

**UNIVERSIDAD POLITÉCNICA DE
VALENCIA**



**DEPARTAMENTO DE LINGÜÍSTICA
APLICADA**

TESIS DOCTORAL

**AN ESP COMPARATIVE ANALYSIS IN MEDICAL
RESEARCH ARTICLES: SPANISH-ENGLISH**

Presentada por:

ALICIA RICART VAYÁ

Dirigida por:

MIGUEL ANGEL CANDEL MORA

VALENCIA, 2008

Prof. D. Miguel Ángel Candel, profesor contratado doctor del Departamento de Lingüística Aplicada de la Universidad Politécnica de Valencia

CERTIFICA:

Que Dña. Alicia Ricart Vayá, Licenciada en Filología Inglesa y Filología Alemana en la Universidad de Valencia, ha realizado bajo su dirección la presente tesis titulada:

**AN ESP COMPARATIVE ANALYSIS IN MEDICAL
RESEARCH ARTICLES: SPANISH-ENGLISH**

para la obtención del título de Doctora.

Y para que conste a los efectos oportunos, firma la presente certificación.

Valencia, a 28 de Mayo de 2008

Fdo. Prof. D. Miguel Ángel Candel Mora

**TO MY PARENTS,
MY BROTHER AND SISTER
AND MY AUNT JUANA.**

There are many people who have helped me in the process of writing this dissertation. First of all I would like to thank my supervisor for his helpful advice, his support and understanding especially in those times when I felt lost.

My colleagues at the department of Linguística Aplicada in the Polytechnic University of Valencia have also been a great source of inspiration and I am very grateful to them. Among them I would especially like to thank Cristina Pérez Guillot for her encouragement and good sense of humour, Frances Watts for her support and Asunción Jaime for her knowledgeable advice and suggestions.

I am also very grateful to Mr. Sperling for his help and kindness and for allowing me to consult their medical on-line library in the University Hospital Benjamin Franklin during my stay in Berlin. His kindness has enabled me to turn my dissertation into a European one. I would also like to thank Frau Preinesberger, Frau Krüger and Frau Schulten for their help.

I owe much gratitude to my best friends at the university, who have always been there for me for any consultation despite the distance. Among them I would especially like to thank my friends Ana María Pons Sanz, Ana Sánchez, Lutz Holländer, Eva Serrano and Laura Macías who have been an essential support throughout my studies and professional career both for their friendship and for their professional guidance.

There are many people who have not read one single word of my dissertation but without whose help I would not have been able to complete it: all the doctors interviewed and all those who have helped me to contact the specialists. But my greatest debts are to my family. My parents, my brother and my sister and my aunt Juana have always supported me in everything I have done, and it is to them that I would like to dedicate this dissertation.

Finally, I would like to thank Marco, for his understanding, guidance, his sensible advice, his patience, his trust in me and love. He has been a source of constant encouragement and moral support.

It is my family and him that have enabled me to finish my dissertation. I would also like to apologise for all the times I was in a bad mood. I deeply thank you all very much.

CONTENTS

CONTENTS

I.	Introduction	
1.1.	Objectives.....	6
1.2.	Parts of the dissertation.....	9
II.	Theoretical background	
2.1.	The Research Article.....	32
2.1.1.	Genre and Discourse Community.....	32
2.1.2.	The evolution of the Research Article.....	36
2.1.3.	The Structure of RAs: The Introduction, Method, Results and Discussion model (I-M-R-D model).....	40
2.1.4.	Adaptation of the conventional I-M-R-D model and the analysis of the “moves” to the medical corpora.....	45
2.1.5.	Contrastive analysis of RA Conclusion.....	47
2.2.	Academic Writing in Research Articles.....	52
2.2.1.	Main features of Research Articles.....	55
2.2.2.	Linguistic features of medical Research Articles.....	61
2.3.	Contrastive Rhetoric.....	66
2.3.1.	English and other European Languages.....	71
2.3.1.1.	English and non- European Languages.....	76
2.3.2.	Different points of view on Contrastive Rhetoric.....	77
2.3.3.	Contrastive Rhetoric in this research work.....	78
2.4.	Medical language.....	80
2.4.1.	Origin of medical language and its influence on the Spanish language.....	80
2.4.2.	New emerging vocabulary	82
2.4.3.	Different ways to increase medical	

Vocabulary.....	85
2.5. Corpus linguistics.....	95
2.5.1. Size of corpus.....	99
2.5.2. Computer tools for corpus analysis.....	104
2.5.3. Different areas of study.....	105
2.5.4. Corpus linguistics and translation.....	107
2.5.5. Characteristics of the corpora used for this research work.....	109

III. Analysis of the corpus

3.1. Corpus design.....	115
3.1.1. Selection of the corpus.....	117
3.1.1.1.Overcoming limitations and determining the aim and field of the corpus.....	117
3.1.1.2.Specifying Conclusion sections in journals.....	118
3.1.1.3.Selecting the journals.....	122
3.1.1.4.Online Sources to access medical Journals.....	124
3.1.1.5.Data: selected English and Spanish journals for the corpus.....	132
3.1.1.6.Index Factor of the Journals.....	137
3.1.2. Codification of the conclusion sections.....	141
3.1.3. Lexical Analysis with WordSmith™.....	144
3.2 Survey to investigate medical needs.....	147
3.2.1 Survey's rationale.....	147
3.2.2 Targeted audience.....	145
3.2.3 Survey design validation.....	150
3.2.4 Structure of the survey.....	153
3.2.4.1 First Section: personal information....	154
3.2.4.2 Second Section: access to medical journals.....	158

3.2.4.3	Third Section: publishing papers.....	159
3.2.4.4	Fourth Section: style sheets.....	164
3.2.5	Transcription to SPSS statistical system	166
3.2.6	Results of the survey.....	172
3.3	Lexical Analysis with WordSmith TM	184
3.3.1	Results.....	184
3.3.1.1	Nouns.....	189
3.3.1.2	Adjectives.....	198
3.3.1.3	Adverbs.....	208
3.3.1.4	Conjunctions.....	213
3.3.1.5	Verbs.....	218
3.3.2	Linguistic Analysis of the Moves.....	262
3.3.2.1	Analysis of the move <i>Background</i>	264
3.3.2.2	Analysis of the move <i>Summarizing</i>	303
3.3.2.3	Analysis of the move Evaluation of the Study: <i>Limitations and Advantages</i>	325
3.3.2.4	Analysis of the move <i>Further Research</i>	377
3.3.3	Neologisms in the Spanish medical corpus analyzed.....	415
IV.	Conclusions.....	451
V.	Conclusiones.....	459
VI.	Bibliography.....	471
	Index of Tables and Figures	
	Abbreviations	
	Summaries: Summary in English, Summary in Spanish and Summary in Valencian.	
	Appendixes (attached CD)	
	I (Corpus)	
	II (Surveys)	
	III (Lists)	

INTRODUCTION

I. Introduction

Language is functional and its development depends to a large extent on society's needs. Language is always evolving and needs reforming because the world changes and our mental consciousness has to adjust to catch up with the changes in the world. This fact can be seen in the work done by scientific researchers, who need to find new terminology for constantly emerging new technology. They have to translate experience into words, to translate their findings into language to make us aware of new realities. Progress in science needs to be revealed through texts. In fact, it is through language that people get socialized to adapt to a new or a different setting.

The idea that social structures as they evolve impact on language while, at the same time, language instantiates and is a reflection of those social structures has led researchers to be concerned with the development of discourse models, particularly regarding the relationship between lexico-grammatical realisations, semantics and society.

The command of foreign languages has been generally accepted as a vital requirement for professional and academic life and thus for international communication (García Hoz, 1990; Guillén Díaz, 2000; Morales Gálvez, 2000; Nussbaum, 2001). This command is not only relevant for the dissemination of new technological and scientific developments, but it is also important for educational, cultural and tourist reasons.

Some authors state that, if English is the preferred language of international communication in Science, English should be seen as the language of Science and not as the language of Englishmen or Americans. As explained by Seidlhofer (2005:339-340), given the intrinsic characteristics of English as a Lingua Franca (further referred to as ELF) as a language used by people belonging to many different language backgrounds, the language has been suffering changes as a result of non-native users transferring discourse patterns of their first language or mother-tongue (further referred to as L1).

Seidlhofer (2005:339-240) points out that a lot of research has been carried out recently by different authors who try to analyze the features of this emerging and evolving “international language” at different levels; at the level of phonology (Jenkins, 2000), pragmatics (Meirkord, 1996) and at the lexicogrammar level (Seidlhofer, 2004:209-239). She explains that the international use of this language has led to the need to study its nature. It should be borne in mind that non-native speakers, who are the vast majority, namely those who speak it as a foreign language or second language, are shaping the language at least as much as native speakers do and consequently this language is turning into a different language to that of English as a native language. To illustrate this idea, she explains that despite the fact that most ELF forget to put the “-s” to mark the third person singular in the present tense, this does not give rise to communication problems. This implies that the English to be taught for students who intend to use it mainly in international settings should be different to that for learners

who intend to master fine nuances at a native speaker language level.

The increasing demand in the working world of professionals able to communicate with people around the world has resulted in the subject of English for Specific Purposes being included in most university programmes (Hyland, 2002).

The predominance of Latin from the Renaissance until the end of the 18th century is comparable to that of English at the present time. The predominance of one language permits researchers to be up to date by just the command of that language. However, on the counterpart, there is, as stated above, an unconscious association between the quality of a text and the language in which it has been written. In fact, nowadays investigators do no longer read or translate an academic article written in a different language from English, since they take for granted that the most important researchers publish their papers in English. Thus anything published in a language other than English is necessarily believed to be of an inferior quality or is unimportant. According to Flowerdew and Peacock (2001:10) “the international language of research and academic publication is English”. Also relevant is the comment of Alcaraz (2000:15): “en la mayor parte de las distintas especialidades, las publicaciones de revistas en lengua inglesa son las de mayor prestigio y difusión internacional”.

In the field of Medicine, Spanish speakers have turned the English sententia “*publish or perish*” into “*publish or muere*”, i.e., there is no other way to gain recognition as a researcher in the medical international community than by

publishing in English (Bernard, 1989:5) This reaches an international level, since other linguists explain that the same happens in other countries (see, for instance, Ren, 1999:286 on China and Walvoort, 1997:5-7 on Holland). Navarro (2000, XIX) states: “El inglés es el idioma de la medicina en todo el mundo (...) sin su conocimiento no se puede acceder a los principales libros de texto y las revistas con los últimos avances en cualquier campo de la medicina”.

It is the need of Spanish researchers to publish in English that has motivated the present research. Concretely, the need of the former to command the language so that they find it easy to belong to the scientific community and are not hindered because of a lack in the command of the language.

1.1. Objectives

Much research has been carried out in the study of language structure and *Moves* in research articles as means for dissemination of knowledge and in the analysis of the constituent parts of the discourse organization of research articles.

It can be emphasized the work carried out on linguistics regarding medicine (Salager-Meyer, 1990, Luzón, 2000; Ballesteros, 2003; Amador Iscla, 2003; Navarro, 2005; Piqué and Posteguillo, 2006, etc.), analysis on collocation frameworks (Baker, 1992; Posteguillo, 1999; Luzón, 2000; Moreno, 2004, etc.), and studies comparing two languages (Mauranen, 1993; Valero-Garcés, 1996; Trujillo, 2001; Ciapuscio and Otañi, 2002;

Connor, 2002, etc.), but there is still a gap for a study focusing on a linguistic comparison searching for lexico-grammatical patterns to establish equivalences in English and Spanish in the Conclusion Section of Medical research articles.

This lack of research in the literature together with the assumption that there is a need on the part of Spanish medical doctors to command the English language in order to succeed in writing medical Research Articles has been the basis for the current research.

Thus, the motivation of this research work starts upon confirming that the Conclusion Sections – more specifically in the field of medicine – had been devoted less attention since, many times, this Section was included under the heading Discussion. Therefore, it has been considered appropriate to contribute to the better knowledge of this Section to facilitate the task of medical researchers when publishing their papers in international journals.

It should also be taken into account that it takes non-English speakers longer to write in English than to do so in their L1. Furthermore, the predominance of the English puts non-English speakers at a disadvantage when attending international congresses because they cannot intervene the way they would if they commanded the language. Thus, only a minority is able to contribute orally in congresses.

The problems faced by non-English speaking scientists are manifested by different authors (Crystal, 1997:14), and especially by the ex-chief editor of *Nature Biotechnology*, Christopher Edwards:

Scientists who do not speak or write English well may be subjected to discrimination or misunderstanding if they try to publish in top general-science journals. If they present their work at international scientific meetings and their English language skills are poor, their findings may not receive the attention and acknowledgment they deserve. The overall losses are great for English as foreign language speakers who do not refine their skills. Discrimination can hurt them in terms of getting hired, promoted, tenured and funded (cp. Vanderbrouck, 1989:1461-1462).

Therefore, the main objective of this research work is to analyze the existence of parallel structures in English and Spanish regarding academic writing. In order to achieve this general objective it has been considered appropriate to base this research work on the following specific objectives:

- To assess the real linguistic needs on the part of Spanish medical doctors collecting data by means of a survey.
- To check the existence of a representative number of Conclusion sections in Medical research articles in English and Spanish which should enable the creation of the corpora for analysis.
- To divide these Conclusion sections into moves on the basis of the semantic content of the conclusions.
- To verify the existence of certain recurrent lexicogrammatical patterns in the Conclusion Section of Medical research articles both in English and Spanish.
- To compare Spanish and English recurrent lexicogrammatical patterns in the moves: *Background*,

Summarizing, Limitations and Further Research and find differences and similarities.

- To find possible equivalences among the linguistic patterns in the 4 different moves.

Once established the objectives, the next step is to explain the process of the research.

1.2. Parts of the dissertation

This research work has been divided into six chapters.

After this chapter which presents an overview of the main features, objectives and motivation of the present research, Chapter two presents a framework of the theoretical background for the research work. Section 2.1 on the importance of English as a lingua franca in the scientific arena covers the characteristics of general academic writing including key concepts such as “genre”, “discourse community”, “moves” and “I-M-R-D” conventions.

Section 2.2 focuses on research articles and their main characteristics revising aspects of interest such as “quality parameters”.

Next, Section 2.3 presents the linguistic features of scientific English. It deals with Contrastive Rhetoric among English and European and non European languages.

Section 2.4 revises literature on medical research articles. A study on possible incorrect terms resulting from the spread of the English language in the medical community is carried out in

this section to give substantial data of this influence on the Spanish language.

A final essential section of the second part of this research work is based on the contribution of Corpus Linguistics to the analysis of large amount of data which allowed comparing and studying different linguistic features.

Chapter three, based on the theoretical background presented in the second part of the research, presents the core of this research work from the section on corpus design, the selection of the journals and the Sections for analyses to the description of computer tools used.

In order to obtain real data from active practitioners in the field of medicine a survey was designed and carried out, therefore section 3.2 describes the complete process from the design of the survey to the analysis of the data obtained. This survey would determine the needs of Spanish doctors regarding the use of the English language for their publications.

Section 3.3 provides an in-depth analysis of the most characteristic and recurrent features mentioned in the literature on corpus linguistics: from lexical analysis, including tables showing the most recurrent morphological items in the corpora, to analyses of the moves.

First, it was necessary to collect a representative number of English and Spanish medical journals containing a Conclusion Section. In order to establish the most recurrent lexico-grammatical structures of the Section and compare the two languages, Conclusions were analyzed in terms of moves. To do so, the model of Yang Ruiying and Desmond Allison

(2003): “*Research Articles in Applied Linguistics: moving from Results to Conclusions*”, was followed but adapted to the new setting of medical research articles. It was decided to focus the linguistic analysis on 4 moves: *Background*, *Summarizing*, *Limitations* and *Further Research*. Next, similar structures containing the same lexical items were listed in each move. A list containing recurrent lexical items present in both languages was created. Lexical items in each move were given a number and equivalences in the two languages were searched for.

Finally, the chapter entitled “Conclusions” summarizes the most outstanding aspects found during the research work carried out based on the objectives established in section 1.1.

In order to contribute to the illustration of the samples analysed it was considered appropriate to include three appendixes in electronic format containing the data analysed: the corpora, the tables with the analyses and the survey.

THEORETICAL BACKGROUND

II. Theoretical background

In recent years a fast development in technology and commerce has led researchers to put their experiences into words. In this process, English has become the “lingua franca” of scientific and commercial sectors (Alcaraz, 2000:14). It is obvious that English is vital for academic life and international communication and thus anyone who intends to succeed in the scientific community has to write necessarily in English.

Linguists agree that English is the most widely foreign language used in Europe (Petzold and Berns, 2000:113-124; Berg et al., 2001:305-319). For Crystal (1997:3-54) in order to reach the status of global language, a language needs to be taken up by other countries. This happens for instance in some countries that make another language the official language and use it for the media, educational system, the government, etc. (e.g. India). Their inhabitants are bound to command this official language, called foreign language, as early as possible, because it is a complement to their native language. This is the status that English has in over seventy countries, such as Ghana, Nigeria, India, Singapore, etc. The second way in which a language is taken up is by making it a priority in the country’s foreign-language teaching, despite not having an official status. Kachru (2001) illustrates the spread of English around the world and its acquisition and use by means of 3 concentric circles (Figure 1.1):

a) the inner circle, containing between 320 and 380 million people, represents those countries in which English is the primary language. People in this circle speak Standard British or American English (e.g. the USA, the UK, Ireland, Canada, Australia and New Zealand),

b) the outer or extended circle, with 150-300 million people, represents those countries where English was spread in a non-native setting and the language has become part of a country's main institutions, playing therefore a relevant "foreign language" role in a multilingual setting (e.g. Singapore, India, etc),

c) and finally the expanding circle, represented by 100-1,000 million people, which includes those countries which consider English an international language but have no history of colonization by those in the inner circle. The language does not have an administrative status either (e.g. China, Japan, Greece, etc). In this case English is called a foreign language. English has this status in over 100 countries, such as China, Russia, Germany, Spain, etc.

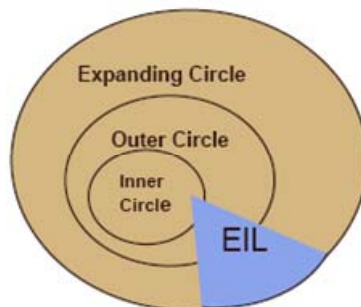


Figure 1.1 Use of international English across Kachru's circles.

It is therefore not surprising that English is currently the most widely taught read and spoken language (Kachru and Nelson, 2001). According to a study carried out by Aragonese et al. in 2002, (Ruiz, 2003:14) 90% of the students in all E.U. countries study English, whereas only 34 % study French, 15% German and 10% Spanish. As far as the percentage of language speakers in the E.U. is concerned, 47% of all speakers have English as either their mother-tongue (16%) or as a foreign or second language (further referred to as L2) (31%).

Increasingly, the status of English as an *international language* (Bhatia, 1997b:313-319; Widdowson, 1997:134-146), or a *global language* (Crystal, 1997) has led to the importance of its teaching and learning not only in Europe but all over the world. Alcaraz (2000:14) adds that English as a foreign Language is one of the most important components of the syllabus in primary and secondary schools all over the world.

Graddol (1997a) suggests that in the near future the number of speakers of English as a foreign language will be greater than that of those who speak it as their L1 and as their L2. According to Crystal (1997:130), despite the number of L1 speakers being greater at the moment, if current population and learning trends continue, L2 speakers will overcome those with English as an L1 in 10 years time because the countries of L2 speakers have a much greater growth rate. These expectations lead us to the idea that the only possible concept of ownership of the language in some years will be a global one.

In fact, in recent years, Seidlhofer (2005:339) refers to the emergence of the term “*English as a lingua franca*” (ELF) to

define a language spoken among non-native speakers. This is because only one out of every four speakers in the world is a native speaker of English (Crystal, 2003). For Firth (1996: 240) ELF is “a contact language” between people who share neither a common native tongue nor a common national culture, and for whom English is the chosen foreign language of communication”. Despite this definition, native speakers are not excluded from ELF communication; in fact, native speakers use English as a lingua franca to communicate with non-native speakers, but the term ELF is more suitable to refer to interactions among non-native speakers. For this reason, there are other labels for ELF such as “English as a medium of intercultural communication” (Meierkord, 1996) or “English as an international language” (Richards et al., 1992:187).

Among the definitions of “*international language*” Richards et al. (1992:187) description should be noted: “a language in widespread use as a Foreign Language or Second language, i. e. as a language of international communication”.

Likewise, Halliday (2002), states that an “*international language*” might have been a regional one that turns into the most spoken at a certain point of time or into a global communication vehicle. He adds that there has been a change regarding economy: in the past it was based on products and services, while nowadays economy relies on information. Given that the vehicle language for the transmission of that information is English, it has turned into an international language.

Crystal (1997:5-7) suggests that no language can become an international medium of communication without a strong political, military or economic power-base; there is a close link between language dominance and cultural power. Therefore, it is necessary to oppose those who state that the spread of the English language has to do with aesthetic qualities, clarity of expression, a logical structure or a simple grammar lacking genre and number and inflectional endings. In fact, Latin was once an international language despite having these linguistic features. Some learners find English internationally appealing due to the fact that English has borrowed many words from other languages with which it has been in contact over the centuries and thus it sounds “familiar” and “cosmopolitan” to them, others consider the language more “democratic” than others due to the absence of a grammar system coding social class differences. However, these positive traits are incidental and contrast with other negative features such as the very irregular spelling system of the language. Thus, ease of learning a language and appealing or inconvenient linguistic features are not relevant for a language to become international, nor can any of these traits stop it from achieving this status either.

There are many reasons why English has become an international language; firstly, as stated above, English is one of the languages with the highest number of L1 speakers (Alcaraz, 2000:14). Additionally, there are several political and historical reasons that have led many countries to adopt English as an L1 or L2 (Alcaraz, 2000). Generally speaking, the establishment of English-speaking colonies in North America at the beginning of

the 17th century was the first decisive stage in the colonial expansion of England which made English an international language. Furthermore, the fast technological development of English speaking countries, especially the USA, explains its condition as a “lingua franca” (Graddol, 1997b). This technological development has consequently spread the English language to other countries turning English into the “lingua franca” of scientific and commercial sectors (Graddol, 1997b; Alcaraz, 2000:14).

It seems then that one language rather than another is used because, at a certain point in history, its speakers impose its use through commerce, culture and politics. History shows that there is clear evidence that the language is always “compañera del Imperio” (Empire companion) and that political structures of domination are not permanent. To illustrate this fact it is worth mentioning the role that other languages (viz. Latin, Italian, Spanish and French) have successively played as a “lingua franca” since the Renaissance (Capel, 2004:742).

In a recent work, Graddol (1997:171-179) studies the difficulties that the English language had to face prior to being considered suitable for scientific writing. According to linguists, there is no language better than another or one that is simpler or more suitable than another to be used around the world. In fact, quite curiously, in the 17th century the writing of original science works in English was delayed due to its inadequacy: it lacked the necessary technical vocabulary and the grammatical resources required to represent the world in an objective and impersonal way.

It was the Royal Society that in 1664 encouraged the development of a suitable writing style and fostered the publication of Science in English. One year later, the first international scientific journal in English was inaugurated: *Philosophical Transaction*. This century was a formative period for the establishment of scientific English. As one of the needs of the new scientific community lay in terminology, translators of classical works used several techniques to face this lack, but the most used was the one consisting in borrowing the Latin term and adapting it to English morphology. German became however the leading European language of science in the 18th century (Duque García, 1999:15), 401 journals in German had been established in contrast to 96 in French and 50 in English by the end of the century, but English gained the first place in the 19th century thanks to a substantial lexical growth in the fields of technology and science due to the Industrial Revolution. By the beginning of this century, English had become the world's leading industrial and trading country and no nation could equal Britain's economic growth. Most inventions of the Industrial Revolution were of British origin (e.g. the harnessing of coal, water and steam to drive heavy machinery, the emergence of new means of transportation, etc). Britain produced at that time in textiles and mining so many goods for export, that this country was called the "workshop of the world". Besides, during the 19th century the British colonial expansion in Africa and the South Pacific took place with its consequent adaptation of English as a semi official language by many newly independent states.

Additionally, in the 19th century the USA was the most productive country and the growth of its population was the fastest in the world (Duque García, 1999:15). In fact, the current spread of the English language is rather a direct consequence of the USA's political, military and economic supremacy since the end of World War II. At this time Europe was devastated. Its research centres, mainly those in Germany, France and England, lost their importance and the economic and technical pre-eminence of the USA made European countries, which were looking forward to be rebuilt, turn to the English language so as to promote their scientific findings internationally. The situation provoked that many people who wished to take advantage of the innovations in English-speaking countries learned the language so as to benefit, and at the same time, many Britons travelled abroad and earned a good living by teaching the new methods of industrial production. Crystal (1997:73) explains that Britain became a magnet of opportunity attracting inventors from the continent. By the end of the century, America had overtaken Britain economically and it also acted as a magnet for European scholars. The power of the American Science in the 1970s turned the English language into the international language of science. The American and British research between 1750 and 1900 represents half of the influential scientific and technological output worldwide.

