997. Genre constrains across languages: causal metatext in Spanish and English. *ESP Journal* 16/ 3: 161-179.
- Moreno, A. I. 2004. The explicit labelling of the premise in premise-conclusion sequences: an English-Spanish contrastive study of research articles on business and economics. In Sanz, I. & A. Felices. (Eds.). *Las nuevas tendencias de las lenguas de especialidad en un contexto internacional y multicultural*. Granada: Universidad de Granada: 159-168.

- Moreno, F. 2000. *Qué español enseñar*. Madrid: Arco Libros. Cuadernos de Didáctica del Español/ LE.
- Moreno-Garcerán, A. 2001. Lengua y cultura, activos intangibles de la nueva economía. *Nueva Revista* 74. URL:<http://www.nuevarevista.net>
- Morris, C. 1938. *Foundations of the theory of signs*. Chicago: University of Chicago Press.
- Morrison, J. W. 1974. *An investigation of problems in listening comprehension encountered by overseas students in the first year of postgraduate students in science at the University of Newcastle upon Tyne*. Unpublished M. Ed. Thesis. Newcastle: University of Newcastle upon Tyne.
- Munby, J. 1978. *Communicative syllabus design*. Cambridge: Cambridge University Press.
- Murphy, D. F., & C. N. Candlin. 1979. Engineering lecture discourse and listening comprehension. *Practical Papers in English Language Education* 2. Lancaster: University of Lancaster: 1-79.
- Myers, G. 1992. Texts books and the sociology of scientific knowledge. *English for Scientific Purposes* 11/ 1: 3-17.
- Nakane, C. 1970. *Japanese society*. Berkeley: University of California Press.
- Newmeyer, F. J. 1980 *El Primer cuarto de siglo de la gramática generativo-transformista (1955-1980)*. Madrid: Alianza.

- Oliver, R. T. 1965. Foreword. In Nathanson, M. & H. W. Johnstone Jr. (Eds.). *Philosophy, rhetoric and argumentation*. Panama: University Park: x-xi.
- Olsen, L. A. & T. N. Huckin. 1990. Point-driven understanding in engineering lecture comprehension. *English for Specific Purposes* 9: 33-47.
- Perelman, C. H. & L. Olbrechts-Tyteca. 1969. *The new rhetoric: a treatise on argumentation*. Notre Dame, IN: University of Notre Dame Press.
- Péry-Woodly, P. M. 1990. Contrasting discourses: contrastive analysis on a discourse approach to writing. *Language Teaching* 23/ 3: 143-151.
- Polanyi, L. & R. Scha. 1983. The syntax of discourse. *Text* 3: 261-270.
- Portolés, J. 1993. La distinción entre los conectores y otros marcadores del discurso español. *Verba* 20: 141-170.
- Portolés, J. 1998. *Marcadores del discurso*. Barcelona: Ariel Practicum.
- Purves, A. (Ed.). 1988. *Writing across languages and cultures: issues in contrastive rhetoric*. Newbury Park, CA: Sage.
- Quirk, R. 1960. Towards a description of the English usage. *Transactions of the Philological Society*: 40-61.
- Raimes, A. 1991. Our of woods: emerging traditions in the teaching of writing. *TESOL Quarterly* 25/ 3: 407-430.
- Räisänen, C. *The conference forum as a system of genres*. PhD dissertation. Göteborg: Acta Universitatis Gothoburgensis.