Another factor that has contributed to the paramouncy of the English language is the birth of the Internet in the 21st century (Alcaraz, 2000:15). According to Crystal (1997:107) 80% of the information stored in the Internet is written in

English. This is not surprising taking into account the statistics provided by the BBC News in 2001 which indicate that “9 out of ten computers connected to the Internet are located in English-speaking countries”. Graddol defines the Internet as *the flagship of global English* and Gil Pou (2000:24) describes the Internet as:

Increíble fuente de datos, artículos, imágenes, fotos, etc. semejante a una biblioteca gigantesca en la que es posible encontrar todo tipo de información en la que el idioma preponderante para viajar por sus autopistas es el inglés.

Regarding Medicine, Navarro (2001:37-38) deals with the increasing eagerness to write in English. He also refers to the relevant role that the appearance of the Internet has played in the storage of data for this science. Some years ago, medical doctors who did not speak English as their L1 started increasingly publishing their papers in English both in international journals and in national ones, since many of the latter began to be published in English. This phenomenon is illustrated by Maher (1986, 206-218), who states that only 8% of the medical papers published in Germany in 1966 were written in English. However, this percentage rose to 18% in 1970s, 32% in 1980s and 45% in 1985. From this moment on in the Western countries it is generally accepted that good quality research is published in English and in international journals. In fact, as Navarro points out (1992:575): “El inglés es el medio de comunicación internacional en el ámbito de las ciencias de la salud y se considera como una segunda lengua común a médicos e

investigadores” (“English is the international communication means in health sciences and it is considered a common second language to doctors and researchers”).

The present eagerness to write in English has led the scientific committees of international congresses and the editorial boards to be formed mainly by American scientists (Aberch, 1996:257-264):

The language can foster cronyism in the scientific community and thus, unwillingly, handicapping the non-English-speakers members, in the sense that they are outsiders in the circle of people that influence decisions and editorial policies.

Pérez-Eid (1990:75-79) carried out a study in which she showed that 75% of scientists in editorial boards of 433 journals with a high impact factor were English-speakers and 8 out every group of 10 were American. (Some further evidence of the dominance of the English language in the scientific world is the fact that only 4 Americans were given an award for the Nobel Prize between 1901 and 1930 whereas 109 were given it in the period of time between 1970 and 1999). The main drawback of this fact is that as people who form these committees are the ones who accept or reject articles to be published, they tend to select those that are more related to their topics of research, their concepts and their methodologies, which are those present in the USA. Among the authors who condemn this situation it is worth mentioning López Piñero and Terrada (1992, 98, 64-68) in their statement:

La dependencia del imperialismo científico, actualmente el angloamericano, significa importar no solamente conocimientos e ideas, sino también patrones de conducta y valores ajenos, que son asumidos, por lo general, de modo inconsciente. Ello conduce a mecanismos extremadamente peligrosos... (...), como bloquear el peso de la tradición científica propia y llenar la laguna resultante con información procedente en este caso de la angloamericana, desde planteamientos que consideran como una norma su cultura, cuando no están formulados desde un mezquino nacionalismo estadounidense o británico.

It is obvious that there is a tendency to concentrate on those issues relevant in the USA as a direct consequence of the keenness to write in English and publish in international journals, thus Europe, Japan and Latin America frequently imitate the USA and deal with projects similar to those in that country rather than concentrating on the local ones that have a more direct influence on them. Thus, main health problems in the USA are devoted maximum attention and countless articles are written to solve problems of obesity, prostate cancer or senile dementia. Out of the 1300 pharmaceutical patents, only 13 were related to tropical diseases, therefore this situation is a clear drawback for underdeveloped countries.

In sum, as a result of the English hegemony, researchers only read American books and journals, and thus American ideas, concepts and reasoning. This is the case from the very early stages of scientists' academic experience; in fact, students at the university use coursebooks in English or books which have been translated from manuals that were originally written

in English, thus, with data frequently limited to the USA territory.

Some authors (Navarro, 2001:49) suggest acting against this international scene of predominance of a language and consequent discrimination of others because otherwise they consider to be accepting a secondary position in the world's science:

¿No sería mucho más lógico adoptar las medidas necesarias para poner fin a las actitudes discriminatorias y evitar su perpetuación? De lo contrario estamos dando por sentado que tanto la medicina hispanoamericana como la europea se conforman con ocupar indefinidamente una mediocre posición secundaria en el gran teatro de la ciencia mundial.

Also Fortanet (2002:20) is against consulting only Anglo-American bibliography and advocates checking first national bibliography and then European or from any other country in the world. Crystal (1997:12-13) also underlines the fears for the danger of a possible disappearance of minority languages as a consequence of a presence of a global language that could make people lazy about learning other languages, since the existence of a global language would make all other unnecessary. Actually, the process of dominance and loss of languages has been known to take place throughout history, but for Crystal this occurs independently of the emergence of a global language (1997:17).

In contrast to those who oppose a global language, there are authors who see nothing wrong with the idea of a global language and suggest that it would result in unity and peace as

linguistic misunderstanding would disappear. However, having a look at history it seems clear that a single language does not necessarily mean social harmony or mutual understanding. Evidence is represented by the Spanish or the American Civil War, The Vietnam War, etc. House (2001:73-89) also underlines the advantages of English as a “*lingua franca*” in Europe, for instance, the fact that it gives rise to a functional flexibility by its wide international extension and thus becomes a very useful tool of communication for different regions and identities that can have English as a common means of communication.

Crystal (1997:9-11) states that the strong need of a *lingua franca* for the whole world is something that has emerged since the 1950’s with the creation of the chief international forum for political communication, the United Nations, which dates from 1945. This body stems from the League of Nations that was created as part of the Treaty of Versailles in 1920. By that time it had 42 members, several from outside Europe. The UN was followed by other international bodies forming the UN system such as UNESCO and UNICEF in 1946, the World Health Organization in 1948, and the International Atomic Energy Agency in 1957, among others. All of them gave rise to many countries being represented in a single meeting-place. As a multi-way translation would result extremely expensive, the existence of a *lingua franca* to facilitate communication in these settings became vital. Usually there was more than one language designated as official. In the UN for instance, 5 languages were recognized: English, French, Spanish, Russian and Chinese. But it makes sense to cut down on the number and use a single

language, which is happening in meetings around the world as a result of the growing competence in the English language. In fact, English is the only official language in most international political assemblies all over the world (the Association of South-East Asian Nations, the Commonwealth, the Council of Europe, the European Union and the North Atlantic Treaty Organization, the Organization of Petroleum Exporting Countries, etc.). Besides, even in restricted membership meetings (e.g. one consisting of only Spanish-speaking states), where their proceedings are not held in English, the reports made at the end for the wider public and the statements for the media are in English (Crystal, 1997:78-79). In the Union of International Association's Yearbook figures showed that in 1995-1996, 85% of the international organizations in the world used English as official language followed by French with 49%. These numbers give evidence that regardless of the location in the world of an organization, English is the chief auxiliary language.

Also the international business communities need a global language to communicate in meetings among company directors coming from different countries to plan multinational deals.

Thus, the technology of modern communication, namely internet, and the technology of air transportation has given rise to the growth of international contacts and has led to the need of a global language in the 20th century.

Irrespectively of linguists' attitudes in favour or against the existence of a global language or international language, it is a fact that scientists necessarily need to be competent in writing

in English if they want to succeed nowadays as researchers (Alcaraz, 2000:15), especially medical doctors researchers because as pointed by Navarro (1994:298) “... el inglés es el idioma internacional de la medicina” (“...English is the international language of medicine”).

Matías-Guiu (1996:505) explains that, in order to obtain financial support for research or to gain professional prestige, researchers depend on the impact factor that the journal in which they publish their articles is given by the “Science Citation Index” (*SCI*). In 1961 computers made it easier to store and analyze information, which led to the foundation of the Institute for Scientific Information (further referred to as *ISI*) that was aimed at measuring rank of scientific outputs.

The most commonly used databases for studies of publications carried out by the *ISI* are the following: Social Sciences Citation Index (further referred to as *SSCI*), the Arts and Humanities Citation Index (*A and HCI*), and the Science Citation Index Expanded (*SCI Expanded*). In these indexes, articles with their references and other material in about 8600 journals are registered. The *ISI* database does not include all scientific journals. For instance, the *SSCI* covers articles in about 1700 social science journals plus selected articles from about 3300 science and technology journals. Nearly 2000 journals a year are reviewed but only 10 to 12% are sorted out to be included in the database, so the latter is frequently renewed. In fact, every two weeks journals are deleted and added to the database. The evaluation of the material carried out by the staff responsible for including journals are: the journal’s basic

publishing standards, its editorial content, the international diversity of its authorship, and the citation data associated with it.

Bearing in mind the aforementioned low percentage representing the number of articles accepted to be included in the ISI, it is observed that the evaluation is quite strict. As far as language is concerned, which is our main interest, emphasis is put on the fact that “English language article titles, abstracts and keywords are essential”, additionally “English language cited references” are also recommended (<http://www.isinet.com>).

Although this website states that “ISI seeks to cover the best regional journals as well”, they insist on the presence of English bibliography for those cases. In fact they consider it an essential component of the articles, otherwise they are excluded.

This constraint has led to the overrepresentation of English language journals (75% of the published articles are written by North American, British, Canadian and Australian authors).

A patent proof that follows these circumstances in the field of Medicine for instance is the appearance of *Medline*, an internet tool designed for searching for journals in a journal database. This tool restricts or limits its search to articles published after 1966 and to English by means of a restrictive criterion. Besides bibliographic data incorporate preferably English journals, and articles contained in these journals seldom have references in the bibliography to articles in other languages. Another way to illustrate the above mentioned issue

is that out of the 8023 bibliographical references included in the 400 articles published by British Medical Journal and *JAMA* between 1980 and 1995, only 40 were written in a non-English language.

As a consequence of the use of this tool, several authors (e.g. Claude Hagène, 1987:189, Timo-Ilaria, 1998:566-570) condemn the undue priority assigned to English authors. They complain about the tendency to read only current articles published in English that leads to the rediscovery of data and ideas previously published in other languages. To give an example, the *Nature Medicine* journal published an article which dealt with the appearance of a new muscle in May, 1996, which actually had been discovered over a hundred years before by German and French researchers (Poirier and Charpy, 1901; Zenker, 1955:355-368 and Henle, 1971).

Another example of the restriction of bibliographical research to the English language is that of the journal dated from 29 of April, 1995 called *Evidence Based Medicine*. The journal is published every two months; according to its publishers, its goal is “to publish the gold that intellectually intense processes will mine from the more of 100 of the world’s top journals”. The interesting part is that the words mentioned in the quotation “the world’s” only referred to articles written in English.

2.1. The Research Article

While mainstream research in linguistics in the United States has focussed on more cognitive aspects of language, British researchers have taken a more sociolinguistic approach (for example, Halliday's systemic linguistics). Within the British tradition, researchers analyzing scientific texts during the 1980s started to focus on the concepts of "genre" and "discourse community".

2.1.1. Genre and Discourse Community

Defining genres may be problematic, but even if theorists were to abandon the concept, in everyday life people would continue to categorize texts. At a slightly simplistic level, one can define genres as realisations of social actions, in other words, "how things get done, when language is used to accomplish them" (Martin, 1985:250).

Adding to this definition, genres can be described as referring to a conventional category of discourse based on large scale typification of rhetorical action. An action acquires meaning from the situation and from the social context in which that situation arose. Honing in on the kind of definition needed for the research of scientific discourse that is to be presented here, Swales has defined *genre* and the concept of *discourse community* as follows:

A genre comprises a class of communicative events, the members of which share some set of communicative purposes.

These purposes are recognized by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. The rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style (1990:58)

Genre analysis is then the study of how language is used within a particular social setting. Examples of genres in academic written English are research articles or papers, abstracts, theses and dissertations. It is the discourse community that determines the characteristics of a certain genre and establishes standardized criteria or schematic structures.

Clearly, there is a close connection between the two terms: genre and discourse community. To illustrate this link, Swales (1990:9) explains that discourse communities are socio-rhetorical networks that are created in order to achieve common objectives. The members of this community are familiarized with the particular genres that are used in order to achieve the mentioned aims. Therefore, genres belong to discourse communities and not to individuals or other kinds of groupings or to wider speech communities.

In his book “*Genre Analysis*” (1990), Swales proposes several characteristics to define a group of people as a discourse community.

According to this author, a discourse community:

- has a broadly agreed set of common public goals
- has mechanisms of intercommunication among its members
- uses its participatory mechanisms primarily to provide information and feedback

- utilizes and hence possesses one or more genres in the communicative utterance of its aims
- has acquired some specific lexis (specialized terminology, acronyms)
- has a threshold level of members with a suitable degree of relevant content and discursal expertise.

The basic definitions of genre and discourse community used in this research work are based on Swales as a recognized authority in the field.

However, Bathia (1993) considers Swales' definitions on genre incomplete. According to Swales, genres are communicative events and the communicative purpose of a genre shapes the genre and provides it with an internal structure. Nevertheless, experts in a professional community may not always agree about the purpose of a genre, thus assigning communicative purpose to a text is not always an easy task. Based on this idea, Bathia widens Swales' definition of genre in the following manner:

It is a recognizable communicative event characterized by a set of communicative purpose(s) identified and understood by the members of the professional or academic community in which it regularly occurs. Most often it is highly structured and conventionalized with constraints on allowable contributions in terms of their intent, positioning, form and functional value. These constraints, however, are often exploited by the expert members of the discourse community to achieve private intentions within the framework of socially recognized purpose/s. (Bathia, 1993:13)

With his definition, Bathia underlines the fact that Swales does not take into account “psychological factors” and “tactical aspects” of genre construction. In the last sentence, the author claims that some purposes might probably be recognized by expert members but are not likely to be officially acknowledged by the institution. To exemplify the idea of possible tactics used, which may mislead the recipient, Bathia (as cited in Askehave and Swales, 2001:199), names skilful news reporters, who may convey their own political point of view under an apparent objective broadcast.

Genre analysis has devoted much attention to Research Articles (further referred to as RAs), which should be considered a kind of genre. In fact, they consist of a communication act in which its participants, namely, researchers, share a common interest of conveying their scientific findings. This purpose leads these experts, i.e. the discourse community, to determine the characteristics of that genre in the sense that they establish standardized criteria or schematic structures. They use a specific kind of language, style, etc. within this particular setting.

Duque García (1999:40) affirms that effective writing or good style varies from one discourse community to another, i.e. it is necessary to know the scientific, linguistic and stylistic criteria of the community in question to be able to produce good style. Therefore, it is highly recommended that the student learns the genre and conventions that characterize the community they want to belong to as future professionals, in order to use this particular style. For this reason, detailed

information about the origin and characteristics of scientific writing and of RAs is going to be dealt with in the next section.

2.1.2. The evolution of Research Articles

For Day (1998:5-6), scientific communication appeared with the first scientific journals: *Journal des Scavans* in France, *The Philosophical Transactions of the Royal Society* in London, both dated from 1665 and *The Acta Eruditorum* in Germany, from 1682. Duque García (1999:12-14) adds that these publications were extremely important for creating a discourse community, as they were the means of disseminating scientific findings. At that time, findings had no value until they were made known. In order to reach a broad audience, the Royal Society supported a clear and simple style.

According to Gross (2002:13), the scientific article was a deliberate invention that took place during the 17th century. It was created in England and France by Henry Oldenburg and Denis de Sallo. The scientific article has evolved and changed with time; in fact in 1920 Fulcher suggested the inclusion of an abstract for manuscripts in the *Astrophysical Journal*. This should imply time-saving for researchers that would get the information on the manuscripts faster. His idea was accepted and is still recognized today in scientific articles.

The first pieces of scientific writing published were presented in a letter format with the typical features of this kind of genre. One outstanding characteristic was the use of the

personal pronoun “I”, which is in accordance with the fact that the observer’s role was more important than that of the writer. This is because by that time findings were handmade instruments; thus, personal skill in dealing with new techniques was fundamental. By means of this brief reports scientists “seek to establish the author’s credibility more by means of reliable testimony than by descriptive accuracy, more by qualitative experience than by quantitative experiment” (Gross, 2002:229).

According to Lannon (1982:3), “in technical writing you report factual information objectively for the practical use of your readers”. Therefore, papers published in the early journals showed a descriptive style. However, as Gross (2002:229) states, style evolves with time moving from a personal style in the 17th century to a very impersonal in the 19th century. Nevertheless, many other changes can be observed regarding style through the above mentioned centuries:

- there is a change from a verbal to a more nominal style; complex noun phrases tend to occur in the subject position.
- Syntax becomes simpler through shorter sentences
- texts contain more “presentational features” that help the reader understand complex information

Stylistic features are given more uniformity in the 20th century as a consequence of the scientific article growing across national boundaries and “scientific English” becoming the

international discourse of science (Gross, 2002:230). The most relevant features in the 20th century in terms of style are: short sentences, simple syntax, complex noun phrases with multiple modification, presence of passive voice, technical abbreviations, quantitative expressions and equations, and citational traces.

This specialized style corresponds to the new audience in the 20th century, which varies from that of previous centuries in the sense that scientific articles are addressed to the discourse community, namely, to other professionals sharing the same specialty or researching in the same discipline.

Regarding the content of the texts, the mere observation of facts in the 17th century where only claims of facts were stated as a result turn into specific explanatory structures in the different disciplines during the 18th century (e.g. “geological explanations”, “chemical explanations”). This is a consequence of the emergence of special-interest journals and societies, new research institutes, etc. This tendency to specific explanatory claims to specific sciences is consolidated by the 20th century. Gross (2002:231), points out other characteristics of scientific articles of this century like the presentation of large data sets, the application of mathematics and support of texts by visual components, i.e. information is often presented in form of tables and figures.

The rapid advance in science in the 19th century and the importance given to the role of the reader encouraged scientific writers to put emphasis on the reporting of methodology.

The work of Louis Pasteur in the second half of the 20th century is the first example. In order to defend his findings from

critics, Pasteur was known to provide very detailed reports of his experiments. This enabled his peers to reproduce them and originated the “*principle of reproducibility*” in science.

As a result of this, the Methods section of articles was described with exquisite detail and thus led to the highly structured Introduction-Method-Results-Discussion format of Research Articles (further referred to as IMRD). This format has proved to present a logical model that makes it easier for scientists to organize their papers, as well as for editors and readers to understand them. The letters stand for the sections that make up an article: Introduction (What problem was studied?), Method (How was the problem studied?), Results (What were the findings?) and Discussion (What do these findings mean?) (Duque García, 1999: 83).

Findings such as penicillin in 1929 and other drugs that eradicated serious diseases (e.g. tuberculosis, diphtheria, etc.) after World War II gave rise to dramatic investments in research by the USA. Consequently, a large amount of papers were written and editors became more demanding asking for tightly written and well-organized papers so as to save expense and space in journals. This is how the IMRD format came into universal use in research journals.

2.1.3. The Structure of RAs: The Introduction, Method, Results and Discussion model (I-M-R-D model)

This organizational patterning in Research Article sections has continued up to the present and has developed into the prototypical I-M-R-D model. This model has been readily accepted and taken for granted by most researchers in their articles (Brett, 1994; Holmes, 1997; Stanley, 1984; Swales, 1990).

In the *Introduction*, researchers typically make topic generalizations, review items of previous research, mention their findings and justify the development of their study. In the *Method* section, they explain the procedure of their study. In the *Result* section, analysts convey results. Finally, in the *Discussion* section they focus on “commenting results” by interpreting, accounting for, evaluating or comparing with previous work (Yang Ruiying, Desmond Allison, 2003: 380).

The I-M-R-D model was later on completed with the analysis of Moves and steps of each Section of the research article. In this sense, Swales introduces the terms “move” and “step”, which can be defined according not only to Swales, but also to Dudley-Evans, Holmes and others, as follows: Moves are made up of steps. Sometimes there is only one step in a move, but a move can be divided into a number of steps. The concept of *move* captures the function and purpose of a segment of text at a more general level, while *step* spells out more specifically the rhetorical means of realizing the function of the move. The

set of steps for a move is the set of rhetorical choices most commonly available to RA authors to fulfil a certain purpose.

These different sections have been studied by different authors. For instance, Thompson, 1993, Brett, 1994, Nwogu, 1997, Yang and Edwards, 1995 and Posteguillo, 1999 worked on the Result section; Holmes, 1997, Hopkins and Dudley Evans, 1988 and Swales and Feak, 1994 on the Discussion section; and Bruce, 1983 and Weissberg 1984, on the Methods section. With regard to Introduction sections, Crookes 1986, Jakoby 1987 or Anthony, 1999 are some of the most important researchers, but one of the most prominent studies is that of Swales.

Swales' original work (1981) on article Introductions offered 4 basic moves:

- Move 1: *Establishing the Field*
- Move 2: *Summarizing Previous Research*
- Move 3: *Preparing for Present Research* (often by identifying a gap in previous research)
- Move 4: *Introducing Present Research*

In the introductions researched by Swales, all the moves are present in most of them (75%). According to Swales (1990), in most genres, moves will be either obligatory or optional; they may be in a fixed or variable sequence, they may be subject to embedding one within the other, and they may be recursive.

However, this compulsory presence and ordering does not apply to the Discussion section studied by Hopkins and Dudley-Evans (1988).

These authors suggested the following moves as characteristic of the Discussion section:

1. *Background Information*
2. *Statement of results*
3. *(Un)expected result*
4. *Reference to previous research (comparison)*
5. *Explanation of Unsatisfactory Result*
6. *Exemplification*
7. *Deduction and Hypothesis*
8. *Reference to Previous Research*
9. *Recommendation*
10. *Justification*

They claim that in this section moves occur in cycles in which an appropriate sequence of moves is selected. Consequently, they do not divide moves into obligatory and optional but rather it is the selection of a certain move that turns the following into an obligatory one. Namely, once the writer chooses a certain move he is obliged to follow with other specified moves. For instance, move number 3, “*Unexpected Result*”, would normally be followed by move number 5, “*Explanation of Unsatisfactory Result*”, so as to justify the previous move.

Swales’ original idea in his first model (1981) that moves were text elements considered obligatory if the text is to be accepted as an example of a genre was rejected in his adapted model named CARS (Create a Research Space in 1990).

Move 1: Establishing a territory	step 1 claiming centrality step 2 making topic generalization step 3 reviewing items of previous research
Move 2: Establishing a niche	step 1 counter-claiming step 2 indicating a gap step 3 question-raising step 4 continuing a tradition
Move 3: Occupying the niche	step 1 outlining purposes or announcing present research step 2 announcing principal findings step 3 indicating RA structure

Table 2.1: Swales' model of moves in Introduction Sections (1990, 141).

He now suggests that schematic structures are prototypes which can be subject to different amounts of variation according to the degree to which the genre is conventionalized. Swales also recognizes the difficulty in applying his original 4-move model to Introductions.

Swales' 1980s model received much criticism for concentrating exclusively on the text content while disregarding the writer's rhetorical strategies when formulating their arguments (Bazerman, 1989). Consequently, Swales' 1990 model (as cited in Dudley-Evans, 1997:150-159) pays attention to the writer's rhetorical or social purposes when structuring and selecting words for their introductions. In fact, "his categories appear rather more sociological than linguistic".

As far as the conventional structure of RAs is concerned, Ciapuscio and Otañi (2002) highlight that the conventional IMRD model evolves with the decades and talk about the possibility that the RA structure might have changed their in the decade of 2000.

Thus for example, Yang Ruiying and Desmond Allison in their article *Research articles in applied linguistics: moving from results to conclusions* carried out a study of the structure of RAs in applied linguistics paying attention to 4 sections: Results, Discussion, Conclusion and Pedagogic Implications. They reported on the observed similarities and differences between them. Their study was based on 20 RAs reporting empirical research. In the Result section, they identified 3 dominant moves: “Preparatory information”, “Reporting Results” and “Commenting on Results”. And secondary moves. 4 to 6 were: “Summarizing results”, “Evaluating the study” and “Deductions from the Research”.

In the Discussion section, they found a very similar structure. However, “Commenting on Results”, unlike in the Results section, was the most frequent move. In this way, they showed that the moves in these two sections overlap considerably.

As for the “Pedagogic Implications” section, these researchers also observed overlapping with the Discussion and Conclusion sections except that the occurrence of the step “Dealing with pedagogic issues” within the move “Dealing with pedagogic issues” occurs in this section at a higher average rate than in the other two sections. There are 4 moves in this section: “Summarizing the study”, “Dealing with pedagogic issues”, “Evaluating the study” and “Deductions from the research”.

According to Yang Ruiying and Desmond Allison (2003:381), the overlapping of the moves in the studied sections of applied linguistics justifies the fact that 3 of the mentioned

sections can function as the closing section of RAs. However differences between the sections lie in their communicative purposes and this explains the appearance of different section headings. (Yang Ruiying and Desmond Allison, 2003:381).

It is worth mentioning that some authors such as Posteguillo (1999) claim that “Discussion and Conclusion are typically presented as alternative equivalent sections”. Yang Ruiying and Desmond Allison (2003:368) argue that there is no evidence in the literature to justify this and that this cannot be always the case for in many occasions both sections appear in the same article as two separate headings.

2.1.4. Adaptation of the conventional I-M-R-D model and the analysis of the “moves” to the medical corpora

In the field of medical RAs Conclusions, the model can still be applied in the present time, in fact, most journal guidelines require this model, and most of them did not ask for a Conclusion Section, rather, this was left to the researcher’s choice. Therefore, most times Conclusion Sections were included in the Discussion Section and occasionally Conclusion stood under its own heading. This accounts for the difficult task of collecting a corpus containing a separated Conclusion Section.

The analysis of the corpora of the present research work intended to follow initially the model created in the research “*A genre rhetorical and lexico-grammatical analysis of a corpus of academic articles*” dated from 2004 by Ricart. The original idea

of identifying the 5 moves in Conclusion Sections of medical RAs was not easy as the structure of medical RAs was not so clearly divided into these moves. In fact, many times the Conclusion Section just provides the reader with a short report in the form of a summary of the paper. This means, it was rare to find Conclusion Sections containing all the 5 moves. 80% of the Conclusions presented from one to two or 3 moves, but only in 20% of them could all the moves be identified. The least common move was that of *Advantage*, in fact, it was only present in 15% of the corpora.

The low rate of appearance of the move *Advantage* led to the decision of leaving it out of the analysis and thus reducing the initial model containing 5 moves to a new model with 4 moves, namely:

- a) *Background*
- b) *Summarizing the study*
- c) *Evaluating the study: indicating disadvantage/limitation*
- d) *Indicating Further Study*

Relevant information regarding the moves is the presence of overlaps. The most common overlap was that of *Advantage* and *Limitation*. This fact also accounts for the decision of discarding the move *Advantage*. Many times the move *Advantage* included a disadvantage as in the example “*Los resultados han sido buenos a pesar del alto índice de complicaciones*”. Additionally, words with inherent negative connotation were the most common when expressing an

advantage, in fact, researchers use them in combination with other words that make them end up losing their semantic value and expressing the opposite. This manipulation of the words explains the frequent overlap of the moves *Advantage* and *Limitation*.

Other overlaps take place between *Background* and *Limitation* and between *Limitation* and *Further Research*. Regarding the first overlap, it is common to mention an existing problem up to the present that is going to be dealt with in the paper, this fact accounts for the overlap *Background* and *Limitation*. As for the second overlap, *Limitation* and *Further Research*, it occurs because researchers tend to inform on a limitation of the study and the consequent need for more investigation to solve the problem.

It can be concluded that the move *Limitation* is the most typical move in the conclusion of medical RAs, and that it is the move that overlaps with the rest of the moves in the analysis. Medical doctors normally emphasize the problematic aspects and underline the need for more investigation in their conclusions rather than the positive or the achievements.

2.1.5. Contrastive analysis of RA Conclusions

In a recent study, Ciapuscio (2002:117-133) and Otañi performed a contrastive analysis in order to compare the Section of Conclusions in RAs. Although their study is mainly focused on a Spanish-German comparison, they also devote part of their

article to English. To carry out their analysis, they based their study on Gnutzman and Oldenburg's model, who in 1991 suggested a rhetoric-linguistic analysis for the introduction and conclusion sections in RAs. By means of this study Gnutzman and Oldenburg intended to improve German writer's RAs in English. They thought that making them aware of rhetoric and linguistic differences in RAs would result in a better production of the English RAs on the part of the German writers.

Similar to the research carried out in 2004 (Ricart) on genre analysis of RAs, Gnutzman and Oldenburg also make use of Swales' model to focus on the Conclusion section of RAs using the journal "Language" as their corpus. They identified 5 moves:

- a) summarizing the author's results
- b) summarizing previous research
- c) strengths and weaknesses of the investigation
- d) open questions and probable solutions
- e) assessment and implications of the author's results

Gnutzman and Oldenburg carry out research on the linguistic realizations of the different moves in Conclusion sections in the journal "Language" and in an equivalent German journal called *Linguistische Berichte*.

According to the above mentioned authors, there is evidence through many studies following theirs that scientists are under the influence of linguistic-cultural factors in their writing productions. However, Gnutzman and Oldenburg affirm that, given the limited corpora used for the analysis and the differences regarding the characteristics of the various

disciplines, these statements' validity should be questioned. Nevertheless, they suggest that some conclusions are common to most of them and therefore relevant. To illustrate this with an example, they mention a news article written by Kaiser in which he states that Latin-Americans judge Germans' text productions as impersonal and boring whilst the latter consider Latin-American texts subjective and personal.

In order to carry out their study Ciapuscio and Otañi take 23 RAs Conclusion Sections of 3 Spanish journals on Linguistics: *Signo y Señal*, *Revista de Lingüística Teórica y Aplicada* and *Revista Argentina de Lingüística*. Among the conclusions they draw from their analysis, they indicate that the presence of the first move (summarizing the author's results) is highly frequent in the 3 languages. However, their linguistic realizations differ. In German, writers adopt an objective and impersonal style, resorting to grammatical resources such as the passive voice. On the contrary, Spanish researchers prefer a more personal style using the first person. Although the plural is preferred, they also use the first person singular. Nevertheless, Spanish texts also show the presence of other elements such as the passive construction with "se" and impersonal sentences which turn the previous personal style into a less personal one. A significant and frequent feature that contributes to making the text less subjective is the use of a noun in the position of the subject or new agent in the passive structure. E.g. "El análisis permite", "la investigación ha demostrado", "los datos evidenciaron", etc.

Regarding other aspects in this move, it is relevant to note that the verbs are all in the indicative and the preferred tense is the present perfect. The authors also mention the importance of hedges and the modals “poder” and “deber” in this move.

The high rate of appearance of the first move in German, Spanish and English, contrasts with the scarce appearance of move number 4 (open questions) in all of them. Most differences were mainly found in moves number two (summarizing previous research) and 5 (assessment and implications of the author’s results). These moves appear frequently in English. This is not so in German and Spanish. German and Spanish scientists often refer to other colleagues as a base for their study or to support or confirm their own work. In contrast, English scientists tend to refer to their peers’ previous work to challenge their statements. That is, they consider their articles have given rise to the need for further research or they even judge them as being to some extent unsatisfactory.

As for move number 3, strengths and weaknesses of the study, according to the authors, it appears that Spanish-speakers have the need to prevent their peers from criticizing their work. For this reason they anticipate the limitations of the study by means of this move, whereas in the works of English scientists this move is not so relevant and in those written by German authors it does not appear.

The above mentioned information on move number 3 does not coincide with the results obtained on genre analysis of RAs (Ricart, 2004). This investigation showed the high presence

of the moves “limitation” in Conclusion Sections of RAs in English. This mismatch of information can be due to two factors. First, the different disciplines studied for the analysis. Their articles belonged to Linguistics whereas the articles selected in 2004 were all scientific articles related to Biotechnology, Robotics, etc. On the other hand, their articles dated from 1993 to 2000, while the articles in the investigation in 2004 dated from 2000 to 2003. It is worth pointing out, though, that Ciapuscio and Otañi mention the evolution that RAs experience with time. In fact, they also compare Spanish RAs in the 80s with Spanish RAs in the 90s and affirm that there has been a tendency for Spanish RAs to adopt what they call an “ABE” rhetorical structure. These letters stand for moves one, two and 5 in their article “*Las conclusiones de los artículos de investigación desde una perspectiva contrastiva*”. Based on this tendency, these authors state that Spanish RAs are taking up the Anglo-Saxon structure (possibly as a consequence of scientific globalization). Therefore, likewise, it is not surprising that English RAs might equally have changed in the decade of 2000.

As regards the last move, number 5, assessment and implications of the author’s results, Ciapuscio and Otañi affirm that in Spanish it presents mainly a subjective tone with the use of first person suffixes in the verbs. They also highlight the optimistic tone and mitigations that intend to stimulate a positive attitude towards the text on the part of the receiver.

Another recent study in this field in connection with RAs is Martin’s research on abstracts in the area of experimental social sciences. His 2003 study is oriented to help Spanish

academics write their abstracts according to the international scientific community's expectations. He also wants to demonstrate that scientific discourse is subjected to socio-cultural factors. He proves that most Spanish abstracts follow the international convention, that is, the English structure. This structure consists of 4 sections: Introduction, Methods, Results and Conclusion. However, they show a summarized version, since only 25% present the 4 sections. This is due to the frequent omission of the Results section in Spanish.

After considering the structure of a RA, stylistic features are going to be considered in the following section.

2.2. Academic Writing in Research Articles

According to Duque García (1999:35), the style of a text depends on a number of factors such as the author's personality, likes, age and nationality, knowledge about the language, genre and stylistic norms they are using, the journal in which they intend their piece of writing to be published, the audience, and the period in which they are writing.

Duque García (1999:15-18) provides us with several definitions for technical writing taking into account varying aspects. Some authors focus on the audience, others on the context, the pragmatic purpose, the style, etc. Among the definitions that underline the importance of the audience it makes sense to comment on a quite recent one:

Technical writing conveys information about a technical subject to a specific audience for a specific purpose (...) that is, to communicate a body of factual information that will help an audience understand a subject or carry out a task. (Markel, 1992: 2)

Duque García highlights Mathes and Stevenson's (1976:3) early definition of technical report as one of the most complete definitions, as they consider the writer, text, audience and context: "The technical report is an act of communication by a professional in an organized system to transfer information necessary for the system to continue to function".

Finally, taking all the aspects mentioned above into account she comes up with her own definition:

La escritura científico técnica es un acto de comunicación dinámico que implica una interacción humana entre científicos (...) que pertenecen a una misma comunidad científica. En este proceso comunicativo el escritor transmite ideas, información y descubrimientos inéditos de manera clara, objetiva, concisa y precisa a una audiencia especializada en su mismo campo científico a través de artículos que se publican periódicamente en revistas también especializadas.

Para que esta transmisión sea efectiva, el escritor debe ajustarse a las necesidades de su audiencia, seleccionando el tipo de lenguaje que puede entender (...)

In her definition she also emphasizes the need on the part of the writer in a technical report to be aware of the factors that influence the context in which the interaction takes place: the different kinds of articles and journals with their norms for

publication, the role played by publishers and referees with their expectations, the professional audience they are addressing, etc.

From this complete definition there two aspects should be focused on. On the one hand, for Day (1998:8) and Duque García (1999:44) a scientific paper is a written report describing original research results, i.e. unpublished, that has to be accepted for publication. These results also have to add “some new understanding, observations, proofs” (Hengl, 2002:1). Thus, regarding technical style, it is not surprising that most authors agree on its main features: clarity, concision, accuracy and objectivity (Day, 1992). Other authors also support transparency (Markel, 1992), content organization, legibility or accessibility, grammatical correctness (Markel, 1992) and lexico-grammatical elements (Halliday, 1993) among others. Fortanet (2002:24-26) underlines the importance of cohesion and coherence at a textual level. The researcher has to plan in advance the organization of the text. Each paragraph has to deal with one argument and all the paragraphs have to form a unity to reach cohesion. In order to succeed in terms of coherence, it is necessary to link the paragraphs in a logical manner. To this end the scientific writer can use different strategies such as repetition of key words, use of pronouns and synonyms, link words, conjunctions, etc. All these expressions help the reader understand the text more easily. Finally, she highlights the importance of avoiding ambiguity and the relevant role of what she calls the “antecedent” or word that is referred to by means of a pronoun, for example. It is utmost important that the writer makes it easy and clear for the reader to localize the antecedent. Another

aspect that Fortanet comments on (2002:25-26) is that of the passive voice. Sometimes, its use is crucial to make the text sound objective. It mainly appears in the methodology section to describe methods used and it is frequently present when the investigation has several authors. However, it should not be overused. For Fortanet, its use should be one fourth of the instances in which it could be used. She claims that it is also very convenient to use active voice, which also has many advantages: it uses fewer words, it is more direct and it gives the idea that the author was the performer of the investigation. In Fortanet's mind, this latter aspect should not be considered immodest on the part of the writer, for he is actually the author of the paper.

2.2.1. Main features of Research Articles

Similarly to Fortanet, Hengl (2002:1) favours a logical, coherent, focused and well-argued writing to ensure that the paper will be not only published but also read. For Hengl (2002:7) a good article is not just one that gets published in a top journal but one that is also read and cited. It is the recognition and consideration of other peers that makes it valuable. This leads us to the second aspect to be focused on, the necessity that the paper be published in a "primary journal" or other primary publication.

Among the many definitions of valid publication, i.e. primary publication, Day (1998:9) underlines the one by The Council of Biology Editors (CBE):

An acceptable primary scientific publication must be the first disclosure containing sufficient information to enable peers (1) to assess observations, (2) to repeat experiments, and (3) to evaluate intellectual processes; moreover, it must be susceptible to sensory perception, essentially permanent, available to the scientific community without restriction, and available for regular screening by one or more of the major recognized secondary services (e.g. Currently, Biological Abstracts, Chemical Abstracts, Index Medicus, ...). Day explains that “susceptible to sensory perception” includes not only written text but material in nonprinting or non-visual forms such as publication in the form of audio cassettes, current electronic journals, etc.

There has been controversy regarding whether material posted on a Web site should be considered as “*publication*”, but the American Society for Microbiology clarifies in 1998 that this is neither permanent nor available to the scientific community without restriction and therefore, primary publication must be in a journal readily available within the scientific community:

A scientific paper published or its substance published in a conference report, symposium proceeding, or technical bulletin, posted on a host computer to which there is access via the Internet, or made available through any other retrievable source, including CD-ROM and other electronic forms, is unacceptable for submission to an ASM journal on

the grounds of *prior publication*. A manuscript whose substance was included in a thesis or dissertation posted on a host computer to which there is access via the Internet is unacceptable for submission to an ASM journal on the grounds of *prior publication*.

Hengl (2002:6) provides us in his article *Rules of thumb for writing research articles* with what he calls “golden rules for easier publishing”

- a) write for your audience
- b) keep a clear focus in the paper and present only results that relate to it
- c) be yourself: write like you speak and then revise and polish
- d) make it simple: use simple examples to explain complex methodology
- e) make it concrete: use concrete words, strong verbs, avoid noun clusters abstract and ambiguous words
- f) make it short: avoid redundancy, repetition and over-explanation of familiar techniques and terminology
- g) take responsibility: make a clear distinction between your work and that of others
- h) make strong statements (e.g. “We conclude” instead of “It may be concluded”)
- i) Be self critical; consider uncertainty of conclusions and their implications.

Regarding the form of the article, Fortanet (2002:17) states that most professional journals do publish stylistic norms to be considered if the researcher wants his article to be accepted. She adds that there are also different manuals for the different scientific communities they address. For instance for humanities there is the MLA style guide (1995), the CBE manual style (1978) for the field of biology, the APA style guide (1994), the Chicago manual of style (1993), the work of E.J. Huth (1999), for biomedical sciences, etc. Duque García (1999: 147-148) states that many professional organizations publishing scientific journals have devoted attention to the creation of general guidelines and rules in order to provide researchers with information about how to write their RAs. (Some Professional Organizations are: The Institute of Electrical Engineering (IEEE), The American National Standards Institute (ANSI), The American Institute of Physics (AIP), etc.)

These guides contain norms regarding the format, the organization of the document, appropriate use of references, abbreviations, etc. However, most of them do not deal with the aspect of style and therefore do not explain how to achieve the most important aspect of scientific writing: concision, accuracy, clarity, simplicity. What are the lexico-syntactical elements to be used in order to have a good style? The only guides with some information about this are: *The American Institute of Physics Style Manual* and *The American National Standard for the Preparation of Scientific Papers for Written and Oral Presentations*. The first manual (1990:12-19) deals with style in

a general way: grammar, punctuation, the use of numbers, brackets, spelling, plural, capital letters, personal pronouns, etc. The second one (1979-1989) provides more detailed guidelines such as: be logical, clear, precise, direct and concise, use a simple language, substitute verbs for abstract nouns, do not modify nouns by more than two other nouns, proper use of passive and active voice, etc. One of the norms that link us to a topic of our concern is the following:

Authors whose English is not completely fluent and idiomatic should make every effort, by consultation with colleagues who write correctly and well, to present the report in an acceptable form. It is unreasonable to expect either editors or referees to rewrite the paper in order to correct inadequacies of language.

Hengl (2002:3) provides us with a list, which comprises the most important reasons for rejection of RAs:

- a) irrelevant topic or topic of local interest only
- b) the RA offers nothing new
- c) lack of focus: the topic, objectives and conclusions are not connected
- d) unclear and misleading argumentation
- e) weak methodology or results
- f) unclear and unfocused style and incoherent text
- g) insufficient data quality: flawed design; insignificant sample number; preliminary findings only

Hengl highlights the most common reasons for L2 articles being rejected: incorrect grammar, incoherent development of the topic in the paragraphs and unclear text.

Gosden (1995:45) carried out an investigation in 1992 indicating that 74% of native speaker science journal editors surveyed indicated non-native speakers' poor reporting as being the main factor in diminishing the importance and quality of their research. He acknowledges the need for non-native speaker novices to be assisted with revisions and polishing of their writings and suggests they should be trained by means of courses on academic writing skills. Despite the non-native speaker novice researchers being at disadvantage regarding publication, Gosden (1995:46) says that "initiation into the international academic community through publications is common for all novice researchers around the world, irrespective of L1 origin". Consequently he is in favor of researchers' awareness of "the hard, norm-developing processes of RA drafting, feedback, negotiation, and redrafting" which he considers indispensable for publication success. A study carried out by Dudley-Evans (1991:41-51) shows that the mistakes made by native speakers of English and non-native speakers writing their PhD in biology were the same type, the only difference between them being a more significant quantity by the non-native. Despite the fact that all researchers have to learn how to write scientifically, English speakers find less difficulty because there are aspects of genre of specifically English style.

2.2.2. Linguistic features of medical Research Articles

Chapter one presented the need to command the English language with certain proficiency on the part of scientific researchers if they have the intention of reading about new advances and publish, since "... doctors around the world basically communicate in a single language: English" (Piqué and Posteguillo, 2006:651).

The fact that medicine appeared centuries ago explains that "medical genres have become stable in their form, structure and style" (Piqué and Posteguillo, 2006:651). This explains for example the fact that in the present time, RAs in medicine tend to follow the IMRD pattern, which has also been adopted by other scientific disciplines.

According to the above mentioned authors, medical communicative events can be divided into many written genres: editorials, research articles, abstracts, case reports, review articles, peer reviews, replies to these reviews, letters of acceptance/rejection of a paper, conference programs, medical popularizations, letters of application, book reviews, and letters to the editor. Having a look at previous literature, Piqué and Posteguillo (2006:383) explain that several linguists have dealt with medical language in the different above genres. Magnet and Carnet analyzed Letters to the Editor in 2001 and 2002, Vázquez and del Árbol, Letters to the Editor and Editorials in 2005, Chubin and Hackett focused on Peer review in 1990, and Salager studied academic conflict from a cross-cultural and

cross-generic point of view in case reports, research papers, review articles and Editorials in 2001, 2003 and 2005.

In their article *Peer Positive and Negative Assessment in Medical English Written Genres* (2006), Piqué and Posteguillo examine variation among medical written genres regarding positive and negative assessment. They focus on reporting verbs and use WordSmith™ Tools for the analysis. In the genre of research papers, they come to the conclusion that book reviews are the most negative and case reports the most positive genres regarding the percentage of positive reporting verbs. As for research papers, positive reporting items are much higher than negative, namely 42.48 and 4.03 respectively. To account for this fact, these authors mention Salager (2005:287) who explain the necessity on the part of researchers to be tactful so as to avoid a face-threatening critique on their peer's work. The authors analyzed the IMRAD model in detail and state that the use of reporting expressions whether positive or negative is not consistent across different sections of Research Papers and that it is the Discussion section where reporting verbs are more frequent, but they do not analyze the Conclusion section. Among the different reporting verbs they highlight "show" and "develop", which appear the most. Piqué and Posteguillo suggest that further research could be done into other genres with clearly defined sections to see if reporting verbs also vary depending on the sections or not.

Studies on the idea that language follows certain patterns that tend to recur in the form of sequences of words have been

undertaken by many linguists (Altenberg 1993; Moon, 1994, etc.).

Biber revised scientific discourse and came to the conclusion that the main features are “very frequent occurrences of nouns, long words, prepositions, conjuncts, agentless and by-passives, past participial adverbial clauses and markedly infrequent occurrences of private verbs, contractions and that-deletions” (Biber, 1988).

A few years later, Lemke (1990:439) explains that two relevant characteristics of scientific discourse are the use of the agentless passives and of nominalizations. In fact, scientists frequently omit the agent to avoid in this way “authorial personality” and give the impression that it is an objective description of facts that it is being dealt with.

Lemke (1990:440) also mentions the existence of nominalization in scientific discourse as an example of another important feature of this kind of discourse, namely, what he calls “condensation”.

Luzón carried out a study in 2000 in which she focused on collocational frameworks in medical research papers. With this aim, she used a corpus of 100 medical papers taken from two leading journals in the discipline of medicine: *the New England Journal of Medicine* and *the British Medical Journal*. By means of computer corpus analysis Luzón (2000:67) found that the 3 most frequent frameworks in the corpus were: *the... of*, *be ... to* and *a... of*. These 3 frames appeared with 1150, 81 and 98 different collocates respectively. Some examples of frequent collocates or fillers of the first framework, namely “*the... of*”,

which is the most frequent, would be: “*the number of*” or “*the risk of*”. She highlights the triplet “*the start of*” as the highest collocation in medical papers since “*start*” appears only forming part of this triplet in the whole corpus. This would contrast with other fillers such as “*risk*”, which appears in the corpus more often than “*start*”, in fact, there are 506 occurrences of the former and only 35 of the latter. However, the filler “*risk*” only appears 150 times in the triplet “*the risk of*”, representing a total of 29.8% of the total occurrences of this collocate. Despite the number of occurrences being higher than that of the collocate “*start*”, the percentage of the triplet “*risk*” 29.8 % is much lower than that of *start*, which is a 100%. Therefore it can be concluded that the collocation with “*risk*” is more frequent but “*start*” is more significant in the medical paper.

Another interesting feature Luzón (2000:69) adds taking into account both the most frequent and the most significant collocates is that in the mentioned framework “*the... of*” all collocates indicate measure or quantification (e.g. “amount”, “degree”, “dose”, “incidence”, etc). Additionally, the triplet is followed by nouns in the research that need to be quantified such as illness, drugs, patients, etc (e.g. “*the number of striol-treated women*”). As for the words preceding the triplet, she mentions the appearance of reporting verbs in combination with the subject *table* and *figure* (e.g. “*Table 1 shows the number of...*”). A second option occupying the place in front of this collocate is that of verbs or nominalizations expressing changes in quantity (e.g. “*increase in the incidence of*”). The last option

is a verb conveying the aim of the research (e.g. “*determine the number of*”).

Luzón also highlights the appearance of the triplet in nominalizations (e.g. “*the use of*”) or in combination with the preposition *in* (e.g. *in the absence of*).

As for the triplet “*a...of*”, Luzón underlines the collocate “*history*” as the most frequent collocate with a total of 70 occurrences and “*variety*” as the most significant with a percentage of appearance in the triplet of 100%. According to Luzón (2000:73) most collocates of this triplets are subtechnical nouns expressing quantity (e.g. *a subgroup of*).

Regarding the triplet “*be...to*” the most frequent collocate is “*likely*” and the most significant is “*asked*” with a percentage of 95%. The triplet “*be likely to*” is often preceded by “*more*” and “*less*”. Luzón (2000:75) explains that the high occurrence of this triplet is a consequence of the tendency of research papers to express tentativeness. Luzón also points that this triplet appears mainly with adjectives and past participles. The fact that it appears with participles reflects the fact that the passive voice is frequently use in this genre (e.g. “*be assigned to treatment*”). Fillers contained in the triplet “*be ...to*” usually express cause, similarity or modality (e.g. “*be related to*”, “*be similar to*”, “*be thought to*”).

The most relevant conclusions of Luzon’s study on medical papers are the following:

1. some fillers occur in a sole syntagmatic structure (e.g. “*the start of*”)

2. triplets are sometimes part of longer structures and are followed or preceded by certain items with specific semantic values (e.g. “*increase the frequency of*”).
3. the triplets found are associated with the move “occupying the niche” (Swales, 1990) in research papers (e.g. “*estimate the number of*”).
4. because the triplets found are used for specific discourse functions and they correspond to a concrete move, it is possible to provide writers of medical papers with formulaic patterns.

2.3. Contrastive Rhetoric

After reviewing the characteristics of scientific research papers, the next step is to analyze the differences between Spanish and English RAs since our study is based on the comparative analysis of these two languages. A general review of the literature will be carried out to establish if some differences between the aforementioned languages have already been found.

To start with, it is worth having a general view of the evolution of Contrastive Rhetoric. This concept was introduced by Kaplan in the mid sixties and it refers to the idea that logic is the basis of rhetoric and it has evolved out of culture, but neither rhetoric nor logic is universal. This implies that language and writing are cultural phenomena and, consequently, different cultures have different rhetorical tendencies. Often, the linguistic patterns and rhetorical conventions of the own culture are transferred to the L2 when writing. These basic notions

generated a lot of positive and negative reactions on the part of linguists.