- Redeker, G. 1990. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14/ 3: 367-381.
- Redeker, G. 1991. Review article: linguistic markers of discourse structure. *Linguistics* 29/ 6: 1139-1172.
- Reid, J. 1988. *Quantitative differences in English prose written by Arabic, Chinese, Spanish and English students*. PhD dissertation. Colorado State University.
- Rendle-Short, J. 2003. So what does this show us?: Analysis of the discourse marker so in seminar talk. *Australian Review of AppliedLinguistics* 23: 46-62.
- Richards, J. C. 1983. Listening comprehension: approach, design, procedure. *TESOL Quarterly* 17/ 2: 219- 239.
- Rivas, F. 1993. Modelo integrado de situación educativa (MISE): una aproximación desde la psicología de la instrucción. In Pelechano, V. (Ed.). *Psicología, mitopsicología y postpsicología*. Valencia: Promolibro: 293-338.
- Rivas, F. 1997. *El proceso de enseñanza/ aprendizaje en la situación educativa*. Madrid: Ariel.
- Rodríguez-Ponga, R. 1998. El español, lengua universal. *Nueva Revista* 60: 75-99.
- Rounds, P. 1987. Characterizing successful classroom discourse for NNS teaching assistant training. *TESOL Quarterly* 21/ 4: 643-669.

- Russel, D. 1991. *Writing in the academic disciplines, 1870-1990: A curricular history*. Carbondale: Southern Illinois University Press.
- Salehzadeh, J. 2006. *Academic listening strategies*. Ann Arbor, MI: The University of Michigan Press.
- Sanders, T., W. Spooren & L. Noordam. 1992. Towards a taxonomy of coherence relations. *Discourse Processes* 15: 1-35.
- Santana-Seda, O. 1970. *A contrastive study in rhetoric: an analysis of the organization of English and Spanish paragraphs written by native speakers of each language*. PhD dissertation. New York: New York University.
- Santiago, R. L. 1970. *A contrastive analysis of some rhetorical aspects in the writing in Spanish and English of Spanish speaking college students in Puerto Rico*. PhD dissertation. New York: New York University.
- Sapir, E. 1912. Language. In Mandelbaum, D. G. (Ed.). *Selected writings of Edward Sapir in language, culture, and personality*. Berkeley: University of California Press: 7-32.
- Saussure, F. 1916. *Cours de linguistique générale*. Publié par C. Bally et A. Sechehaye avec la collaboration de A. Riedlinger. Edition critique préparée par Tullio de Mauro. Paris: Payot (Payothèque), 1978. Trad. cast. de Alonso, A.: *Curso de lingüística general*. Buenos Aires: Losada, 1945.

-
- Sawa, G. H. 1985. *Modified Input as an aid to comprehension*. Hawaii: University of Hawaii at Manoa. Manuscript.
- Schiffrin, D. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Scollon, R., & S. B. K. Scollon. 1981. *Narrative, literacy, and face in interethnic communication*. Norwood, NJ: Ablex.
- Seidlhofer, B. 2001. Closing a conceptual gap: the case for a description of English as a lingua franca. *International Journal of Applied Linguistics* 11/ 2: 133-158.
- Seidlhofer, B. 2004. Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics* 24:209-239.
- Seidlhofer, B. Forthcoming. English in the expanding circle: What it isn't. In Rubdi, R. & M. Saraceni. (Eds.).
- Shourup, L. 1999. Discourse Markers. *Lingua* 107: 227-265.
- Silva, T. 1993. Toward an Understanding of the Distinct Nature of L2 Writing: The ESL Research and its Implications. *TESOL Quarterly* 27/ 4: 657-677.
- Simpson, R. C. & J. M. Swales. (Eds.). 2001. *Corpus linguistics in North America*. Ann Arbor, MI: The University of Michigan Press.
- Sinclair, J. 1987. *Looking up: an account of the COBUILD project*. London: Cobuild.

- Skulstad, A. S. 1996. Rhetorical organisation of chairmen's statements. *International Journal of Applied Linguistics* 6/ 1: 43-63.
- Skulstad, A. S. 2002. *Established and emerging business genres*. Kristriansand: HoyskoleForlaget.
- Sperber, D. & D. Wilson. 1986. *Relevance: communication and cognition*. Oxford: Blackwell.
- Stubbs, M. 1983. *Discourse analysis*. Chicago, IL: The University Chicago Press.
- Swales, J. M. & B. Malczewski. 2001. Discourse management and new episode flags in MICASE. In Simpson, R. & J. M. Swales. (Eds.). *Corpus linguistics in North America*: 145-164.
- Swales, J. M. & A. Burke. 2003. "It's really fascinating work": differences in evaluative adjectives across academic registers. In Leistyna, P. & C. F. Meyer. (Eds.). *Corpus analysis. Language and structure and language use*. Amsterdam: Editions Rodopi B. V.: 1-18.
- Swales, J. M. 1981. *Aspects of article introductions*. Aston ESP Research Reports 1. Birmingham: University of Aston, Language Studies Unit.
- Swales, J. M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M. 2002. Integrated and fragmented worlds: EAP materials and corpus linguistics. In Flowerdew, J. (Ed.). *Academic discourse*. London: Longman Pearson Education: 150-164.