In his study “Cultural thought patterns in intercultural education” in 1966, Kaplan established the basis for Contrastive Rhetoric (further on referred to as CR) by analyzing the organization of paragraphs written by students of English as a Second Language who had different origins. He contrasted Anglo-European, Semitic, Oriental, Romance and Russian languages. He distinguished 5 kinds of paragraph development attributed to different rhetorical strategies. In figure 2.1 Kaplan represents the conclusions drawn from his analysis by means of a graphical representation, which led to his article being referred to as “the doodles article”. He states that Semitic languages, like Arabic, use fix parallel structures giving preference to coordinate rather than to subordinate clauses, thus he represents the textual organization on these languages with parallel lines. Oriental languages prefer an indirect approach to the topic to be dealt with, that is, the main idea of the paragraph takes place at the end, so for its graphical representation he chooses a spiral. For Romance languages which present the material in a non-linear but digressive way, moving away from the main topic, he opts for a rough line. Similarly, Russian languages, whose graphical representation is a discontinuous rough line, present parallel and subordinate structures, irrelevant to the main topic.

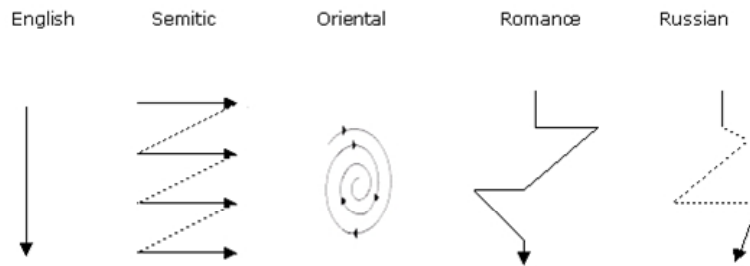


Figure 2.1: Graphical representation of language (Kaplan, 1996)

This analysis makes Kaplan the pioneer in CR; not only did he show that each language has unique rhetorical conventions, but additionally he focused his study on a rhetorical level, that is to say, he was the first to analyze units larger than a sentence, i. e., the paragraph.

This study was highly polemical. In fact, Paul Kei Mastuda states: “it is one of the most widely known and criticized articles in the history of applied linguistics” (2001:257). According to Trujillo (2001:316), Kaplan has always been modifying his searching methodology and making corrections on what he established as his premise for CR in 1966. Connor (1966:16 as cited in Trujillo, 2001:315) made one of the first and hardest critiques. She disapproved of both Kaplan’s theoretical statements and of his methodology. She accused him of favoring the writing of native English speakers. For Connor (2002:9), it is the reader’s expectations that determine what is considered coherent and direct writing; consequently, Kaplan’s conclusions on the linear line argument favored and reflected only what the English native speakers in this case regard as coherent. However, this may not coincide with the views of other language speakers on coherence, since

they may have different expectations. Another aspect criticized by Connor is the fact that he disregarded linguistic and cultural differences in writing among closely related languages (e.g. Chinese and Korean languages belonged to the same group in his study). She also criticizes Kaplan for considering the transference of the L1 into the L2 as only a negative aspect.

In the 1980s Kaplan, Connor and Purves give rise to a new concept of CR by means of 3 works, respectively: *Annual Review of Applied Linguistics* (1984), *Writing across Languages* (1987) and *Writing across Languages and Cultures* (1988). In the 1990s, several factors supported the rebirth of CR: the development of writing research, the reappearance of the concept of genre with Swales work *Genre Analysis* in 1990 and the improvement of research techniques within CR.

After Swales work it seems evident that two texts can only be compared if they belong to the same genre; that is, if they share a similar purpose. Swales (1990:58, as cited in Trujillo) stated that “the objective of a text is to shape the schematic structure of the discourse and influence and constrain choice of content and style” (1990:58), thus, texts with different objective present different lexical items, morphological structures, etc.

Connor (1996:5, as cited in Trujillo, 2001:318) affirms that despite linguists’ disagreement regarding CR and the many critiques it has received, there are some basic principles regarding this discipline on which most researchers agree:

- a) Language and Writing are cultural phenomena

- b) As a direct consequence, each language has unique rhetorical conventions.
- c) Linguistic and rhetorical conventions of the L1 interfere when writing the L2.

Summing up, the function of CR is that of comparing two texts written in two different languages so as to find out differences and similarities and to research the historical, cultural and linguistic reasons that explain these differences. Its purpose is to teach how to write while bearing in mind the textual idiosyncrasy of the languages.

Obviously, CR research involves the study of various disciplines. For its study Clyne (1987) suggested the inclusion of linguistics, psycholinguistics, psycho-sociology and sociocultural aspects, among others. Connor also supports this idea a few years later (1996, 27): “CR is interdisciplinary; it draws on several related fields of study such as text linguistics, composition pedagogy and literacy development. These influences enrich the scope and depth of CR research, enabling recommendations for teaching L2 writing”.

It is clear then, that CR is no longer limited to the study of the paragraph in terms of its grammatical correction and spelling as stated by Kaplan at the beginning the 1980s. It has evolved and considers other aspects not only at a textual level. Besides, it also includes other sciences related to language such as the philosophy of the language. It studies the relationship between culture and language. It also studies writing processes and mental models associated with these processes, known as

Psycholinguistics. Lastly, it analyzes the macro-levels of language and language teaching, or Text Linguistics.

2.3.1. English and other European Languages

For Andrade (1990:19) cultural models are present in a variety of domains (e.g. the cultural model of telling a lie, the cultural model of the American marriage or a restaurant interaction, etc). He explains that Swales' concept of genre represents a cultural scheme accessible for a writer, but this also implies that genres are a useful tool for readers as well, in the sense that they also share this cognitive scheme regarding the text and expect their expectations to be fulfilled.

Connor (2002) presents us with the most important research on CR comparing English with other languages. Ventola and Mauranen carried out a study in 1991 in which they compared scientific articles written in English by Finnish and by native English scientists. They came to the conclusion that English scientists used connectors more frequently and they presented a wider range of them. Additionally, English writers placed the main idea sooner in the text than the Finnish, who used less metatext and had problems concerning the use of the article. According to Moreno (2004:322), Metatext or metadiscourse is "the linguistic material in texts, whether spoken or written, that does not add anything to the propositional content but that is intended to help the listener or

reader to organize, interpret and evaluate the information given”.

Some years later, in 1993 Mauranen carried out a similar study but this time focusing on economic journal articles written by Anglo-American and Finnish authors. He came to the conclusion that the former presented a more reader oriented text. They showed a more positive politeness and a more explicit textual rhetoric. To him, the Finnish rhetorical strategies based on implicitness though probably effective in this language, are not when transferred into English.

A study contrasting English and Polish introductions in RAs was performed by Duszak in 1994. Comparisons between these two languages were also studied by Golebiowski in 1998 that focused on psychology journal writing and considered English texts more direct, assertive and positive.

According to Trujillo (2001:322), Spanish has been given much attention in the field of CR. Santiago, in 1970 and Santana-Seda, in 1975, wrote their PhD dissertations on this subject. The latter author's title being: *A Contrastive Study in Rhetoric: An Analysis of the Organization of English and Spanish-Speaking Paragraphs Written by Native Speakers of each Language*, and the former's: *A Contrastive Analysis of some Rhetorical Aspects in the Writing in Spanish and English of Spanish-Speaking College Students in Puerto Rico*. Álvarez (2005:9) explains that with their study they showed how Spanish texts are more complex in terms of syntax due to the important role that indirectness and abstraction play in this language. Trujillo (2001:322) states that most studies have come to the

conclusion that Spanish writing is more detailed and to that end uses more coordination and subordination than English.

To give some examples on researchers in the field of English-Spanish CR, Joy Reid compared 4 languages in 1988-1992 including Spanish and English and agreed to the general conclusions stated by the authors above. Also Monataño-Harmón shares this view in her 1991 study where she contrasts Mexican Spanish with American English.

As for Spanish researchers in Spain, it is worth mentioning Valero-Garcés, who in 1996 wrote an article to analyze English and Spanish metatext in business academic papers. She proves that Spanish writers use less metatext when writing English than Native-English writers do. Besides, she notes another different aspect in the way in which they refer to their presence in the text and their references to other authors. Spanish writers use brackets whereas English writers start their sentences with the names of the authors. Another difference is that in English the writer gives the reader little room for thought, whereas in Spanish writers stimulate the reader to interpret the text. According to Valero-Garcés, “the reader, not the author, is primarily responsible for effective communication”. Following this more active role on the part of the reader, another difference emerges regarding the conclusion of the text. The Spanish writer provides the reader only with a conclusion while the English writer gives a summary of the text and an accompanying evaluation.

With all the differences mentioned, she concludes that in spite of the uniformity imposed by the RAs genre, “there is

intercultural variation in the rhetorical preferences of writers". Thus, she stresses the need to investigate a greater corpus. According to Valero-Garcés, limitations on CR are due to the lack of resources to perform research with larger corpora.

In 1997, a similar study was carried out by Moreno in an article entitled: *Genre Constraints across Languages: Causal Metatext in Spanish and English Research Articles*. She also used business articles for this purpose. However, Moreno's conclusions are contrary to Valero-Garcés', for she draws the conclusion that intercultural differences regarding rhetorical preferences in metatext disappear due to the uniformity imposed by the requirements of the genre. Nevertheless, she admits that her study focuses only on a minor aspect of RAs. She calls this the "CEISRs", the causal metatext that should help the reader's comprehension regarding the cause-effect relation between sentences. For this reason at the end of her article she says:

Thus, this study does not permit any inference about the RA genre variable in controlling the rhetoric of text organization in Spanish and English. In other words, "these conclusions do not mean that the RA genre variable is powerful enough to control every rhetorical aspect of text organization in Spanish and English (Moreno, 1997:174).

Trujillo (2001:324) at the end of his article *Escritura y Cultura: La Retórica Contrastiva*, insist that although it seems reasonable to affirm that there are differences between the English native speakers and the Spanish writing in English, there are still many questions to be researched. To give an example, he mentions the need to prove whether the studied differences appear irrespective of the genre or text being

analyzed, of the purpose of the text, of the reader, etc. Additionally, the analysis should extend to elements such as the macrostructure, superstructure, etc. He also writes about the possible study concerning the comparison of Spanish with other languages. Therefore, he insists on the need to add more languages to the study of CR as well as the need to increase the size of the corpora.

In a very recent study, Moreno (2004) also tried to prove intercultural variation in rhetorical preferences despite genre uniformity in RAs. To this end, she analyzed a parallel corpus made up of 72 RAs on Business and Economics. Moreno analyzed researchers' productions in their L1. Her aim was to compare the preferred mechanisms used by English and Spanish researchers when presenting a claim by means of metatext. More exactly, she analyzes retrospective labels in the "premise-conclusion intersentential coherence relation". She finds some differences in the degree to which the English and Spanish scientists use different strategies. She underlines the importance of her study for both EAP and translation. She explains that it would be highly beneficial to make students aware of the differences and similarities of labelling premises in a premise-conclusion sequence and provide them with practice using equivalent labels in English. As for translation, she states that:

Given a retrospective label of a given degree of explicitness, which incorporates such and such meanings in the source language, they could help translators to predict the type of expression that would have plausibly been used in the target language. (Moreno, 2004:338).

2.3.1.1. English and non-European Languages

In addition to European languages, another significant contribution to CR according to Connor (2002:11-12) is the comparison of Arabic-English argumentation styles in 1997 by Hatim. He reviews Arabic argumentation from previous researchers such as Koch (1983), who claimed that Arabic tends to repeat, paraphrase and to use double argumentation. Hatim agrees on differences regarding Western and Arabic argumentation, the former being based on counter-arguments or opposition and the latter on what he calls “through argumentation”. However, he insists that for Arabic both kinds of argumentation are equally logical and that characteristics of Arabic argumentation are a result of socio-political factors; thus, it is speakers and not the language itself who are to be considered repetitive. Other linguists who continued studies on English-Arabic comparison are Hottel-Burkhart in 2001, who adds to Hatim’s findings that differences do not only concern the content of the argumentation but also the form and style.

In 1997, Scollon and Scollon performed CR in English and Chinese focusing on newspaper articles. They pointed out the difference in the use of quotations. In English they are not confusing but very clear, whereas in Chinese they seem to be very ambiguous in the sense that it is unclear who the quotation belongs to.

As far as these two languages are concerned, Zhu in 1997 and Kong in 1998 researched into the differences between Chinese and English business letters written by native speakers

of English and by Cantonese. They based their investigations on Swales moves analysis in 1990. Kong comes to the conclusion that moves were not the same and did not appear in the same order, and that the rhetorical structures in them also differed in the two languages. His explanation for this mismatch is similar to that exposed by Hatim to justify differences between Arabic and English. For Kong, differences are based on politeness and face system theories rather than on the inherent rhetorical patterns of the languages. He argues that in business transactions the relationship between the parts and the social expectations and considerations differ greatly between these two cultures. The British go straight to the point and make more face-threatening moves, whereas in Chinese the importance of deference gives no room for being direct or for face-threatening moves taking place. In fact, Chinese need more time or text in the letters in this case to express the request and furthermore they tend to justify the request throughout the text by means of interpersonal elements.

2.3.2. Different points of view on Contrastive Rhetoric

As Connor (2002:16) notes, CR has received much criticism. Some authors such as Spack (1997) argue that CR is insensible to cultural differences; others such as Scollon (1997) say that it is too focused on texts. According to Tannen (1985:212, as cited in Connor, 2002:18), documenting cross-cultural differences is considered by some people as

stereotyping and consequently discriminatory. However, not taking into account cultural differences often gives rise to “misinterpretation, and hence discrimination of another sort” (212).

In her article *New directions on Contrastive Rhetoric* (2002), Connor compares CR to intercultural pragmatics analysis. As she states, differences in L2 writing cannot only be related with L1 interference but with many other factors such as national culture, educational background, genre characteristics and mismatched expectations between readers and writers, etc. This puts CR in the same field of study as pragmatics.

Other linguists criticized CR for making students subordinate their native language and culture identities when writing to native-English speaker expectations.

Despite the critics against CR, linguists agree that cultural differences should be taught so that students bear them in mind and reach successfully the discourse community. As Mauranen states (2001:53), different cultural expectations cause misunderstandings.

2.3.3. Contrastive Rhetoric in this research work

For the elaboration of this research work, Swales’ statement (1990:58) that 2 texts have to belong to the same genre and have a similar aim to be compared was taken into account. Consequently, in this research work the function of CR is going to be that of comparing English and Spanish so as to

find out differences and similarities between the Conclusion Section of medical RAs in both languages.

Several studies on CR comparing English and other languages have been conducted in the literature and in different disciplines, especially in economic articles (Mauranen, 1993) and business articles (Moreno, 1997). CR has also been applied to RAs and to different sections like Introductions (Duszak, 1994) or even Conclusions (Ciapuscio, 2002).

Among the conclusions reached by different researchers, Spanish is said to have more complex texts in terms of syntax, to possess a more detailed writing and to use more subordination and coordination than English.

The content of the previous paragraph and Valero-Garcés' statement (1996) that intercultural variation in the rhetorical choices of researchers exist in spite of the uniformity imposed by the RAs genre have motivated the performance of this research work. Researchers (Valero-Garcés, 1996, Trujillo, 2001) underline the need to investigate larger corpora in CR, which is what this research work aims at. Mayor supports the same idea when she states the lack of CR analysis in the field of medicine especially for teaching purposes:

...dada la falta patente de análisis contrastivos sistemáticos y la inexistencia, que no escasez, de trabajos orientados a la enseñanza-aprendizaje de las peculiaridades de la comunicación médica en la combinación lingüística inglés-español. (Mayor, 2002:83)

2.4. Medical language

Since one of the linguistic features included in the analysis of the corpus is the lexical aspects, this Section analyzes the linguistic features of medical language in Spanish. The aim is to try to explain the process of enlarging medical Spanish language, i.e. the different ways in which words are included in it. The presence and origin of certain neologisms in the Spanish corpus will be studied as well as their evolution, their inclusion or rejection in the dictionaries. This chapter will show that sometimes it is difficult to figure out how words have entered the language, that is, whether by derivation of the own language or as borrowings of other languages. Regarding the use of anglicisms this section demonstrates that Spanish scientists still continue using many previously unaccepted anglicisms in their papers from 2000 to 2006 despite critique of several linguists.

2.4.1. Origin of medical language and its influence on the Spanish language

According to Ballesteros (2003:2), Medical language stems from the times of ancient Greece in IV and V b.C and has therefore Greek roots. Despite the fact that the Arabs stayed in Spain for 8 centuries, their influence on the Spanish vocabulary in general is very low, and it is even lower in the field of medicine. Other languages spoken in the Iberian Peninsula (e.g.

Portuguese and Catalan) or close to it, like Italian have influenced the general Spanish language, but not the medical language.

In general, scientific language has absorbed through time many terms from foreign languages that were dominant in a certain period of time because of cultural and/or economical reasons. As pointed by Ballesteros (2003:2), in the 18th and 19th century, Germany was the most outstanding country regarding the discipline of medicine. However, German researchers and other researchers from central Europe had a tendency to create neologisms based on Latin, which was the language studied at university by that time. Consequently, the Germanic influence is not relevant in the medical language. By contrast, French has had an influence on the Spanish language throughout history, and especially after Philip the 5th's reign and up to the 20th century. The influence of French neologisms on Spanish medical language is considerable. For example, the suffix “-age” was adopted in the Spanish language with a slight change in spelling, i.e. the replacement of the letter “g” with the letter “j”, namely “-age” became “-aje” and words such as “drenaje” appeared as a result of the influence of French in the Spanish language.

English had very little influence on Spanish up to the middle of the 20th century. However, the Second World War and the consequent economical and scientific leadership of the U.S.A brought anglicisms not only to the Spanish language, but to all languages in the world. Ballesteros (2003:3) notes the incomparable influence of anglicisms with his statement

“Resulta estremecedor que el inglés, en apenas medio siglo, haya introducido en nuestra lengua más palabras que el árabe en ocho centurias”.

Scientific hegemony in science was the main via to the spread of anglicisms , in fact, in the discipline of medicine one fourth of Spanish medical books have been translated from the English language, and over 80% of the bibliographical references in Spanish medical books come from journals that have been written English and stem from either Anglo-Saxon countries or non Anglo-Saxon countries.

The eclectic character of medical language with words stemming from different cultures is expressed by Saldaña (2003:103), who affirms that:

Cada nueva disciplina nace con nuevas voces, formadas unas con reglas históricas y otras con la fuerza del neologismo descarnado. Por ello, cualquier texto médico español aparece plagado de raíces grecolatinas, árabes, francesas, alemanas y, en las obras recientes inglesas.

2.4.2. New emerging vocabulary

In the introduction of the present research work, the need on the part of scientific researchers to find new terminology to be able to describe constantly emerging new technology was underlined. In the field of medicine constant advances have led medical doctors to create new vocabulary to refer to the latest concepts and discoveries. Ballesteros (2003) highlights the need of modern languages to be able to adapt to the mentioned medical breakthrough: “Es imprescindible que cualquier lengua

moderna tenga suficiente plasticidad para adaptarse al dinamismo de los continuos descubrimientos científicos”.

The arrival of foreign words to the Spanish language was started by researchers as a consequence of scientific progress and for cultural and social reasons. However, the inclusion of anglicisms after the Second World War was originated through the media by people who did not have much linguistic knowledge of the English language and used foreign words not as a consequence of a need to refer to something new, but rather, many times they resorted to neologisms with the purpose of making information sound more modern.

The use of anglicisms without knowing the style norms of the English language is criticized by Mayor (2006:133), who underlines the importance of knowing style norms especially to communicate in the English and Spanish medical context without language interferences:

Además del conocimiento léxico propio de la materia, aquellos que se dediquen a la divulgación médica han de tener claras las divergencias interlingüísticas con objetivo de evitar la transferencia de normas de escritura típicas del inglés pero extrañas en español.

Apart from Mayor, several authors claim that the inclusion of new vocabulary is sometimes not the result of a need on the part of scientists to explain new concepts but rather of other factors such as ignorance, idleness and pedantry (Iscla and Aleixandre, 2003). According to Lázaro “algunos

extranjerismos obedecen más a la necesidad que a la necesidad”. (Lazzaro, 1997).

According to Ballesteros (2003:2), the inclusion of new vocabulary is many times against the most important characteristics of medical language, namely, clarity and conciseness. Language needs to be precise and simple since it aims at transmitting the message without difficulties.

Las publicaciones médicas, como cualquier lenguaje científico, deben caracterizarse por la precisión, huyendo de las frases largas, de la abundancia de adjetivos y de cualquier pomposidad. La claridad de la información que se desea transmitir debe ser la condicionante esencial de cualquier artículo científico” (Ballesteros, 2003:2)

Ballesteros (2003:3-4) describes the only aspects that justify the incorporation of a new word in the vocabulary:

a) The only reason for the incorporation of a neologism is the inexistence of a term that brings the researcher to the need to create a new one. As an example he criticizes the incorporation of “posicionar” because of the existence of the equivalent Spanish verbs “situar” and “colocar”. Another author that shares Fernández’s opinion is Navarro (1992:588) that defines the neologism “posicionar” as “horrible neologismo calcado del inglés” and suggests more alternatives to the ones proposed by Ballesteros like “tomar posición”, “emplazar”.

b) The term has to be adapted to Spanish linguistic norms. Fernández (2004:34-36) refers to the problem that there are very few norms to fix a criterion in terms of incorporation of neologisms regarding orthography (e.g. Japón, París, both with an accent but “Washington”). Fernández (2004:36) also underlines the additional problem that the Real Academia de la Lengua Española (further referred to as Real Academia) has not established clear criteria in relation to the admission and rejection of words. Navarro (1992:575) refers to the occasionally difficult eradication of anglicisms once they are widely used in medical texts.

2.4.3. Different ways to increase medical vocabulary

The massive incorporation of vocabulary not based on any of the above reasons has led researchers to deal with the different ways in which neologisms are introduced in the Spanish medical language analyzing plausible resulting problems (Martínez, 2006, and Alexandre, 2003, etc.).

In the field of semantics Martínez (2006:84-90) and other authors (e.g. Ballesteros, 2003:4) provide us with different lexico-semantic options in which the Spanish language is enlarged by emerging vocabulary:

a) Words from other languages are slightly modified in terms of spelling and are adapted and accepted by the Language

Real Academia to the Spanish language (e.g. “estrés”, “escáner”, “chequeo”, “estándar”, etc).

b) Words from other languages are added to the language with or without morphological adaptation. Some examples of words with no morphological adaptation are: “shock” and “test”. This phenomenon is the most frequent in the last years (Navarro, 1992:575) and it is criticized because the use of these words, especially of anglicisms, is not due to the need to express something new, rather, these words have an equivalent in Spanish, which should be used instead (e.g. “second-look” instead of “revisión”, “shunt” instead of “derivación”, “cortocircuito”, “comunicación” or “anastomosis” (Navarro, 1992:579), “randomizado” instead of “aleatorizado”, “by-pass” instead of “derivación”, “puente” or “anastomosis”, “adherencial” instead of “adherente”, “plug” instead of “tapón”, “mandatorio” instead of “obligatorio”, “screening” instead of “cribado” and “bleb” instead of “ampolla” or “bulla”, etc.). Iscla and Aleixandre (2003:4) provide us with other examples such as the use of “nodal” instead of “ganglionar”. As Navarro (1994:142) points out: “Los extranjerismos siguen entrando violentamente en nuestro idioma y permanecen en él, desplazando en ocasiones los términos autóctonos”.

The majority of words identified in medical English coming from other modern languages are anglicisms. This fact is not surprising taking into account that the majority of neologisms are anglicisms followed by foreign words coming from French, Italian, Arab and German in this order (Fernández,

2004:32). The high rate of anglicisms and false friends in Spanish medical texts is to a great extent due to the fact that most Spanish publications are based on English bibliography (Navarro, 1994:142). In 1992, Navarro (1992:589) congratulates the DRAE for their task of incorporating neologisms and well as eradicating others (e.g. "morbilidad" and only "morbilidad" is accepted for euphonic reasons) (Navarro, 1992:584). However, he differentiates between anglicisms that are necessary and desirable (Navarro, 1992:587) such as "dopar", "dopado" and "dopaje" and those that are not. In the case of "dopar" he considers it necessary because it conveys an additional meaning to that of the Spanish verb "dorgar/se". The English verb "dope" implies the consumption of drugs with the purpose of improving stamina. As pointed by Rodriguez (1992:178) it is used in relation to competitions, contests. On the other hand, he criticizes the acceptance of some anglicisms by the DRAE like "discapacitado" and "discapacidad, which the DRAE recognises as calques of the English term "disabled". Navarro refers to them as "descarados anglicismos" (1993:587) that have the intention of replacing "minusvalía" and "minusválido". Other anglicisms accepted by the DRAE (1993:589) Navarro considers unacceptable (e.g. "rango", "relax", "test", etc.). Navarro's point of view is shared by Fernández (2004:37) who expresses his concern about the possibility of forgetting the own words as a consequence of the use of unnecessary anglicisms : "El problema es que si empleamos "hall" más que "vestibulo" o "recibidor" ... puede

ser que acabemos olvidando las nuestras, las ricas y expresivas voces españolas”.

Navarro (1994:504) underlines the importance of adapting anglicisms to our language since the future of most languages relies on its capacity to appoint new technical and scientific terms: “Resulta, pues, fundamental hallar con rapidez equivalentes propios para los neologismos ingleses, o bien adaptar éstos a las características de nuestro idioma”. As a consequence he denounces those who consider the language a mere mean of communication in the field of science and do not care about the intrusion of unnecessary anglicisms and disregard the cultural value of languages.

Ballesteros (2003:4) affirms that the Spanish language cannot easily produce short terms to replace Anglicisms, and this accounts for the phenomenon described in the above paragraph, i.e. the increasing adoption of anglicisms . As an example he compares the term “stent” with “prótesis endovascular”. Fernández (2004:21-22) refers to Cabré who explains the preference to use “Alzheimer” rather than “*patología de carácter neurodegenerativo que en el lenguaje médico se caracteriza como demencia*”. This deficiency of the Spanish language to create short Spanish equivalences is also reflected in the fact that sometimes researchers end up resorting to Greek terms instead of the Spanish ones preferring a shorter terminology. For instance, the term “Hairy cell leukaemia” translated for “leucemia de células peludas” would reasonably be substituted by “tricoleucemia”. This word contains the prefix “trico-” from the Greek noun “trikos” meaning “hair”.

Admittedly, “tricolucemia” is more euphonic than “leucemia de células peludas”. Navarro (1994:509) advocates for “tricolucemia” and emphasizes the need for a terminological uniformity. Ballesteros (2003:4) states that “euphony” is also a reason that accounts for the use of one term rather than another in the field of medicine. To illustrate this with an example, “kill cell” corresponding to “célula asesina” in Spanish was substituted by “célula citocida” with its root in the Latin verb “occido” meaning “to kill”. Navarro (1994:510) prefers “célula citolítica to “célula citocida”.

c) Existing words, mainly nouns, are turned after a process of derivation into a new term especially by the addition of a certain suffix or prefix. This addition implies a change in meaning and sometimes it also changes the grammatical category of the word (e.g. “antibioticoterapia”, meaning the treatment of a patient with antibiotics, or the verb “protetizar”, derived from the noun “prótesis”, meaning to implant a prosthesis). Other typical examples have taken place by the addition of the suffix “-al” turning nouns into adjectives like “muestral”, “situacional”, “poblacional”, and the like. This new terminology is not accepted at the beginning by the Real Academia. However, when the use becomes evident and wide many of these terms end up being included in the Diccionario de la Real Academia Española (further referred to as DRAE). Martinez (2006:84) recognizes the advantage of this mechanism that favors the fluency of the text avoiding verbal periphrases, but he underlines the risk of overusing them. In fact, sometimes

words suffer this kind of transformation, when it is absolutely unnecessary. The addition or change of the new suffix does not contribute to expressing a different nuance in terms of meaning. (e.g. “objetivizar” instead of “objetivar” or “lesional” instead of “lesivo”).

The flexible structure of the Spanish language is reflected in the many words that the Real Academia initially rejected but has accepted with time. Some of them are: “enfaticar”, “concretizar” (as a synonym of “concretar”), “chance” (as a synonym of “oportunidad”), “analítica” (for “análisis clínico”), “inusual”, “injuria” etc. The latter terms show a clear influence of the English language terms “unusual” and “injury”. This flexible feature of the language is underlined by Ballesteros (2003:7) when he makes reference to Fray Luis de León’s sentence “nuestra lengua es de cera para los que saben tratarla”.

Other terms which could also be considered unnecessary in terms of semantics are the many Latin terms used that have a Spanish equivalent. Their use in medicine is very spread and accepted, but Martínez insists on not overusing them and provides us with a list of the most used in the discipline of medicine: “a priori”, “a posteriori”, “ad hoc”, “ex profeso”, “in situ”, “sine qua non”, “versus”, “etc”.

d) The use of “false friends”, that is, of words that are translated wrong on the grounds of presenting a morphological similitude with a word in the other language. The most common false friend is the use of “severo” instead of “grave” as a

translation of the English term “severe”. The Spanish adjective “severo” means “strict” and not “serious” as the English adjective “severe”. The frequent misuse of the noun “severo” has turned it to be used in the same context as the English “severe” and consequently it has adopted the English meaning, which does not correspond with its actual meaning. In fact, the first translation registered in the Collins dictionary (1994) for the adjective “severe” is “severo”. Another example is the translation of the English “tablet” for “tableta”, “nodal” instead of “ganglionar” or “protusion” instead of “saliente” or “protuberancia” (Martínez, 2006:85). Another example that is worth mentioning is “evidence” that should be translated for “proof” or “finding”, i.e. “prueba” or “hallazgo” but it is by no means “evidencia”. As pointed by Navarro (2005:345) the word “evidencia” in Spanish means: “certeza clara, manifiesta y tan perceptible de una cosa, que nadie puede racionalmente dudar de ella”, thus, “evidencia” in Spanish is used to imply that something does not need to be demonstrated because it is clear. The level of certainty of the word “evidence” in English is much lower, thus it would correspond in Spanish to “indicios”, “signos”, “datos”, “pruebas”. Therefore, the sentence “*There is no evidence that breast implants cause serious disease*” would correspond in Spanish to “*no hay ningún indicio de que los implantes mamarios provoquen enfermedades graves*”.

Regarding false friends in the field of medicine, Navarro collects many of them in his work *Diccionario crítico de dudas inglés-español de medicina* (2005). In the introduction the author criticizes the existing medical dictionaries that collect

cognates with the same meaning in both languages but do not deal properly with problematic words such as false cognates “...en cambio, cuando el traductor encuentra algún vocablo inglés que plantea la más mínima dificultad y acude a esos mismos diccionarios, o no está recogido, o aparece incorrectamente traducido, o no se ofrecen algunas de sus múltiples acepciones” (Navarro, 2005, XVII). Navarro (2005) devotes his dictionary especially to clarify these complex words. For instance, the word “anthrax” does not correspond to the Spanish “ántrax” but to “carbunco”, i.e “carbuncle”. Similarly, the English “labor” does not correspond to the Spanish “labor” but it means “childbirth”, and so on.

e) Words are given a meaning that has not been registered in the dictionary. The word “bulla” in Spanish is defined in the dictionary as “concurcencia de mucha gente”, thus, it is a synonym of the English word “crowd”; however it is used to refer to “ampolla” or “blister”. Another example is “adolecer” defined in Spanish as “something having a defect or vice” and used however with the meaning of “lack”. Similarly, “obertura” refers in Spanish to “overture” but it is used with the meaning of “opening” or “gap”.

f) In the field of Syntax, the Spanish language is under the influence of other languages from which it copies syntactic structures such as verbal periphrasis or locutions. At this point it is important to highlight that medical texts also contain many expressions that are not specific of this discipline but rather

belong to the common language. (Navarro, 2001, XX). This accounts for the analysis of common syntactic structures coming from other languages. Some examples of these expressions are: “jugar un papel”, which is wrong in Spanish. Unlike in other languages such as English or French “jugar”, “to play” does not mean “to represent” but means “to play” in the sense of “to play a game”. Other locutions and periphrasis from other languages are: “en base a”, copied from the English “on the basis of”, instead the correct expression is “basándonos en” or “teniendo en cuenta”.

Another instance of an English structure taken in Spanish pointed out by Navarro (1994:507) is “depending on”. Despite the spread of this expression in medical texts he suggests the use of the term “según” which accounts for conciseness. (E.g He suggests translating: “*Depending on the individual response, the doctor will decide*” for “*El médico decidirá según la respuesta individual*”).

Ballesteros (2003:5) criticizes the frequent use of the passive as a result of the English influence on the Spanish language. In Spanish, the excessive use of the passive results in a less fluent text, more difficult to read and even to understand. Martinez (2006) recommends the use of the so called pasiva refleja instead. (e.g. “*Mañana se operará a este enfermo*” is preferred to “*Mañana este enfermo será operado*”).

g) Another consequence of the influence of the English syntax is the abusive use of the gerund. According to Ballesteros (2003:5), in Spanish the gerund should be used for simultaneous

actions or immediately prior to a second action. He condemns the habitual use of sentences like “*se practicó una exploración no encontrándose nada anormal*” instead of “*no se encontró nada anormal*”.

h) Another feature of medical language in Spanish is the use of ellipsis. This strategy also facilitates reading the text since it becomes faster. However, according to different authors (e.g. Martínez, 2006:87; Ballesteros, 2003:2) if overused, ellipsis can be against conciseness and clarity and for this reason they suggest avoiding its use. This lack of clarity by means of ellipsis is reflected in the example provided by Martínez (2006:87) that prefers “*no existen adenopatías, ni existen metástasis y la analítica es normal*” to “*no adenopatías, no metástasis, analítica normal*”. Although Martínez (2006:87) does not associate the ellipsis to the influence of another language, the example given in which the verb is omitted (e.g. “no adenopatías”) is reminiscent of the recurrent English structure made of the adverb “non-” plus an adjective (e.g. “non-invasive”). The difference lays in the grammatical value of the word after the adverb, namely a noun in Spanish but an adjective in English, but the omission of the verb takes place in both cases.

2.5. Corpus Linguistics

The analysis presented in chapter three is based on the analysis of a corpus. This Section provides an overview of the most recent literature on Corpus Linguistics in order to extract its relevance to translation and its application for this research.

Computing tools specially designed for the analysis of a corpus have enabled the study of recurrent patterns and grammatical structures as well as the analysis of the specific terminology of a certain field. Thus, the time and effort spent in the elaboration of a corpus is worth for the fact that this corpus can be used as database material in order to analyze the language used by experts.

The main requirements to be taken into account for the elaboration of a corpus according to different experts in corpus linguistics can be summarized as follows:

a) Representativeness: Biber (1998:246) states that the aim of the elaboration of a corpus is to represent a language or some part of it and adds that the concept of “representativeness” on corpus design is essential. That is, a corpus is not merely a gathering of texts but a selection of them according to specific criteria and to a specific aim. As an example he points out that when analyzing a corpus containing for instance conversations among teenagers our outcome cannot be generalized to all kind of conversations. The conclusions achieved by the analysis of the texts, which is based on observation would only represent a specific type of conversation.

b) Diversity: Another term closely associated with “representativeness” and also highlighted by Biber (1998:248)

in the field of corpus design is “diversity”. Researchers have to be concerned with the diversity of texts in their corpora. Biber states that there is no “general language” but each variety of language or register has its own patterns of use, that is, it uses different lexical, grammatical and discourse features. Therefore “a well designed corpus must represent the different registers of the language”, that is “register variation” Biber (1998:249)

c) Subject matter: In addition, subject matter is another aspect to be considered for a corpus since in a lexicographic analysis for example, it would affect frequency of words which clearly varies depending on the subject matter.

d) Another definition of corpus also underlying the idea of “representativeness” but also that of “frequency” is given by Crystal (1987:410) who defines corpus as “a representative sample of language, compiled for the purpose of linguistic analysis”. This underlines its utility to analyze frequent language patterns. Halliday (1991:41) also refers to this aspect when he affirms that the main objective of a corpus is that of “providing evidence of relative frequencies in the grammar, from which can be established the probability of grammatical systems”.

e) Data in a corpus must be authentic, a considerable amount systematically organized, presented in electronic format, not selected on linguistic grounds and selected with an aim of study in mind.

For Sinclair (1991:171) a corpus offers the possibility of recognizing what is characteristic of a language. He explains that a corpus must be limited, present an electronic format, be

selected according to some criteria and be representative taking into account the aim of the study.

According to Leech (1991:8-11), corpus linguistics could be said to have started in the times of structural linguistics after Bloomfield in the USA about the 1950s. For linguists such as Harris and Hill “corpus” was “the primary explicandum of linguistics” consisting of “a sufficiently large body of natural occurring data of the language to be investigated”. At the end of that decade, Chomsky underlined the shortage of corpora for linguistics studies and highlighted the competence of human intuition on the other hand that made up for the corpora scarcity. This idea was supported by the coming generation of theoretical linguists:

Any natural corpus will be skewed. Some sentences won't occur because they are obvious, others because they are false, still others because they are impolite. The corpus, if natural, will be so wildly skewed that the description would be no more than a mere list. (Chomsky, 1962:159)

Chomsky's studies based on “intuitive data” have been strongly criticized (Stubbs, 1996:28). Chomskyan linguists made data up for analysis, that is, they established some grammatical aspect by means of sentences and utterances that had been invented and were based on the linguists' intuitions.

Among the many linguists who clearly opposed Chomsky it is necessary to mention the pre-eminent scholar in corpus linguistics, John Sinclair. He states that examples are “not to be tampered with” and explains “the absurd notion that invented examples can actually represent the language better

than real ones” (1991:5). Sinclair (1991) argues that they cannot be trusted in terms of frequency, grammar and meaning. In fact, he explains that invented examples do not reflect real usage of a language. They should be rather used to clarify and improve explanations. Sinclair also exemplifies superiority of authentic data over intuition in a later study on the identification of ergative verbs carried out with Francis (1994:197).

According to Stubbs (1996:32) Chomskyan linguists preferred invented instances to real ones *knowing that most could never occur, and were also untypical*. This approach contradicts Sinclair’s theory of typicality, which states that one of the main objectives of a corpus is to determine what is central and typical in language. Also Firth’s contextual theory is against the Chomskyan’s point of view regarding analysis of utterances as a unit of analysis: *the complete meaning of a word is always contextual, and no study of meaning apart from a complete context can be taken seriously* (Firth, 1935:37). The above arguments support Stubbs’ announcement that the unit of study must be whole texts.

In his book “*Corpus, Concordance and Collocation*”, published in 1991, Sinclair defines the required data for corpus linguistics studies as authentic, not selected on linguistic grounds, a considerable amount, systematically organized and not annotated according to existing tagging theories. With the idea of data not selected on linguistic grounds, Sinclair (1991:100) refers to the fact that many times linguists choose certain texts because they contain a language aspect which calls their attention. However, this should not be the point for corpus

compilation, for this would show a peculiarity of the language rather than the regular pattern of a certain social role. Sinclair also opposes the reliance on annotation and advocates for recording structure by means of searching by word form and sorting by the superficial similarity of the instances. He affirms that “If... the objective is to observe and record behaviour and make generalisations based on the observations, a means of recording structure must be devised which depends as little as possible on a theory: *The more superficial the better*” (Sinclair 1987:107). Many people consider corpus annotation only the basis of corpus linguistics (Leech 1991:20-25).

2.5.1. Size of the corpus

Sinclair supports that a corpus should be as large as possible because this enables the linguist to reach a quality of evidence not previously attainable. For him quality in the case of corpus linguistics goes hand in hand with quantity. Looking at a lot of data allows the linguist to come to conclusions on frequency. With these arguments he affirms that the only limitation on corpus compilation should be associated with hardware and software storage and running problems.

The increasing size of corpora has been directly related to the increasing power of computers able to contain large bodies of text. Leech (1991:10) distinguishes 3 generations of corpus on the grounds of its size. The first one contains corpora with a million words and is represented by the Brown Corpus.

Other corpora belonging to this generation are: a) the Survey of English Usage Corpus (SEU), a corpus of spoken and British English created by Randolph Quirk in the 1960s. b) The CAMET (Computer Archive of Modern English Text) project, designed to collect a one million word corpus of written British English for use on the computer. In 1978 it became the LOB (Lancaster-Oslo/Bergen) corpus, taking its name from the first capital of the cities of its creators respectively (G. Leech, S. Johansson and K. Hofland) c) The International Computer Archive of Modern English (ICAME) in the 1970s.

The second generation is represented by John Sinclair's Birmingham Collection of English Texts in the 1980s. Another Corpus in the second generation is the SEC (Lancaster-IBM Spoken English Corpus) created in 1987 which consists of modern spoken English monologues.

The collection of corpora belonging to the second generation was possible thanks to a more advanced technology called the KDEM optical character-recognition device. Finally, the third generation is that of corpora containing hundreds of millions of words. To this generation belongs the British National Corpus (BNC), which was created in the period from 1990 until 1995. It is an annotated corpus created to investigate the British variety of the English language. It is made of 100 million words, 10 million of which are spoken material. Its creators were the British Library, 3 publishers (Longman Group Ltd, Oxford University Press and Chambers-Harrap) and two academic institutions: Oxford University Computing Service and Lancaster University. Considering the significant growth in

machine-readable text collections in the last years it can be expected that by 2021 one million million word corpora will be a reality.

Aijmer and Altemeyer (1991:1) underline that computer advances have given rise to new programs, greater storage capacity and faster processing of the data by the equipments. To this fact they add the important role that the internet has played in the easier creation of corpora. The internet provides us with texts from different fields and facilitates coordination and exchange of information on the part of linguists, in fact, the internet has become a corpus itself for many translators who use it to check the use of a certain term or expression. However, one is faced with the dilemma of whether this is a reliable information coming from such a source. For instance, frequency of use is not a piece of information that can be considered and it is also difficult to know whether the text found is representative of the search that is being carried out because the origin of the text or source is not known in many occasions. To counteract this problem, in the field of translation, professionals use search engines in local discs that search in a previous selected and saved corpus of texts that the translator has created (Candel, 2003)

Despite the great amounts of data that the internet offers and irrespective of its reliability as a source of information, Leech affirms that a collection of machine-readable text does not make a corpus. Therefore, he claims that the size of a corpus is not so relevant. The sample has to be representative according to our purpose of analysis. The task of a corpus is to represent a

given language, dialect or other subset of a language. Computerized archives cannot be taken arbitrarily according to the chances we have as it was the case in the second generation for instance, where data were collected beyond the bounds of the purpose. It is necessary to extract from the vast data resource the corpus that is needed for a particular study. In this way the size of the original corpus is reduced. Taking this into account, Leech distinguishes between general-purpose corpora, and gives as an example of a corpus belonging to this first generation, the SEU, and corpora designed for a more specialized function.

In addition, Biber (1998:248-249) notes that the size of a corpus does not only have to do with the number of words it contains, but also with the number of texts from different categories and the number of samples from each text. The number of texts required to represent a concrete aspect for analysis must be one that accomplishes representation of variation across speakers or authors. The number of texts needed per category varies depending on the design of the research. An important feature regarding text that one has to bear in mind is that it can vary internally to a large extent and this is why linguists have to be concerned with the number of samples needed. Biber illustrates this fact focusing on a corpus of RAs, for which samples representing the different sections a RA is made of, i.e. Introduction, Methods, Results and Discussion, would be needed because they have different patterns of language use. As for the number of words, lexicographic studies need very large corpora, but the most important issue to bear in mind is that for any kind of research diversity is more important

than the number of words contained, as Biber (1998:249) affirms: “size cannot make up for a lack of diversity”.

Swales (2006:20) advocates for specialized corpora and affirms that “bigger may not always be better, and size may not win all”. Lee (2001:37) emphasizes that a small specialized corpus is preferable because it is more homogeneous and more suitable for genre-based studies since “it can take into account interactional, pragmatic and contextual features in addition to purely linguistic ones”. To illustrate the truth of these statements, Swales (2006:20) gives as an example Luzon’s collection of medical RAs in 2000 and his own project MICASE based on academic research, which consisted of 152 speech-events of academic speech made of 1.7 million words.

A further aspect to take into account regarding the size of corpora that makes it less significant is the fact that there is a lack of balance between spoken and written corpora. Spoken data are still much more difficult to collect than written and the transcription into written form is complicated. But difficulties regarding the size of corpora may include different aspects. Swales (2006:20) highlights the problem of cost when undertaking the MICASE project. He states that the investment on putting the text together would never be recovered.

2.5.2. Computer tools for corpus analysis

In any case, irrespective of the corpus size, it is important to underline the need for a retrieval facility. It is necessary to access the corpus by means of tools for linguistic analysis, that is, search and retrieval programs such as the KWICK concordance program, the Oxford Concordance Program, etc. However, the collaboration of human intuition is still needed, because concordance programs do not have the capacity to differentiate words or signs taking into account the context in which they appear, for instance, they cannot tell the difference between “I” being a personal pronoun or a roman numeral. Therefore, the computer program is used as a simple sorting and counting data tool and it is the linguist who performs serious data analysis. This collaboration between human and machine supported by different linguists (Baker et al. 1996:27; Ajimer and Altenberg, 1991:15) is displayed by Leech in a descending scale in which he describes 4 models of analysis in which human intervention decreases from most to least (1991:17):

- a) Data retrieval model, in which the linguist analyzes the data.
- b) Symbiotic model, in which the machine partially analyzes the data but the analysis is improved by the linguist.
- c) Self organizing model, in which the machine analyzes the data and the linguist, provides the

parameters of the analytic system and the software.

- d) Discovery procedure model, in which the machine analyzes the data using its own categories of analysis and the linguist, only provides the software.

2.5.3. Different areas of study

Biber et al. (1998:1) point that the various types of linguistic analysis based on a corpus can be classified into two kinds: those on linguistic structures and those on the use of language.

Alan Partington (1998:2-3) provides us with a more detailed overview of the main 9 areas of corpus-based linguistic analysis:

- a) Style and authorship studies
- b) Historical studies
- c) Lexis: to research on the frequency of words, how they collocate, word senses, etc.
- d) Syntax: to investigate patterns of combination of words in phrases, clauses or sentences and the interdependence between structure and meaning of words. As Sinclair (1991:65) affirms: “It seems that there is a strong tendency for sense and syntax to be associated”.

- e) Text: according to Partington, studying corpora above the clause level has not been devoted so much attention except for a few linguists such as Stubbs in 1996 and Halliday (1985:345), who affirms that discourse analysis must go hand in hand with grammatical analyzes.
- f) Spoken language: to study pauses (Stenström, 1990), hedges, back-channel responses, softeners (Altenberg, 1990), etc.
- g) Register studies: to study sublanguages, that is, varieties of a language. Among the different linguists in this realm some of the most outstanding are Biber (1998), Biber, Conrad et al. (1994) and Nakamura (1993).
- h) Lexicography: This area has been very influential on English pedagogy because through these studies the English Cobuild dictionary provides learners with information on words usage, frequency, sense, etc.
- i) Translation studies: to study linguistic, semantic or pragmatic aspects in the process of translation.

These studies are performed using what Partington calls “equivalent corpora”. This expression is used to refer to corpora in two or more languages containing similar text types. On the other hand, Baker et al. (1996:177) take into consideration 4 kinds of corpora in the field of translation: parallel, multilingual, bilingual and comparable. It is important to consider certain

criteria when selecting the texts for the corpus such as the institution or referent, the textual genre, the topic and the function.

2.5.4. Corpus linguistics and translation

In the chapter “Corpus Linguistics and Translation Studies” (Baker, 1993:235-237), Baker supports the idea of translation consisting in looking for similar meanings and functions in the target language. She consequently recognizes that “the need to study authentic instances of similar discourse in the two languages becomes obvious” (1993:236). At the same time she states her disagreement with the priority conferred to the source text and to the notion of equivalence, which has been emphasized in the field of translation for many years. Since the early seventies the notion of “equivalence”, understood as “word-for-word” or “sense-for-sense” translation, has been declining and given way to the idea that for two expressions to be equivalent they need to be equal in their uses. It was this shift in the realm of translation from meaning to usage, or in other words, “from a conceptual to a situational perspective” (Baker et al., 1993:237) what has led towards corpus-based studies. Obviously, to research into usage and context real and extensive data is required. Therefore, shifting attention away from the concepts of equivalence and source text implied giving up studying individual instances in a text at a time and favored the study of many texts of the same kind. According to Baker

(1993:238-239), this change of perspective in translation that undermines the source text goes so far that in the early eighties a Tel-Aviv scholar named Toury affirmed that the source text was a mere “stimulus or source information rather than the starting point for analysis” (Baker et al., 1993:238). Another linguist who shared this view was Vermeer (1983:90), who states that it is the expectations on the part of the target reader and not the source text that must be given priority, and therefore they are the ones that determine the function of the translated text.

Sinclair (1992:395) advocates for an improvement on machine translation systems in the sense that they succeed in resembling the structures and natural patterns of languages.

Baker underlines the “urgent need to explore the potential for using large computerized corpora in translation studies” (Baker et al., 1993:248).

Finally, it is worth pointing out recommendations by Córdón et al., (1999:39) on documentation studies. To carry out the bibliographical research they point out the following phases: Firstly, the aim of the research has to be determined and what it is going to be used for. Secondly, it is important to define searching criteria: the time of the documentation of the research, the language or languages, and the kind of document to be collected.

2.5.5. Characteristics of the corpora used for this research work

The corpus of this research work was created taking into account different theoretical concepts of the different authors that have dealt with Corpus Linguistics.

Biber's (1998:246) proposal for Corpus Linguistics was taken into account when collecting the corpus for this research work. The corpus is "representative" because texts have been selected according to certain criteria:

- Texts had to belong to the field of medicine
- Texts had to contain a Conclusion Section
- An even number of English and Spanish articles were included
- The corpora have been created with an aim, namely that of establishing a comparison between English and Spanish linguistic structures in Conclusions of medical RAs.

According to Biber (1998:248) "diversity" is attained with register variation, which is present in the corpora thanks to the 3 different specialties in medicine: cardiology, paediatrics and psychiatry.

As for Crystal's proposal (1987:410) in relation to the subject matter, frequency of words that belonged to one of the 3 specialties rather than to general medical English have been rejected since the aim is the study of general medical language patterns.

Taking into account Sinclair's (1991:171) suggestion on the corpus, for this research work a similar number of conclusions was selected in both languages, so it can be considered a limited corpus. Additionally, it is presented in electronic format, which was the reason why the medical fields were limited to 3 specialties.

To collect the corpora, Sinclair's approach (1991:5) of authenticity against "invented and intuitive data" supported by Chomsky (1962:159) was taken into account. In fact, all the Conclusion Sections are taken from real medical journals, many of them present in the SSCI, thus no invented instances have been analyzed.