-
- Swales, J. M. 2004. *Research genres. Exploration and application*. Cambridge: Cambridge University Press.
- Sweetser, E. E. 1990. *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Tannen, D. (Ed.). 1982. *Spoken and written language: exploring orality and Literacy*. Norwood, N.J.: Ablex.
- Tauroza, S. & D. Allison. 1994. Expectation-driven understanding in information systems lecture comprehension. In Flowerdew, J. (Ed.). *Academic listening: research perspectives*. Cambridge: Cambridge University Press: 35-54.
- The British Academic Spoken English (BASE) corpus*. 2006. On behalf of the Centre for English Language Teacher education. Warwick University. URL:
<http://www2.warwick.ac.uk/fac/soc/celte/research/base/>
- The British National Corpus*, version 2 (BNC world). 2001. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>
- The MICASE (Michigan Corpus of Academic Spoken English)*. 2002. Simpson R.C., S.L. Briggs, J. Ovens, & J.M. Swales. URL:
<http://micase.umdl.umich.edu/m/micase/>
- Thompson, S. 1994. Frameworks and contexts: a genre-based approach to analysing lecture introductions. *English for Specific Purposes* 13/ 2: 171-187.

- Thorne, K. 2003. *Blended learning: how to integrate on-line and traditional learning*. Kansas: Sterling.
- Trujillo, F. 2003. Culture in writing: discourse markers in English and Spanish student writing. In Departamento de didáctica de la lengua y la literatura (Ed.). *Tadea seu liber de Amicitia*. Granada: Imprenta Generalife: 345-364.
- Ventola, E. 1987. *The structure of social interaction: a systemic approach to the semiotics of service encounters*. London: Frances Printer.
- Ventola, E. 1992. Writing scientific English: overcoming cultural problems. *Journal of Applied Linguistics* 2/ 2: 191-220.
- Vygotsky, L. S. 1978. *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press. Published originally in Russian in 1930.
- Vygotsky, L. S. 1986. *Thought and language*. Cambridge, MA: The MIT Press.
- Waggoner, M. 1984. The new technologies versus the lecture tradition in higher education: is change possible? *Educational Technology* 24/ 3: 7-12.
- Whorf, B. L. 1956. Grammatical categories. In Carol, J. B. (Ed.). *Language, thought and reality: selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press: 87-101.

- Widowson, H. G. 1993. The ownership of English. *IATEFL Annual Conference Report*, Plenaries 1993. Whitstable: IATEFL.
- Wilkinson, R. (Ed.). 2004. *Integrating content and language: Meeting the challenge of a multilingual higher education*. The Netherlands: Maastricht University.
- Wilson, D. & D. Sperber. 1993. Linguistic form and relevance. *Lingua* 90/1-2: 1-25.
- Woods, N. 1978. *College reading and study skills*. 3rd ed. New York: Holt, Rinehart and Winston.
- Yates, J. & W. J. Orlikowski. 1992. Genres of organisational communication: a structural approach to studying communication and media. *Academy of Management Review* 17/ 2: 299-326.
- Young, L. 1994. University lectures-macro-structure and micro-features. In Flowerdew, J. (Ed.). *Academic listening: Research perspectives*. Cambridge: Cambridge University Press: 159-176.
- Yunick, S. 1997. Genre, registers and sociolinguistics. *World Englishes* 16/3: 321-336.