To determine the size of the corpora, Leech's theory (1991:8-11) was followed, in fact the samples of the research work are large enough according to the purpose of analysis. Both corpora, the English and the Spanish are purpose corpora as opposed to general corpora. Thus, the concept of specialized corpora rather than large but general and unspecific corpora supported by Swales (2006:20) and Lee (2001:37) was contemplated for the collection of the RAs. The research was limited to those RAs containing a Conclusion Section with a heading introducing this section that contained the word "Conclusion".

Thanks to computer hardware advances and thanks to new software applications (WordSmith™ version 4), the computer tool used to calculate recurrences, it was possible to calculate that the corpus of this research work contains a total of 163, 059 words.

From the 9 areas for corpus-based linguistic analysis that Partington (1998:2-3) mentions, there are several that this research work will cover: lexis, that is, the recurrence of words and how they collocate, in fact, recurrence of the most common words will be displayed on lists containing the different grammatical categories, i.e. verbs, adjectives, nouns, adverbs and conjunctions. Syntax will also be analyzed, in fact, this research work is going to deal with collocations and how meaning is conveyed with words presenting a different syntax. Naturally, it will be necessary to deal with grammar in the analysis of the syntax. Text will be analyzed since it is above the sentence level that the study focuses, in fact it deals with whole Conclusion Sections that are divided into different moves. As far as register is concerned, the research work focuses on medical English. Lexicography is also focused on in this research work where typical patterns are analyzed and thus words usage and frequency are studied.

The field of translation is the most relevant for the current research work. This research work will rely on “equivalent corpora” in Parrington’s (1998:4) terminology, that is, on two languages, namely English and Spanish containing similar text types, or in Baker et al. (1996:177) terminology, on bilingual and comparable corpora. These bilingual corpora will provide us with a source of information regarding the terminology used in medical English and Spanish.

The current corpus-based study has been conducted according to Baker’s (1993:235) idea that translation studies should be understood as searching for “equivalences” in

authentic instances of similar discourse of two languages. The term “equivalence” has been adopted for equivalent expressions in terms of use rather than for word-for-word translations, although, equivalent literal structures were also analyzed and indicated when occurring in the corpora. The idea is to focus on the concept of “typicality” emerged in the eighties to underline the options that are repeatedly opted for in a certain context to establish equivalence among English and Spanish medical structures.

Finally, following Baker et al.’s statement (1993:248) on the need to explore the potential for using large computerized corpora in translation studies the collection of the corpus started taking into account the recommendations by Cordón et al. (1999:39) on documentation studies. Firstly, the aim of the research was determined, namely, to establish recurrent lexicogrammatical patterns in medical language providing equivalences in two languages. Next, searching criteria were defined. The time of the documentation, which was between 2000 and 2006, the languages, Spanish and English, and the kind of document collected for analysis, which were medical Conclusion Sections of RAs.

ANALYSIS OF THE CORPUS

III. Analysis of the corpus

3.1. Corpus design

The elaboration of the corpus was an essential aspect of this research work. In fact, during the searching process a change in the field of study was necessary. Originally, the corpus for this study was thought to be an extension of *A Genre rhetorical and lexico-grammatical analysis of a corpus of academic articles* (Ricart, 2004), which contained 358 English RAs in the fields of Robotics, Telecommunications and Computer Science. However, a corresponding corpus in Spanish, that is a corresponding number of Spanish journals with a RA structure (i.e. containing the sections: Introduction, Methodology, Results and Discussion) was by no means attainable for several reasons. The first reason is that taking only RAs meant disregarding all the other possibilities that appeared in the journals which seemed on many occasions RAs but were not (i.e. Revisión Temática, Artículo Especial, Caso Clínico, Notas Clínicas, or in English: Reviews or Special Articles, Case Reports, Technical Notes, letters to the Editor, etc). It is important to make clear that most of the times RAs were referred to as “Original Papers” or “Artículos Originales”. A second reason that underlines the impossibility of a Spanish corpus in the fields mentioned above is that RAs needed an electronic format to enable the research of this study. Unfortunately, because of the tendency in the last years to

publish in English, finding Spanish RAs with this format was not an easy task. Different problems were faced; a considerable drawback in collecting these RAs is that the access to these journals in electronic format is under payment conditions; otherwise it is only available in the printed version. Additionally, the Spanish journals found in electronic format are not as technical as the English ones since they are aimed at a more general public and not at a group of experts or researchers (for the “discourse community” according to Swales, 1990). A third reason for the imposed change of field was the fact that Spanish journals’ extension in these fields is considerably reduced in comparison to English ones; in fact, they are on average between 3 and 5 pages. The fourth argument is that on top of the mentioned difficulties, most of the articles do not show a RA structure, except for a few that are written for congresses. This explains the fifth reason against the original idea of the Spanish corpus, since most congresses are held mainly in South America and thus papers presented are written in South American Spanish, which is not the variety of language to be considered in this study.

3.1.1. Selection of the corpus

3.1.1.1. Overcoming limitations and determining the aim and field of the corpus

As a consequence of the constraints of the disciplines chosen in the original research on academic articles (Ricart, 2004), there was a need to deal with a different technical field where the sample of Spanish publications was large enough to create an equivalent corpus. As the research on genre analysis of academic articles (Ricart, 2004) resulted rather arduous due to its highly technical nature, the alternative considered was to look for texts that were less related to machines and abstract concepts, difficult to understand for someone outside those fields. This idea led me to check Spanish publications in Medicine. Despite being a technical subject, this discipline has to do with human beings.

Firstly, in the research on genre analysis of academic articles (Ricart, 2004) only the recognition of recurrent lexicogrammatical realisations in the conclusion section of English RAs was carried out, whereas now .texts are going to be analyzed in two different languages. Texts belong to a new, specific and unique field, i.e. Medicine, and this fact will allow for less general and more precise conclusions, since every pattern described will belong exclusively to this discipline. Moreover, the study will not be limited to the recognition of certain patterns but it will be oriented to its application in translation. This research will enable and facilitate translation,

considering Baker's view (1993:237) that translation deals with finding corresponding structures in the target text rather than with providing a word-for-word translation. As emphasized throughout this chapter, translation relies on usage rather than meaning.

3.1.1.2. Specifying Conclusion sections in journals

During the collection of texts, it was observed that there was a lack of uniformity in the articles regarding the name of the headings standing for the Conclusion Section. For this reason, to compile this Section in English and Spanish articles it was necessary to establish certain criteria to give the study coherence and reliability. First of all, the Conclusion could not be included under a different heading (mainly in the Discussion Section), even in the case that there was a clear full stop followed by a new line with an introductory sentence informing about the beginning of the conclusion (i.e. "En conclusión..." or "La conclusión del presente estudio...", "To conclude..." etc.). The section should appear as a separate heading containing the word "conclusion". In most of the conclusions the section consisted of the word *conclusion* on its own, however, other combinations with this word were also accepted. Sometimes the conclusion section and the discussion section merged and appeared under one heading "discussion and conclusion". Regarding Spanish journals, Table 3.1 shows that in Paediatrics only one of the journals presented some examples, *Vox Pediátrica*, and in Psychiatry most of the journals displayed examples of this kind:

Paediatrics

Vox Pediátrica (PD.VP.Vol8 (1)-00-1),

Psychiatry

Cuadernos de Medicina (P.C.MPP.Vol.60/61-01/02-7 or P.C.MPP.Vol.60/61-01/02a-8)

Psicothema (P.PT.Vol.17.n1a-05-1 or P.PT.Vol.17.n1b-05-2)

Revista de Psiquiatría de la facultad pública de Medicina de Barcelona (P.RPFB.Vol.27(3)-00-5 or P.RPFB.Vol.28(1)-01-9)

Psiquiatría Biológica (P.PB.Vol.9n.4-02-7)

Revista de Psiquiatría y Psicología del Niño y del Adolescente (P.RPNA.Vol.4(2)a-04-3)

Psiquiatría Pública (P.PP.Vol.12a(3)-00-2)

Table 3.1 Spanish journals where the “discussion and conclusion” sections are merged

On two occasions, in the Spanish corpus, the heading was the combination of the Conclusion section and the Results section and appeared with the following names: “Resultados y Conclusión” (in *Psicothema* 2005-2003) and “Resultados y Conclusiones” (in *Revista de la Asociación Española de Neuropsiquiatría* (P.RAEN.Vol.XX, n. 73a-00-1).

Other combinations of the headings containing the word “Conclusión” were found in the following journals displayed in Table 3.2:

<i>Revista Española de Cardiología :</i>
“Conclusiones e Implicaciones Prácticas”
<i>Cuadernos de Medicina :</i>
“Conclusiones y perspectivas de futuro”
“A modo de conclusión” (P.C.MPP.Vol.71/72a-04- 26)
“A modo de conclusión y perspectivas futuras” (P.C.MPP.Vol.65d-03- 19)
“Conclusión: El uso clínico de la meditación” (P.C.MPP.Vol.69/70a-04-24)
“Intervención y conclusiones” (P.C.MPP.Vol.75b-05-39)
<i>Psicothema:</i>
“A modo de conclusión” (P.PT.Vol.16.n1c-04-12)
<i>Trastornos Adictivos :</i>
“Conclusiones y recomendaciones” (P.TA.Vol.6(3)d-04-4)
“Conclusiones: buprenorfina-Subutex ?y las condiciones de prescripción y financiación” (P.TA.Vol.7n (1) d-05-11)

Table 3.2 Different headings containing the world “conclusion”

Surprisingly, an alteration in the structure of the RA was found on one occasion, the Conclusion section preceded the Discussion Section (*Psicothema*: P.PT.Vol.13n1-01).

With regard to English journals, Discussion and Conclusion also merged in one section in two out of the three specialties: in Paediatrics, in the journal *Paediatric Neurosurgery*: (PD.PN.Vol.331-00-26 or PD.PN.Vol.38h-03-88) and in Psychiatry, in both the journal *Psychopahtology* (P.P.Vol.35c-02-21) and *Psychotherapy and Psychosomatics* (P.PP.Vol70c-01-7, P.PP.Vol69c-00-3 or P.PP.Vol.71h-02-19)

The heading “Summary and Conclusions” appeared occasionally in the 3 specialties; in Cardiology: (C.C.Vol.99ñ-03-98), in Psychiatry, in the journal *Psychopathology*

(P.P.Vol.38a-05-42) and in Paediatrics, in the journal *Paediatric Neurosurgery* (PD.PN.Vol.34b-01-31).

Another uncommon but repeated heading was “Conclusions and Future Directions”, which appeared in the specialty of Psychiatry, in the journal *Psychotherapy and Psychosomatics* (P.PP.Vol.73e-04-28) and in Paediatrics, in the journal *Paediatric Neurosurgery* (PD.PN.Vol.33j-00-24).

Also “Conclusions and Implications” appeared in the specialty of Cardiology, in the journal with the same name (C.C.Vol.95b-01-29) and in Psychiatry, in the journal *Psychotherapy and Psychosomatics* (P.PP.Vol.73f-04-29).

Other headings for the Conclusion Section that turned out in the journal *Cardiology* are shown in Table 3.3:

“Conclusion and Clinical Implications” (C.C.Vol.99c-03-86)
“Conclusion and Practical Implications” (C.C.Vol.98i-02-70)
“Concluding Remarks” (C.C.Vol.94e-00-14 or C.C.Vol.94f-00-15)
“Conclusion and Possible Clinical Application of the c-TVI Method” (C.C.Vol.102b-04-116)
“Conclusions and Future Research” (C.C.Vol.93i-00-9)
“Limitations and Conclusions”

Table 3.3 Different headings in the “conclusion” section in the journal *Cardiology*

Finally, another heading only present in Psychiatry, in the journal *Psychotherapy and Psychosomatics* was “Conclusion and Suggestions for Further Research”.

Therefore, all the above headings standing for the Conclusion Section, which contained the word *conclusion*, were accepted in the corpus unlike other headings such as “Further Comments”, “Comentario”, “Future Research”, etc. not containing this word, which were disregarded.

3.1.1.3. *Selecting the journals*

The selection of journals was carried out randomly with the aim of ensuring that a considerable number of Medicine RAs presented the corresponding IMRD model and that they also contained a Conclusion Section. The number of articles containing a Conclusion Section was checked in two journals of Dermatology dating from 2005 up to March 2006, which was the most recent volume published at the time of research. The selected journals were: *Piel* and *Actas Dermo-Sifiliográficas*. Another journal selected not belonging to Dermatology was *Química Clínica*. The results obtained from the latter journal were a bit disappointing. Despite the period of time consulted being wider than that of the journals of Dermatology, namely from 2002 to 2006, only 4 articles contained a conclusion Section. These were rather few, but expectations increased on verifying the existence of a total of 7 conclusions in *Piel* (namely, one in 2006: Vol. 21 num. 2 February p.67 and 6 in 2005: Vol. 20 num. 8 Oct p 370, num. 7 Sep p. 314, num. 5 May p 211, num 3 march p 112, num 2 Feb p 63 and num 1 January p.3) and of two conclusions in *Actas Dermo-Sifiliográficas* (Vol. 96 num.7 Sep 2005 p 424 and Vol. 96 num. 4 May 2005 p 222). These results accounted for the existence of Conclusion Sections in medical RAs. The investigation started by checking at the medical university of Valencia periodicals section those Spanish journals that showed a higher number of volumes a year, since the chance of finding a Conclusion Section was thought to be higher. In order to get an idea of the amount of articles with a

Conclusion Section that were available yearly, the years 2005 and 2006 were checked until the month of March. The first journals checked were *Medicina Clínica* and *Revista Clínica Española*. However, the former, which has two volumes with 20 issues a year, thus a total of 40 issues a year, only presented one article containing a Conclusion Section (Vol. 126 n.14). As for the latter journal, only one article with a Conclusion Section was found in the same years (Vol. 205 num 12 Dec 2005 p 595).

After this first step it was rapidly concluded that a higher number of volumes did not necessarily imply a higher number of Conclusion Sections. Bearing this in mind, it was decided to concentrate on a few areas of medicine rather than selecting journals belonging to very different fields, since fewer specialties would also make the lexicographic analysis more valuable. Focusing on those specialties that presented a significant amount of journals, the study was restricted to three disciplines: Psychiatry, Anaesthesiology and Dermatology. On collecting the conclusions belonging to the first specialty a second problem was observed: although some journals showed a representative number of articles, many could not be accessed on line. This was the case of *Anales de Psiquiatría* which publishes 6 numbers a year. This journal contained two Conclusion Sections in 2006: (Vol. 22 January 2006 num 1 p.8-16 and Vol. 22 January 2006 num 1 p.17-22) and 15 Conclusion Sections in 2005 (Vol. 21 Oct-Nov 2005 num 6 p.274-279, p.280-288, num 5 p.223-229,230-236, num 3 p.95-101,p.102-110,111-116,117-123, num 2 p.45-55, p.56-66, p.67-72,p.73-81, num 1 p.15-23, 23-31, 39-44).

To solve the problem described in the paragraph above, two decisions were made. Firstly, the period of time originally selected for research (from 2005 to 2006) was extended from two to 7 years to make sure that a representative number of samples could be attained. Conclusions were therefore sought from 2000 to 2006. Secondly, in order to avoid finding samples which lacked a corresponding electronic format, research on printed papers was disregarded and the exploration was limited to online resources. Consequently, it became essential to investigate sources available in internet that could provide electronic Spanish medical papers. This second action imposed the specialties to be chosen for research forcing the abandonment of two of the first specialties selected initially. From this moment on the study was focused on: Psychiatry, Paediatrics and Cardiology.

3.1.1.4. Online Sources to access medical Journals

In her paper “Acceso abierto y revistas médicas españolas”, Abad (2006:4-8) states that there were no free access medical journals with a complete text in the internet in 1999. However, the situation has changed dramatically in the last years. In 2004 there were 360 medical journals accessible online, out of which 159 (this would represent 44. 7%) were freely accessible to a complete text. This number rose to 170 in 2005, although only 127 contained original articles. Access to these journals is sometimes confusing. There is more than one

URL for the same journal and sometimes one of these URLs has been updated more recently. This means that journals are many times included in different databases. As a consequence, the same journal can be found in a platform but at the same time in the web page of a certain editorial. Additionally the information in these two pages may not coincide.

Table 3.4 shows the main internet sources found in the web on searching Spanish medical journals:

A. The internet website www.fisterra.com
B. The main searching source in Hospital La Fe in Valencia with a link to an editorial called “Doyma”: http://www.doyma.es .
C. Other internet sites reached through the search engine/ directory: “Google” http://dialnet.unirioja.es http://www.secp.org/SB11listrev.asp , http://www.bibliopsiquis.com/asmr/numeros.htm , http://dinarte.es

Table 3.4: Internet sources available for the elaboration of the Spanish corpus.

The first source found was the internet website www.fisterra.com. It is a virtual library in which journals are sorted out under different specialties. (See appendix III).

Considering the list in appendix III and focusing our attention on the three specialties selected for this research work, several journals were found although not all of them were worth for our study. The web-site address below shows a list of the journals belonging to the mentioned three disciplines. It can be observed that, after the journal’s name, different abbreviations

are added to give some information regarding the availability of the journal. These abbreviations are the following: “Sum” or “Res”, which stand for “summary” or “resumen” in Spanish, “TC”, which stands for “Texto completo” or “Complete Text” and “PDF”, shown always into brackets “(PDF)”, indicating the existence of the complete text in “PDF” format. Obviously, only those texts showing “TC” or “PDF” could be used for this research. Table 3.5, Table 3.6 and Table 3.7 below show all the journals in www.fisterra.com in the specialties of Psychiatry, Paediatrics and Cardiology and their availability:

Psychiatry

Title of the journal	Availability
<i>The Journals of American Psychiatric (Edición en español)</i>	
<i>Actas Españolas de Psiquiatría.</i>	Sum; Res
<i>American Journal of Psychiatry (Edición en español)</i>	
<i>Anales de Psiquiatría.</i>	Sum
<i>Ansiedad y Estrés.</i>	Sum
<i>Aula Médica Psiquiatría.</i>	Sum; Res
<i>Bipolar Disorders (Edición en español)</i>	
<i>Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace.</i>	Sum; Tc (PDF)
<i>European Journal of Psychiatry (ed. Ingles).</i>	Sum; Tc (PDF)
<i>European Journal of Psychiatry. (ed. Español).</i>	Sum; Tc (PDF)
<i>Universidad de Zaragoza, Facultad de Medicina</i>	
<i>Monografías de Psiquiatría.</i>	Sum; Res
<i>Norte de Salud Mental.</i>	Sum; Tc (PDF)
<i>Asociación de Salud Mental y Salud Comunitaria</i>	
<i>Papeles del P. Jofré.</i>	Sum, Tc
<i>Revista de la Sociedad de Psiquiatría de la Comunidad Valenciana</i>	
<i>Psicología.com.</i>	Sum; Tc
<i>Psicothema.</i>	Sum; Tc
<i>Revista de Psicología del Colegio Oficial de Psicólogos del Principado de Asturias.</i>	
<i>Psiquiatría Biológica.</i>	Sum; Res
<i>Psiquiatría.com.</i>	Sum; Tc
<i>Psiquiatría Pública.</i>	Sum; Tc (PDF)
<i>Psiquis.</i>	Sum
<i>Psychiatric Services (Edición en español)</i>	
<i>Revista de la Asociación Española de Neuropsiquiatría.</i>	Sum; Tc(PDF)
<i>Revista de Psicogeriatría.</i>	Sum
<i>Revista de Psiquiatría.</i>	Sum; Tc (PDF)
<i>Universidad de Barcelona. Facultad de Medicina</i>	
<i>Revista de Psiquiatría y Psicología del Niño y del Adolescente</i>	Sum; Tc (PDF)
<i>Revista Goze.</i>	Sum; Ts
<i>Academia Vasca de Ciencias de la Salud Mental</i>	
<i>Salud Mental Atención Primaria.</i>	Tc (PDF)
<i>Siso-Saúde.</i>	Sum;
<i>Asociación Galega de Saúde Mental</i>	
<i>Trastornos Adictivos.</i>	Sum; Res

Table 3.5: Journals from the web page www.fisterra.com belonging to the specialty of Psychiatry

Paediatrics:

Title of the journal	Availability
<i>Anales Españoles de Pediatría.</i>	Sum; Res;Tc
<i>Asociación Española de Pediatría</i>	
<i>Boletín de la Sociedad Canaria de Pediatría. Canaria Pediátrica.</i>	Sum, Tc (PDF)
<i>Boletín de la Sociedad Vasco-Navarra de Pediatría.</i>	Sum; Tc (PDF)
<i>Boletín de Pediatría.</i>	Sum; Tc (PDF)
<i>Sociedad de Pediatría de Asturias, Cantabria y Castilla y León</i>	
<i>Canarias Pediátrica.</i>	Sum; Tc (PDF)
<i>Sociedad Canaria de Pediatría</i>	
<i>Evidencias en Pediatría</i>	
<i>Asociación Española de Pediatría de Atención Primaria</i>	
<i>Foro Pediátrico.</i>	Sum; Tc
<i>Sociedad de Pediatría de Atención Primaria de Extremadura</i>	
<i>Monografías de Pediatría.</i>	Sum; Res
<i>Pediatría Catalana.</i>	Sum; Tc (PDF)
<i>Pediatría Integral.</i>	Sum; Res
<i>Asociación Española de Pediatría. Sección de Pediatría Extrahospitalaria.</i>	
<i>Pediatrika.</i>	Sum
<i>Revista Española de Pediatría.</i>	Sum; Res
<i>Revista Pediatría de Atención Primaria.</i>	Sum; Tc
<i>Vox Paediatrica.</i>	Sum; TC (PDF)
<i>Sociedad de Pediatría de Andalucía Occidental y Extremadura</i>	

Table 3.6: Journals from the web page www.fisterra.com belonging to the specialty of Pediatrics**Cardiology**

Title of the journal	Availability
<i>Clínica e Investigación en Arteriosclerosis.</i>	Sum;Res
<i>Hipertensión. Sociedad Española de Hipertensión.</i>	Sum; Res
<i>Monocardio.</i>	Ts
<i>Sociedad Castellana de Cardiología</i>	
<i>Revista de Cardiología Extrahospitalaria.</i>	Sum; Tc
<i>Revista de Investigación Cardiovascular.</i>	Sum; Res
<i>Revista de la Sociedad Castellano-Leonesa de Cardiología.</i>	Sum; Tc
<i>Revista de la Sociedad Andaluza de Cardiología.</i>	Sum; Tc
<i>Revista Española de Cardiología.</i>	Sum; Res; Tc
<i>Sociedad Española de Cardiología</i>	
<i>Revista Latina de Cardiología</i>	Sum, Res

Table 3.7: Journals from the web page www.fisterra.com belonging to the specialty of Cardiology

It was necessary to cope with some constraints on the accessibility to journals with a “TC” or “PDF” abbreviation. Several problems were identified: some journals had no link to the articles despite having the letters “TC” or “PDF” after their name (e.g. *Revista de la Sociedad Castellano-Leonesa de Cardiología*). Others could not be opened (*Papeles Del P.Jofre*).

There was also a group of them which needed a password to be accessed (e.g. *Psicologia.com*). Besides, in some cases, access was only possible up to year 2000 (e. g. *Psiquiatría Pública* or *Revista de la Asociación Española de Neuropsiquiatría*). In other cases, a certain year could not be checked (e.g. year 2006 in *Psicothema*). On other occasions access was even more limited, because only one year was accessible (e.g. *Revista de Psiquiatría* and *Psicología del Niño y Del Adolescente*, where only the year 2004 was available and *Foro Pediátrico* with only 2005 available).

Additionally, several journals had to be rejected for other reasons: some contained articles that were not RAs and therefore did not present the corresponding sections of the IMRD model (e.g. *Salud Mental*, *Atencion Primaria*, *Revista de Cardiología extrahospitalaria*). In fact, there were journals called “boletines” which contained “protocolos” rather than “artículos originales” or “papers” (e.g. *Boletín de Pediatría*). Finally, there were also journals in English which were therefore not useful to create a Spanish corpus (e.g. *European Journal of Psychiatry*).

The second source was that of the journals available through the intranet of the University Hospital La Fe of Valencia (See appendix III).

In this appendix it can be seen that the database of the University Hospital La Fe only contains 18 Spanish journals out of the over 380 available in English. This fact accounts once more for the tendency to publish in English rather than in any other language. The Spanish journals available in this page are the shown in Table 3.8 below:

Anales Españoles de Pediatría	Index de Enfermería
Atención Primaria	Medicina Clínica
Cirugía Española	Metas de enfermería
Clínica e Investigación en Ginecología y Obstetricia	Psiquiatría Biológica
Educación Médica	Rehabilitación
Enfermería Clínica	Revista Clínica Española
Enfermería Intensiva	Revista de Calidad Asistencial
Fisioterapia	Revista de Ortopedia y Traumatología
Gastroenterología y Hepatología	Revista Española de Cardiología

Table 3.8: Spanish Journals in the database of La Fe University Hospital

From the 18 journals cited in the above paragraph, those belonging to the specialties to be studied in this research work are: *Anales Españoles de Pediatría*, *Atención Primaria*, *Psiquiatría Biológica* and *Revista Española de Cardiología*.

Except for *Index de Enfermería* (this journal was linked to the platform *Index de Enfermería Digital*) and three other journals that could not be accessed (*Educación Médica*, *Anales de Pediatría* and *Metas*), all the Spanish journals on this website were directly linked to “Doyma”. This is a medical publisher portal accessible via internet, which can be reached under the internet address: www.doyma.es. It is the most important provider of Spanish scientific and medical information. It publishes medical journals in 32 different areas of medicine. The access to these journals was occasionally possible without being registered (e.g. free access to the Journal *Trastornos Adictivos*), especially the consultation of one article or one year. However, to cover the information required for this research work, i. e. articles in three determined specialties in a period of time of nearly 7 years (from 2000 to the 1st of April

2006) a password was many times indispensable. Doyma provides us with a more extensive list of Spanish journals than the intranet of the University Hospital La Fe, which includes 62. (See appendix number III).

The corresponding journals to the 3 selected fields of this study found in the medical editorial Doyma are displayed in Table 3.9:

Psychiatry	<i>“Psiquiatría Biológica”, “Rehabilitación Psicosocial”, “Trastornos Adictivos” and “Vigilia del Sueño”.</i>
Paediatrics	<i>“Anales de Pediatría”, “Anales de Pediatría Continuada”</i>
Cardiology	<i>“Clínica e Investigación en Arterioesclerosis”, “Hipertensión” and “Revista Española de Cardiología”.</i>

Table 3. 9: Journals in the medical editorial Doyma belonging to the specialties of Psychiatry, Pediatrics and Cardiology

Unfortunately, as stated above, because of the difficulties to access the journals, only one journal in each field was attainable and included in the corpus: *Revista Española de Cardiología*, in the field of Cardiology, *Anales de Pediatría*, in Paediatrics and *Psiquiatría Biológica*, in the field of Psychiatry.

Other sources were internet web-sites found by searching the web such as: <http://www.secp.org/SB11listrev.asp> for Paediatrics or the directory called “latindex” with the following internet address: <http://dialnet.unirioja.es>. In this directory, there were 22 journals belonging to the three concerning specialties for this research work under the link “Ciencias de la salud”.

Table 3.10 shows a list of the mentioned journals:

<i>Acta Pediátrica Española</i>	<i>Actas Españolas de Psiquiatría</i>
<i>Anales de pediatría</i>	<i>Anales Españoles de Pediatría</i>
<i>Archivos de Psiquiatría</i>	<i>Clínicas Pediátricas de Norte América</i>
<i>Conductas adictivas</i>	<i>Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace</i>
<i>Conductas de Psiquiatría Comunitaria</i>	<i>Informaciones Psiquiátricas</i>
<i>Monografías de Psiquiatría</i>	<i>Pediátrica</i>
<i>Psiquiatría Biológica</i>	<i>Psiquis</i>
<i>Revista de Asociación Española de Neuropsiquiatría</i>	<i>Revista de Neurología</i>
<i>Revista de Psicopatológica y Salud mental del Niño y del Adolescente</i>	<i>Revista de Psiquiatría de la Facultad de Medicina de Barcelona</i>
<i>Revista Española de Neurofisiología</i>	<i>Revista Española de Pediatría</i>
<i>Salud Mental</i>	<i>Trastornos Adictivos</i>

Table 3.10: Journals from the web page <http://dialnet.unirioja.es> belonging to the specialties of Psychiatry, Pediatrics and Cardiology

Despite the large number of journals in Table 3.10, only four could be consulted without a password: *Anales de Pediatría*, *Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace*, *Informaciones Psiquiátricas* and *Trastornos Adictivos*. Access to the first journal; *Anales de Pediatría*, had already been possible through www.doyma.es, but with this source it was possible to get the articles from *Transtornos Adictivos* as well, that had not been possible to reach without a password in Doyma. As for *Informaciones Psiquiátricas*, many of its articles were published in *Geriatrika*; *Revista Iberoamericana de Geriatría y Gerontología*. However, the journals have been written by Spanish doctors. Although the papers found in 2006 were not structured according to the

IMRD model, they contained a Conclusion Section. Regarding Psychiatry, another journal was available under the internet address: <http://www.bibliopsiquis.com/asmr/numeros.htm>. This journal, *Avances en salud mental relacional*, could not be used for the corpus because the existence of 3 or 4 volumes a year dating from 2002 until 2006, only one number is available up to 2006 and additionally, articles have no RA structure.

As for Paediatrics, it is remarkable that many South American journals were disregarded in this specialty for the fact that the interest is the study of the Castillian usage of the language and not the Latin American usage, which is a variety of the language not considered for this study. (*Acta Pediátrica de México, Archivos Argentinos de Pediatría, Archivos Venezolanos de Puericultura y Pediatría, Revista Cubana de Pediatría, Revista Mexicana de Pediatría*, etc.).

3.1.1.5. Data: selected English and Spanish journals for the corpus

After the description of the different sources found to create the Spanish corpus and the explanation accounting for the journals that could not be included for different reasons, the list below shows the journals finally selected for the corpus and the number of conclusions that could be obtained in each of them. Table 3.8 below is divided into three parts according to the 3 specialties chosen for this research work: Spanish journals of Psychiatry, represented by a total number of 10 journals,

Spanish journals of Paediatrics, with 7 journals and Spanish Journals of Cardiology, represented by only one journal.

Within the 3 specialties three sections are displayed according to the sources where the journals have been taken from as shown in Table 3. 11:

Specialty	Source	Name of Journal	Number of conclusions	
Psychiatry	A : www.fisterra.com	Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace	40	
		Norte Salud Mental	3	
		Psicothema	29	
		Psiquiatría Pública	3	
		Rev de la Asociación Esp de Neuropsiquiatría	2	
		Revista de Psiquiatría de la fac de med de Barcelona	41	
		Rev de Psiquiat y Psicología del Niño y del Adolescente	5	
	B : www.doyma.es (Intranet of the Hosp LaFe)	Psiquiatría Biológica	13	
		Transtornos Adictivos	17	
		Informaciones Psiquiátricas	4	
	Paediatrics	www.fisterra.com	Boletín de la Sociedad Canaria de Pediatría.	8
	Boletín de la Sociedad Vasco-Navarra de Pediatría		7	
	Foro Pediátrico		2	
Revista Pediatría de Atención Primaria	5			
Vox Pediatría	6			
www.doyma.es (Intranet of the Hos La Fe)	Anales de Pediatría		9	
C : http://www.secp.org/SB11listrev.asp http://dinarte.es	Sociedad Española de Cirugía Pediátrica		44	
Cardiology	B : www.doyma.es	Revista Española de Cardiología	71	

Table 3.11: Spanish journals selected for the corpus of the study

Despite the difficulties in the implementation of the corpus, it was possible to collect 18 Spanish journals that contained a total of 311 articles with a Conclusion Section.

From the very start, while searching for articles with a Conclusion Section, it became clear that journals including this section did not explicitly state this requirement in their author's guidelines or "normas de publicación". In fact, only five out the

18 Spanish journals selected for this research work required explicitly a Conclusion Section (i.e. *Revista de Psiquiatría de la Facultad de Medicina de Barcelona*, *Trastornos Adictivos*, *Revista de Psiquiatría y Psicología del Niño y del Adolescente*, *Revista Española de Cardiología* and *Revista Pediatría de Atención Primaria*). The author's guidelines in the rest of the papers required the IMRD model except for 3 of them. In these three journals a variation of the IMRD was expected. Instead of the section "Material y Métodos", in the journal *Psiquiatría Biológica* the name given to this section was "Pacientes o sujetos y Métodos". In the journals *Vox Pediátrica* and *Anales de Pediatría*, "Materiales" or "Pacientes y Métodos" was the expected heading for this section. The least conventional journal regarding structure was *Informaciones Psiquiátricas*. The sections in the author's guidelines were as follows: Introduction, Methods, Ethics, Statistics, Results, Discussion and Acknowledgements.

Regarding the selection of English journals, it is important to point out that the three specialties used for the compilation of the Spanish corpus imposed the data compilation for the English corpus. The constraints and problems in the compilation of the Spanish corpus disappeared when dealing with the English language. Finding English publications on Cardiology, Paediatrics and Psychiatry was much easier. In fact, the number of journals needed to reach an equivalent corpus to the recently created in Spanish was much lower. English journals had a higher number of volumes and often of articles as well.

Evidence that publication in English is very prolific is also provided by PubMed, which is the main medical library in internet. It is a service of the National Library of Medicine and the National Institutes of Health, which can be accessed via the internet address: www.pubmed.com. With the help of this service it was possible to check the productivity of journals in English language in the three specialties. To obtain some relevant data supporting this hypothesis it was necessary to limit the research in Pubmed taking into account 3 aspects: language, namely English, the date of publication: journals published between the period of January 2000 up to June 2006 were selected, and the default tag: the option “journals” was selected.

Once this specific information was entered in the Pubmed interface, the outcome confirmed the productive publication in English with the existence of 5,819 articles for Paediatrics, 1,041 for Cardiology and 375 for Psychiatry. It is necessary to point out, though, that to read the articles in Pubmed sometimes required a password.

However, the English corpus was collected when carrying out research at the *Institute for Human Genetics* in the University Hospital Benjamin Franklin (“Medizinische Fakultät der Charité”) in Berlin. Publications with an electronic format were available in the medical library by means of an online-catalog: www.charite.de/charite. (See appendix III)

From this appendix, the journals belonging to the specialties at hand are: *Cardiology*, *Psychopathology*, *Psychotherapy* and *Psychosomatics* and *Paediatrics Neurosurgery*.

Despite the fact that the amount of journals with electronic format was not significant, the number was enough because the journals presented many volumes and articles. The journals selected for the corpus account for this fact, but the clearest example is illustrated by the journal *Cardiology*, with two volumes a year containing 4 numbers each. After the research of the years 2000 until 2006, 176 articles with a conclusion were found in this journal.

Publications were not as abounding in Psychiatry as in Cardiology and consequently, a second journal was needed to obtain a significant number. The first journal selected was *Psychopathology* with one volume per year containing 6 numbers. From this journal a total of 50 articles with a Conclusion Section were obtained. From the second journal, *Psychotherapy and Psychosomatics*, only 35 conclusions were obtained in spite of having exactly the same volumes and numbers as *Psychopathology*.

With regard to Paediatrics, publications in this specialty were very outstanding in number. *Paediatrics Neurosurgery* presented two volumes a year containing 6 numbers each from 2000 to 2003 and one volume a year from 2004 to 2006. A total of 147 conclusions were obtained from this journal reaching a total number for the English corpus of 408.

Table 3.12 shows the journals selected for our analysis. It also displays the number of conclusions found in each of them as well as the source where they were found:

Specialty	Source	Name of Journal	Number of conclusion
Cardiology	www.charite.de/charite	Cardiology	176
Psychiatry	www.charite.de/charite	Psychopathology	50
		Psychotherapy and Psychosomatics	35
Paediatrics	www.charite.de/charite	Paediatrics Neurosurgery	147
Total number of conclusions			408

Table 3.12: English journals selected for the corpus of the study

Interestingly, with regard to the author’s guidelines of these journals, no Conclusion Section was required under the heading “arrangement”. Nevertheless, many of them included this section as it was the case with the Spanish journals. In fact, in the author’s guidelines, there was no reference to the sections of the papers. The aspects considered in all of them are: Title page, Full address, Key words, Abstract, Foot Notes and Tables of Illustrations.

3.1.1.6. Index Factor of the Journals

As the chapter above indicates, the criteria taken into account for the selection of the journals did not contemplate the Index Factor. In the present chapter an explanation for this choice will be provided.

The English journals selected have a higher diffusion than the Spanish. A means to control the extent to which a journal is spread and cited is determined by the Impact Factor (further referred to as IF). In the last years a growing interest in measuring journals quality has emerged. IF is a result to this new interest that allows a comparison regarding quality among scientific journals. It is one of the most frequently used

parameters for evaluating scientific publication. This factor was created by the Institute of Scientific Information (ISI) a few years ago. With regard to journals, their IF is published annually by means of the *Science Citation Index- (SCI) Journal Citation Report* (further referred to as JCR), which is a recognized authority among the worldwide research community for evaluating journals. The JCR can be accessed via internet under the following address: <http://portal.isiknowledge.com/>. Among other data, such as Immediacy Index, Total Cites, Total Articles and Cited Half-Life or Journal Title, it provides information on the mentioned IF of journals. The IF evaluates the number of citations a journal receives on behalf of other journals, that is, it measures the frequency with which an article has been cited in a journal in a particular year or period. The JCR presents statistical data from 1997 onward that allow a systematic and objective evaluation of the leading journals in the research community. It covers over 7500 of the world's most cited, peer-reviewed journals in approximately 200 disciplines. (JCR is available in two editions: a) *The Science Edition*, which covers over 5,900 leading international journals from the Thomson Scientific database and b) *The Social Sciences Edition*, which covers over 1700.) JCR allows authors to identify journals relevant for their fields and to find out about the status of the journals in which they want to publish. It is therefore also useful for librarians to measure journal usage, for publishers, to control their competitors and identify new publishing opportunities and for editors, to assess the effectiveness of editorial policies and

aims. Finally, it is also useful for information analysts to study citation patterns within and between disciplines.

The IF was not one of the requirements for the selection of articles for this study due to the fact that the Spanish IF was in most cases non-existent or very low. In fact, to illustrate this fact with real data, the JCR showed only 3 journals belonging to the selected specialties for this research work with an IF when the language specified for research was limited to Spanish (namely, *Revista Mental de Madrid* 0.414 and *Actas Españolas de Psiquiatría* 0.286 and *Revista Española de Cardiología* 1.769). If Spanish is selected as a language in the JCR (i.e. giving “Spain” in the option “territory”) and a list from the highest to the least IF is required, only 30 journals are displayed. It is relevant the fact that *Revista Española de Cardiología* is on the fifth place of the above mentioned list. The relevance of the IF of this journal is also pointed out by Abad (2006:6). She provides us with information regarding Spanish medical journals and their IF. She affirms that between 2000 and 2004 only 17 Spanish medical journals had an IF. This factor was less than 1 in 13 of the cases. Out of the 17 journals, the exception to this low factor was one of the two journals with an IF contained in the Spanish corpus of this research work: *Revista Española de Cardiología*. The evolution of *Revista Española de Cardiología* is amazing, for it started with 0,7 in 2000 and rose to 1.802 in 2004. This was the first time a Spanish journal reached such a high IF. In 2005, the IF of this journal was 1.769. For Abad García, its fast growth is due to its spread in the internet and also to the fact that it started to be edited in English language as well.

The other journal with an Impact Factor in 2005 present in the Spanish corpus was *Psicothema*: 0,874.

The situation described above regarding Spanish IF, explains why it was not a requirement for the journals to be included in the corpus, however, JCR allows it in this research work to get an idea of the relevance of the selected journals among the research community.

Table 3.13 represents the IF of the 4 journals selected in English language in 2005 from the highest IF to the lowest:

Psychotherapy and Psychosomatics	4.966
Cardiology	2.092
Paediatrics Neurosurgery	1.049
Psychopathology	0.808

Table 3.13: Example Impact Factor of the English journals of the corpus in 2005 (<http://portal.isicknowledge.com>)

The data on the IF in Table 3.13 above contrasts with that of the Spanish journals. Firstly, the English corpus for this research work is made of only 4 journals whereas 18 Spanish journals were needed to get a representative number of conclusions out of which only one of them was found in the ISI Web of Knowledge, namely, *Revista Española de Cardiología*. These data put Spanish journals in great disadvantage in comparison to English regarding IF. At the same time these data verify the spread of the English language in comparison with the Spanish language regarding publications already stated in chapter one.

3.1.2. Codification of the conclusion sections

Prior to the task of analyzing the lexico-grammatical structures and vocabulary of the two languages and compare them, it was necessary to arrange the material. The first step was the codification of the conclusions. With this purpose, Conclusion Sections were extracted from all the articles found with a PDF format (a total of 719; 311 Spanish conclusions and 408 English conclusions) and were copied in a new file. Two text files were created, one for each language, where all conclusions were included. The file containing the English conclusions of all the specialties was called “English corpus” and the one containing the Spanish conclusions, “Spanish corpus”. The latter contained 104,194 words and the English corpus 58,865. A code was given to every conclusion on grounds of the specialty and journal it belonged to. The first letter of the code corresponded to the specialty the journal belonged to. In this way, “C” was selected for “Cardiology” in the English journals and also for the Spanish journals belonging to “Cardiología”. A “P” was selected in both languages for “Psychology” and “Psicología” and “PD” was used for “Paediatrics” and “Pediatria”.

After the capital letter indicating the specialty, an abbreviation corresponding to the name of the journal was added. The abbreviation was composed of a selection of letters up to 4 taken from the journals’ names (e.g. for the journal *Psychotherapy and Psychosomatics*, the abbreviation used was

“PP”) Table 3.14 and 3.15 show the abbreviations used for the journals:

ENGLISH			
Specialty	Source	Name of the journal	Code
Cardiology	www.charite.de/charite	Cardiology	C
Psychiatry	www.charite.de/charite	Psychopathology	P
	www.charite.de/charite	Psychotherapy and Psychosomatics	PP
Paediatrics	www.charite.de/charite	Paediatrics Neurosurgery	PN

Table 3.14: Abbreviations used for the codification of the English journals in the corpus.

SPANISH			
Specialty	Source	Name of the journal	Code
Psychiatry	fisterra.com	Cuadernos de Medicina Psicosomática y Psiquiátrica de Enlace	CMPP
		Norte Salud Mental	NSM
		Psicothema	PT
		Psiquiatría Pública	PP
		Revista de la Asociación Española de Neuropsiquiatría	RAEN
		Revista de Psiquiatría	RP
	www.doyma.es	Revista de Psiquiatría y Psicología del Niño y del Adolescente	RPNA
	http://dialnet.unirioja.es	Psiquiatría Biológica	PB
		Trastornos Adictivos	TA
		Informaciones Psiquiátricas	IP
Paediatrics	www.fisterra.com	Boletín de la Sociedad Canaria de Pediatría	BSCP
		Boletín de la Sociedad Vasco-Navarra de Pediatría	BSVN
		Foro Pediátrico	FP
		Revista Pediatría de Atención Primaria	RPAP
		Vox Pediátrica	VP
	www.secp.org/SB11/listre y.asp, http://dinarte.es	Revista de Cirugía Pediátrica	RC
Cardiology	www.doyma.es	Revista Española de Cardiología	REC

Table 3.15: Abbreviations used for the codification of the Spanish journals in the corpus

After these combinations of letters corresponding to the specialty and the journal, the volume and year of the journal were indicated. Regarding the year of the journal, only the last

two numbers of the year in question were given (e.g. “01” would represent the year “2001”, and so on.) In order to avoid long codes and too many numbers, pages were not included in the codification. For this reason, articles belonging to the same year and volume were given a letter after the volume. Sometimes, articles published in the same year and in the same volume were over fifteen. As a consequence, the letters after the volume that were being used instead of the page numbers were given in alphabetical order so as to make clear that the codes represented different articles. Finally, a number was introduced at the end of each code in order to enumerate the total number of conclusions found in each journal separately.

To give an example, conclusions belonging to the *Revista Española de Cardiología* were coded as follows: the code starts with a capital “C” standing for the specialty. This “C” is followed by the letters “REC”, corresponding to the name of the journal. Next, the volume is specified with the abbreviation “Vol.” followed by the number of the volume and the two last numbers of the year of publication are given. Finally, conclusions were enumerated in each specialty. Thus, a final number was added to the codification on the basis of the place of the conclusion in the journal *Revista Española de Cardiología*. Therefore, the first conclusion of Cardiology from 2000 taken from the journal “*Revista Española de Cardiología*” was labelled: “C.REC.Vol.1-00-1”, the second “C.REC.Vol1-00-2” and so on.

3.1.3. Lexical Analysis with WordSmith™

With the electronic format of the conclusions it was possible to make use of WordSmith™ Tools. In order to process the texts with this tool it was necessary to turn the articles found in PDF (Portable Document Format) into “txt”, i.e. “plain text”. After this step the aim was to localize technical terms by means of different applications, collocations and standard or recurrent structures used naturally in both languages when writing a medical conclusion in a paper. WordSmith™ also provides us with a statistical calculation of the appearance of certain terms and the way they are used in combination with other word collocations. This tool gives us plenty of examples of a word or phrase in their contexts. In this way it is possible to have a much better idea of the use of a word. The application named Concorder in WordSmith™ helps us to find out the words that normally belong to a certain term. For instance, it can be observed by reading the examples that in academic writing a “paper” normally “describes”, “claims” or “shows” but it does not “believe” or “want”. Therefore, one should avoid sentences such as: “This paper wants to prove that”.

WordSmith™ also provides some statistical data. It shows us the total number of words in the Spanish corpus, which was 104,194 and the total number of words in the English corpus, which was 58,865. Taking into account that the total number of conclusions in English for analysis was higher than the total number of Spanish conclusions, namely 408 and 311 respectively, it is remarkable that the total number of words in

Spanish is greater than in English. This fact gives evidence that Spanish conclusions are longer than English.

WordSmith™ also provides us with a list in alphabetical order which shows all the words contained in the English corpus and the Spanish corpus with information on the percentage of appearance. Using this application called “Word list”, the first step taken was the creation of ten lists, five for each language indicating the most recurrent adjectives, adverbs, verbs, nouns and conjunctions in the two languages. The lists are divided into 3 columns, the first indicates the name of the word, the second the number of times it appears in the texts and the last one the percentage of running words, that is, the percentage it represents considering the English Corpus or the Spanish Corpus. Unfortunately, WordSmith™ does not separate words into morphological categories; therefore it was necessary to look for the different categories in detail. Additionally, the programme cannot separate different tenses, for this reason, it was necessary to make a thorough analysis which turned to be very time consuming. Also with English verbs it was necessary to look up all the different tenses. Again, only the infinitive form was shown in the list created, but this form included all the tenses in the Wordlist tool. Next to the running words it was relevant to add a new column indicating the most frequently verbal tenses used for this category. Regarding adjectives, also singular and plural were included in the list. Finally, it was also necessary to go through all the occurrences of a certain word sometimes in order to determine the morphological function. For this task several grammar books were consulted (e.g. Quirk, 1990;

Sánchez, 2003; Swan, 1984, etc). WordSmith™ also includes an application called “Concordance” that shows all the sentences containing the word in question. Analyzing all the instances of a certain word was often necessary since many words with a high appearance could have different morphological functions and consequently also different translations depending on their function in the sentence. Thus, to illustrate this with an example, the English word “no” can be an adverb with a negative meaning in Spanish as well, but in front of a noun it acts as a determiner and its translation into Spanish changes into “ningún/ninguna”, for instance: “*Whilst no single treatment for the disorder has been shown to be effective, many different approaches have been tried, ranging from medications to different types of psychotherapy*”. The same problems were faced with the Spanish lists. Moreover, Spanish made it even more complicated since also feminine and masculine forms had to be considered for adjectives and nouns and many more verbal forms were to be considered and added to the infinitive form representing them. An example of words having different meaning and morphological function in Spanish is the word “baja” meaning “sick leave” if considered a noun but “low” when functioning as an adjective.

After the elaboration of all the lists, it was necessary to delimit the study of frequencies and collocations, to concentrate on the aim of this research. Therefore, the first 10 words of each list were taken into account in each language, but only those words present in both English and Spanish were considered for a further study focused on collocations. Nevertheless, not all

morphological categories were taken into account for this second part of the study but only verbs, nouns and adjectives. In this way, conjunctions and adverbs were left out. Finding collocations formed by the combination of adverbs and conjunctions with other words was difficult and it would not have provided us with much relevant information regarding fixed or repeated structures in medical texts. However, analyzing the most common in the two languages and those that coincide with a high percentage of appearance is important.

3.2 Survey to investigate medical needs

3.2.1 Survey's rationale

The assumption that led motivated this research work was the idea that Spanish medical doctors need help when publishing in international journals on the grounds of a lack of proficient English.

As no previous literature was available regarding doctor's command of the English language and their needs when dealing with the writing of a RA, a survey investigating these needs was required that would support the importance of the current investigation.

Therefore, a survey was carried out in which medical doctor's competence of the English language was checked as well as the course of action they take (their writing habits) when they want to publish in an international journal in English.

3.2.2 Targeted audience

However, the current survey is addressed to Spanish and German doctors with the idea to assess whether there are differences between them regarding their level of English and consequently the difficulties they come through when publishing in international journals. To ascertain this and to see what factors could explain these differences, a survey was elaborated. Among the different methods available to carry out a survey in a certain population to consider a certain aspect, the following methods can be underlined: personal interview, phone interview, mailing and Internet. After considering that each of them has its advantages and disadvantages the survey was handed in personally in the case of Spanish doctors and sent via internal post in the case of German doctors. The reason for doing it by means of internal mail in Germany, concretely in Berlin, was that by the time the survey was carried out there was a high possibility of missing many doctors' collaboration since it was during the months of June, July, August and September 2006 and many of them were on holiday. Internal mail was the only safe way to make sure they received the survey. In the case of Spain, the survey was conducted in Valencia, so there were no time restrictions to contact the professionals.

To avoid any possible kind of selection bias that could invalidate the results, the surveys were handed in to the heads of department in the different medical specialties. They were in charge of providing medical doctors with information on the objective of filling the requested survey. Participation was

voluntary and it was carried out randomly to avoid selection bias of the medical population interviewed. Additionally the interviewing process was anonymous so as to increase the rate of answers. A mail box was made available to return the filled surveys in the different medical departments.

As a first step, hospitals in Valencia and Berlin were contacted to check out the number of specialists in the 3 selected specialties. All the hospitals where doctors worked were University hospitals. Spanish doctors were working at the time of the survey at the University Hospital La Fe of Valencia, and at the University Hospital Dr. Peset of Valencia. Regarding German doctors, they were working at the University Hospital Benjamin Franklin in Berlin.

The data showing the number of specialists by the time of the survey are indicated in Table 3.16 below:

Hospital	Cardiologist	Paediatrics	Psychiatrist
Spain			
University Hospital La Fe	30	71	7
University Hospital Dr. Peset	9	14	6
Germany			
University Hospital Benjamin Franklin	70	120	30

Table 3.16: Hospitals involved in the study and number of specialists in each specialty at the time of the performance of the survey

According to Table 3.16, Paediatrics is the specialty with the highest number of specialists followed by cardiology and psychiatry in all four hospitals. German participation was lower than Spanish. Although a total of 100 surveys were sent, only 44 were received for analysis. Regarding Spanish participation, over 100 were handed in, but only 68 were available at the end

for analysis. This makes a total of 112 doctors interviewed. As shown above, not all specialties had equal representation, thus it is not surprising that most collaboration came from paediatrics with 50 surveys, then from cardiologists with 27 and finally from psychiatrists with 11.

As the survey was conducted in two different countries, questions had to be adapted to the different cultures. For this reason, some questions had to be more specific in one of the two languages and sometimes even a different question or an added question was necessary. This adaptation was mainly needed in the first section of the survey dealing with personal details where questions regarding the school system, education system, hospital posts and the like were addressed.

It was my concern to inform doctors with an introductory paragraph in the survey, that it was anonymous so that they did not feel intimidated.

3.2.3 Survey design validation

As no validated survey was found in the consulted literature in order to assess doctors' level of medical English, the information in Martín's article from 2004 describing the design of this survey was followed. Martín (2004:23) defines survey as follows:

Instrumento utilizado para la recogida de información, diseñado para poder cuantificar y universalizar la información y estandarizar el procedimiento de la entrevista. Su finalidad es conseguir la comparabilidad de la información.

According to Martín (2004:24-28) there are several factors to be taken into account when elaborating a validated survey, which are shown below:

- a) Prior to the elaboration of a survey it is crucial to have a clear concept of the aspect aimed to be measured. Therefore, the first step would consist in the definition of the mentioned aspect for which the help of experts may be needed to get informed on the subject matter or existing literature can be consulted for this purpose as well.
- b) To be provided with the necessary computing statistical tools in order to evaluate data with scientific rigour.
- c) To ensure that validity and reliability criteria can be applied to the survey that has been designed. This implies that the survey should present the following features: it should be adequate for the problem to be measured; it should be able to measure accurately those characteristics aimed to be measured and not others. Thus, it should be reliable and accurate, i.e., presenting a minimum measurement error.
- d) The population targeted
- e) The way to reach the population
- f) The formal aspect of the survey.
All these aspects will determine the formulation of the questions.
- g) Each item has to be defined in an exhaustive and excluding manner.

- h) To make questions that can be clearly understood by the interviewed subject. It is important to make sure the content of our questions will be easily understood and the subject will accept it as valid and adequate.

Taking these considerations into account, Martín provides us with some recommendations to create a survey:

- a) Short questions are preferred.
- b) Avoid using the interrogative “why”
- c) Do not make negative questions
- d) Avoid questions where candidates have to make an effort to remember the requested data
- e) Do not make questions containing alternatives that could seem appealing to the candidates and that may induce them to give one answer rather than another.

As far as the answers of a survey are concerned, Martín (2004:25-27) states they can be classified into three groups:

- Answers to questions where only options are given (yes/no or true/false)
- Answers to questions in which several options are given
- Answers where gradual options are given either in terms of numbers or verbally. For instance, answers in which candidates have to circle a

number from one to 10 according to their agreement on a certain topic.

- f) Finally, Martín (2004:28) underlines questions have to be given in a logical order.
- g) Once the definitive draft has been designed, it is recommended to carry out a pilot study in which a group of subjects similar to the one that is going to be actually interviewed answers all the questions contained in the survey. The goal is to check answers' reproducibility and to remove those that can lead to errors in the results.

Taking into account all these considerations a survey was performed following the above consulted literature.

3.2.4 Structure of the survey

The survey was divided into 4 parts. The first part dealt with personal information, the second with access to medical journals on the part of the doctors, part 3 had to do with publishing papers and the last part with style sheets.

The total number of questions in the survey is 25. However, questions in the first section, which show information on personal details, were not given a number. For this reason in the survey only 17 questions are numbered. The first section contains eight questions, the second section contains two, the third has eleven and the last section four.

Figure 3.1 Spanish survey

3.2.4.1 First Section: personal information

Regarding personal questions, the first information required was the doctor’s specialty. All the doctors interviewed belonged to one of these 3 specialties: Psychiatry, Cardiology and Paediatrics as shown in Figure 3.2 below:

Figure 3.2 Medical Specialty

These specialties were the 3 selected for the corpus because the greatest amount of conclusions was found in these disciplines. The reason to include this first question was to see whether differences take place depending on the specialty

doctors belong to. The second question in this section was the age of the doctor.

Edad: _____

Figure 3.3 Age of medical doctors

As shown in Figure 3.3, four options were given concerning age: between 20 and 30, between 30 and 40, 40-50 and more than 50. The aim was to see if differences in English level were associated to age but also to gender. Consequently, in the third question doctors were asked to indicate whether they were male or female.

Other aspects that could influence the doctors' knowledge and experience with English publication was their experience in publishing which could be related to two additional factors asked: the year when they obtained their degrees (see Figure 3.4) and the number of years they had been working as indicated in Figure 3.5 below:

Año de licenciatura: _____

Figure 3.4 Year in which doctors finished their studies

Años de ejercicio profesional: _____

Figure 3.5 Years doctors have been working

But not only these two factors were asked, as Figure 3.6 indicates, they also had to specify the highest certificate they had obtained to evaluate their experience as researchers; i.e., if

they had only a degree or if they had continued their studies and obtained other certificates such as Masters degree or Doctorate.



Figure 3.6 Academic Level : highest certificate obtained by the doctors

As it was mentioned above, the fact that German is a Germanic language unlike Spanish, led me to predict the former doctors would find it easier to read and write in English. However, there were other factors to be considered that could contribute to a faster or better learning of the language that would explain differences of level between the two countries. For instance, most schools and high schools in Spain teach only one foreign language at the present time, and a few years ago many of them taught French rather than English or they did not teach a second language. The situation could be different in Germany, where apparently more languages were and are taught. Research shows that L3 learners are more effective learners of the target language than are L2 learners. A third language learner is a person who has already acquired two languages and is learning a third one. The more languages one knows the more ready is his or her mind to learn other languages. As a consequence, Figure 3.7 shows the importance to find out if there were any differences with respect to the languages taught at school and high school in the two countries.

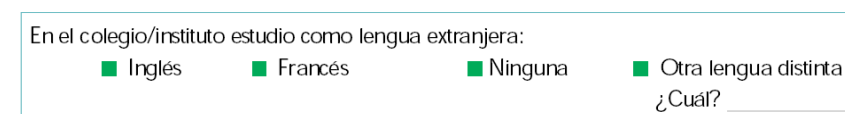


Figure 3.7 Language studied at school and high-school

Taking into account L3 learner's facilities, information was sought to see if doctors did not learn a second language at school or high school at all, whether they learnt only one: English, French or any other, or whether they learned more than one second language. Presumably, German doctors would have learned more languages than Spanish doctors due to the different norms regarding school systems and the teaching of foreign languages in the two countries both in the past and currently.

Finally, another aspect that could influence the language level of doctors is if they had stayed for a period of time in a foreign country where they had to learn another language. Obviously, a stay in an English-speaking country would improve their level of English considerably; however, any country where doctors had stayed for a period of time was of my interest, since they would have been forced to speak a different language anyway. Figure 3.8 shows the question formulated to obtain the aforementioned information:

¿Ha realizado alguna estancia en el extranjero por motivos laborales?			
<input type="checkbox"/> No	<input type="checkbox"/> Sí	¿En qué país? _____	Duración de la estancia _____

Figure 3.8 Working in a foreign country

The last question included in this section was a direct one in which doctors had to evaluate their level of English themselves taking into account different language skills: writing, reading and speaking.

Nivel de conocimiento del inglés:				
Escrito:	<input type="checkbox"/> Nulo	<input type="checkbox"/> Bajo	<input type="checkbox"/> Medio	<input type="checkbox"/> Alto
Hablado:	<input type="checkbox"/> Nulo	<input type="checkbox"/> Bajo	<input type="checkbox"/> Medio	<input type="checkbox"/> Alto
Comprensión lectora:	<input type="checkbox"/> Nulo	<input type="checkbox"/> Bajo	<input type="checkbox"/> Medio	<input type="checkbox"/> Alto

Figure 3.9 Knowledge of the English language

As indicated in Figure 3.9, the options they were given to answer range from no knowledge at all to a high level. However, the answer to this question could not be efficiently interpreted since answers could be quite subjective. For this reason the language level had to be discovered by means of other questions contained in the second part of the survey.

3.2.4.2 Second Section: access to medical journals

In the second section one of the main objectives was to discover what means of information doctors use in order to be updated in their fields, i.e., if they use on-line papers in their own language or in English or if they use any other means to get informed.

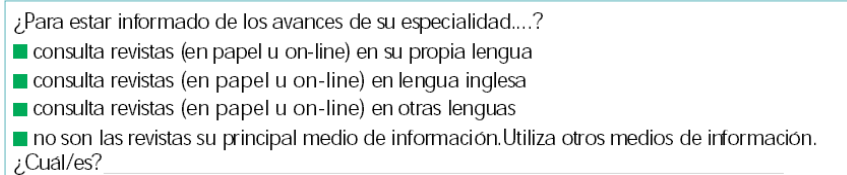


Figure 3.10 Means doctors use in order to get updated in their specialties

Figure 3.10 would already give an idea of the percentage of those who read in English. The aim was to see to what extent doctor's level of the English language allows them to use sources of information in English. Therefore, a second clarifying

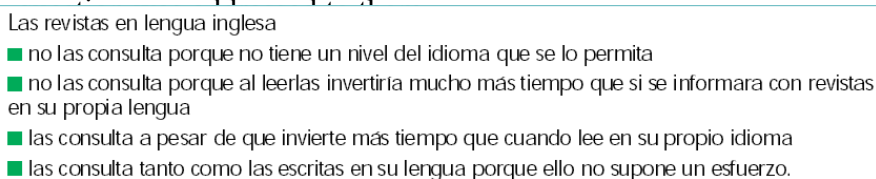


Figure 3.11 Capacity of the doctors to consult English journals

As pointed out in Figure 3.11, doctors had to indicate whether they read English journals without difficulty or whether they read them despite considering this a more time consuming task on the grounds of a lack of the language command. They were also given the option of answering that they admitted not reading them either because their level did not allow them to do so or because in this way they avoid making an extra effort and investing much more time than if they read them in their mother-tongue.

3.2.4.3. *Third Section: publishing papers*

By means of the mentioned questions data regarding reading skills was covered, so the next step was to find out about their writing skills. Therefore, the next question was addressed with the aim of finding out how prolific they were both internationally and nationally:

3 ¿Ha publicado alguna vez un artículo en una revista de medicina?

- No, nunca. (Por favor, pase directamente a la pregunta número 14)
- Sí, en revistas nacionales. (Si ha marcado sólo esta opción, por favor pase directamente a la pregunta número 14).
- Sí, en revistas internacionales de lengua inglesa

Figure 3.12 Doctors' publications in medical journals

With Figure 3.12 the aim was to discard that percentage who had not contributed to the writing of a paper in English. However, the possibility of being included as an author in a paper but still not having taken part in the writing of the paper

but in any other possible activity that made the publication possible should be taken into consideration. For this reason, a more accurate question was needed to determine the number of doctors who had participated in the writing part. A new direct question was stated asking whether they had actively taken part in the writing process of an English paper or not:

4 ¿Ha participado activamente en la redacción de algún artículo en inglés?
 Sí No (Por favor, si ha elegido esta opción, pase directamente a la pregunta número 14)

Figure 3.13 Doctors active participation in the writing process of research articles

Figure 3.13 was a key question because from this moment on, surveyed people were divided into two groups; those who had never taken part in the writing process of an English paper journal and those who had. The former were asked not to answer questions from 5 to 12 of the survey and to continue with question number 13. Those who had participated in the elaboration of a written paper in English continued with question number 5, which was again a question separating one group from another.

5 Si publica en una revista internacional en lengua inglesa
 Redacta el artículo primero en español
 Redacta el artículo directamente en inglés
 En la actualidad, ¿opina que su nivel de inglés es suficientemente bueno para redactar los artículos y enviarlos a una revista sin necesidad de una posterior corrección por una persona especializada?
 No (Pase a la pregunta 13) Sí (pase a la pregunta 13)

Figure 3.14 First language used by doctors when writing for an English publication

The objective of Figure 3.14 was to find out the way they proceed when they are faced with the task of writing a paper. The question was what language they use when writing a paper, that is, whether they write directly in English or if they write in their mother-tongue first and translate into English later on.

It is interesting to see what percentage need to be provided with English structures and how many of them need to be guided in the way they write in their mother-tongue first. This latter group were addressed all the questions in the section, since they are related to the way doctors summarize in Spanish or German, prior to the translation.

The first question addressed to doctors writing first in their mother tongue, that is, question number 6 in section 3 (see Figure 3.15 below), was intended to ascertain to what extent they are aware of the effect that their writing in their mother-tongue will have on later translation (Navarro and Barnes, 1996:302):

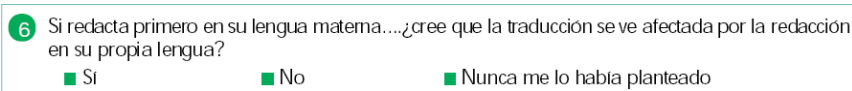


Figure 3.15 Influence of the use of the mother-tongue on a posterior translation

Checking if they actually think about the fact that what they write has to be translated into another language was aimed at making them aware of this fact in order to smooth the translation process. This could be possible by means of simple structures rather than complicated language or long sentences. For this reason in question number 7 (see Figure 3.16 below) of this section, doctors are asked if they use, for instance, a simple language or short sentences, or in other words, plain Spanish or German with the purpose of turning the task of translating easier for themselves or for the professional translator.

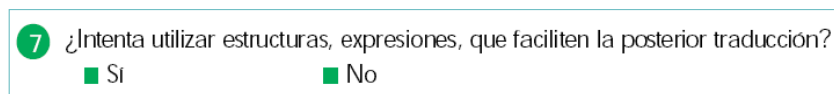


Figure 3.16 Using simple structures with translation in mind

The next question, number 8, is especially interesting because it could provoke a change of attitude on the part of the interviewed person. In Figure 3.17, doctors are asked if they are aware of the time and money cutback that simplifying the language would imply thanks to a more accelerated translation process.

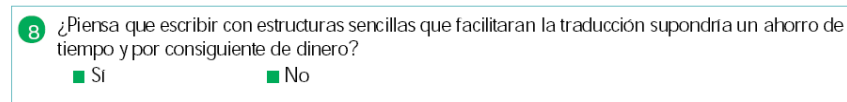


Figure 3.17 Economical advantages and time saving when using simple structures with translation in mind

If they became aware maybe a willingness on the part of the writer to improve his first writing in his/her mother tongue could be possible.

As a consequence, in questions 9 and 10 doctors have to say if they would consider it worthwhile to learn how to write in an easy way with translation in mind (see Figure 3.18)) or if on the other hand they consider this option as rather complicated and consequently a vain effort (See Figure 3.19).

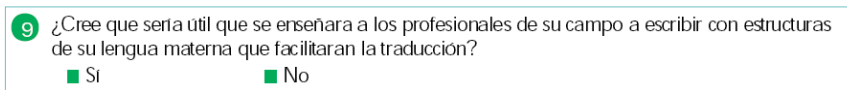


Figure 3.18 Teaching medical researchers to use simple structures that would facilitate translation

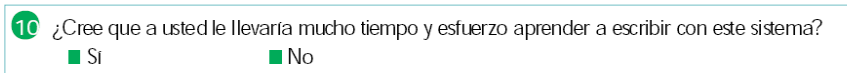


Figure 3.19 Opinion regarding personal skills to learn the use of simple structures

With these questions, the purpose is to find out if doctors would be willing to make use of some simple structures they could be provided with, after studying equivalences between their language and English because they understand these simple

predetermined structures would simplify the translator's task considerably.

After a rough draft in their mother-tongue it was important to find out who translates into English. Data on how many of them carry out the translation on their own and how many turn to a professional translator, agency of translation or even to a native speaker were sought as Figure 3.20 illustrates:

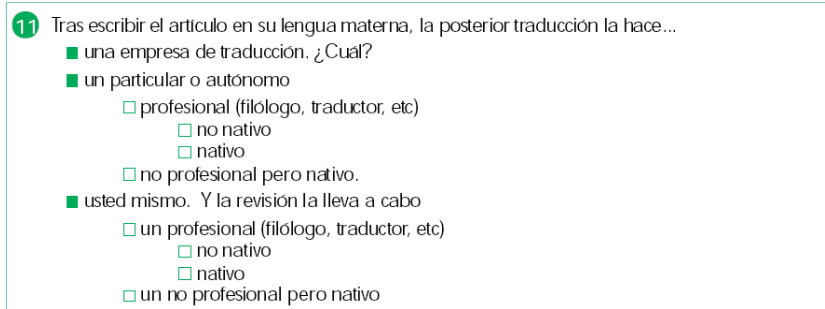


Figure 3.20 English translator of articles that have been written in the mother-tongue

Another aspect of concern was the success of the translation. It was investigated if those who send their articles to be translated have ever been faced to any problems. For instance, if the transference of communication is always successful or if there has ever been a misinterpretation of some part or sentences on the part of the professional translator.

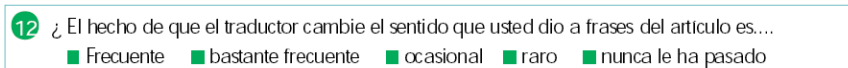


Figure 3.21 Misunderstandings on the part of the translator of the article written by doctors in their mother-tongue

The question that arises taking into consideration Figure 3.21 is to what extent this misinterpretation could be associated to a poor or not appropriate writing in the source language. In this case, providing doctors with a guideline on how to write,

what style and structures to use could avoid this kind of problems.

To conclude this third section, question number 13 was addressed to all doctors having published in English, both if they write first in their L1 and then translate or if they write directly in English. With this question the objective was to find out about doctors' success in their first delivery, that is to say, how well their first try for publication is written.

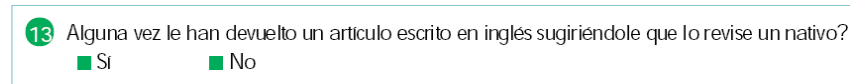


Figure 3.22 Rejected articles for publication due to incorrect use of the English language

For this reason, in Figure 3.22, information was required on how many of them have been rejected one of their intended articles for publication with a suggestion on the part of the referees to write them “in a native style and appropriateness”.

3.2.4.4. Fourth Section: style sheets

As for the last section, number 4, the professional's awareness of the existence of style sheets in their fields was to be assessed. Questions number 14 and 15 (See Figures 3.23 and 3.24 below) deal with this aspect, in fact they are required to write the name of any of them in question number 15.

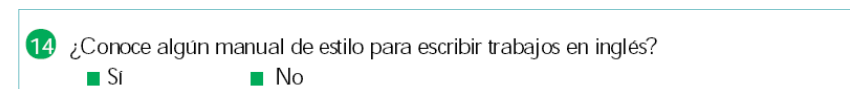


Figure 3.23 Doctors' knowledge about the existence of style sheets

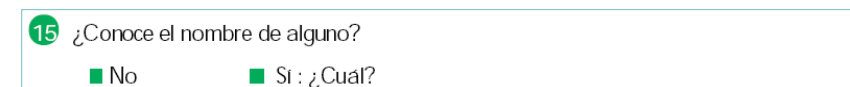


Figure 3.24 Name of some style sheets

This information was required to assess to what extent doctors are committed to their obligations as researchers and publishers of international journals and strive to get as much information on the guidelines of a certain journal as they can prior to writing a paper to succeed.

Finally, as shown in Figure 3.25, another issue of concern is whether there is a language center in the doctors' places of work to assist them when faced with translation problems.

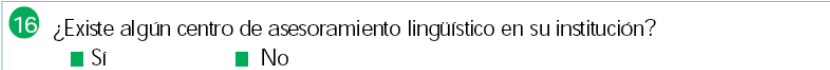


Figure 3.25 Existence of a linguistic supporting centre in the doctors' institution

An additional purpose shown in Figure 3.26 is to get information on the doctors' interest, i.e. if they would consider the existence of such a center useful and if they would make use of them in case one was available.

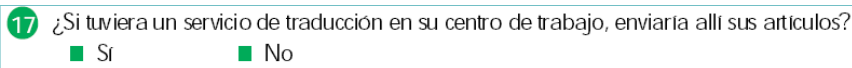


Figure 3.26 Willingness on the part of doctors to make use of linguistic supporting centres

Despite the detailed elaboration of the survey, not all the questions were included in the research but only those needed for analysis to reach the needed results.

3.2.5 Transcription to SPSS statistical system

Data were analyzed with the programme Statistical Package Social System (further referred to as SPSS), by means of which various possible answers to the same question were

contemplated. A graphic representation by means of 12 tables was included to show the results.

Thus, a total of 12 questions were entered in the programme SPSS under the option of “vista de variables”. The variable view provided by SPSS to enter data is limited and strict norms have to be taken into account: no spaces are allowed between words, additionally no signs are allowed except for the full stop. For this reason, the information contained in the questions has been consequently reduced to a few key words representing the whole question. This reduced content standing for the whole questions is what the programme later displays when it shows charts with results. The total number of variables, that is, the total number of questions analyzed was 14 out of the 25 contained in the survey. As mentioned before, questions in section number one, were not numbered, however, they were given a number when commented for analysis. Naturally, the number of the variables introduced in the programme SPSS for analysis is not equivalent to the numbers of questions in the survey. Therefore, numbers in the survey should not be taken into account, the following numbering corresponding to data introduced in the SPSS programme are the ones to be taken into consideration. The variables were introduced with the following reduced labels:

From the first section: personal details, the following were selected:

- 1) Medical. Specialty and Age
- 2) School.L2

3) Studies. Abroad

From the second section: Access to medical journals information, the following variables were selected:

1) English Journals

From the third section: Published papers, the variables selected are:

1. Writing. English
2. First writing language
3. Translation. Task
4. Translation. Misunderstandings
5. Translation. Consequences
6. Saving Time and Money
7. Facilitating. Translation
8. Rejected. Articles

From the fourth section: Style sheets:

9. Language Center

In Table 3.17 the first variable analyzed was the specialty of the doctors. The variable was labelled “Medical Specialty”. To enter the answer given by doctors in the programme, all possible answers had to be specified in the option called “valores” or “values”. Numbers were given to each possible answer starting from 0. To illustrate this with an example, in the case of indicating the medical specialty, Psychiatry was given the value “0”, Cardiology was represented with the number 1 and Paediatrics was numbered with a 2. In this way, the tool provides us at the end with the percentage of

0, 1 and 2, representing the total number of doctors in the 3 specialties respectively.

Also “Age” was studied in this first variable. Here, values were divided into 4 categories and numbered from 0 to 3. Thus, 4 groups were differentiated: doctors between 20 and 30 years old were represented by the value 0. The value 1 was used for the age of 31 to 40, 2 for those between 41 and 50 and 3 for doctors over 51.

Table 3.19 shows the language studied at school or high-school by doctors, which was introduced in the programme with “School.L2”, an abbreviation standing for “second language learnt at school”. Up to 6 values were necessary to analyze this subject. Those who studied English as a second language were represented with number 0. For those who studied French number 1 was introduced. Doctors who did not study a second language at school were represented with number 2. The fourth option was for those who studied French plus any other second language and it was given a 3. The fifth option was English and French, represented with number 4. For the group of doctors that studied English and any other language different from French the number was 5. The sixth and last option was for those who learnt English and two or more other languages and they were given the number 6.

It is worth mentioning at this stage that in Germany, students of medicine had to study until very recently what is called “Latinum”. This is a 3-year (kleines Latinum) or 5-year course (grosses Latinum) of Latin that is required for certain studies in which Latin plays an important role, and this is clearly

the case of Medicine. Obviously, unlike Spanish, German is a Germanic language with no similarity in the root of its words to technical words in medicine, which come from Latin. Taking into account, that Spanish is a Romanic Language and that in Spain only one year at high school is obligatory for those who selected the option of science at the age of 15, i.e. for all the current doctors, it was considered that those German doctors who had a Latinum could be said to have studied one additional language. The option of French was mainly thought for Spain, since a few years ago (in the 1970s) French was more frequently taught at schools than English in the aforementioned country. Regarding the option of English and other languages, it was thought specially for those German doctors who had studied a different language apart from the Latinum, which in most cases was Russian on the grounds that Berlin was after the Second World War divided into 4 parts and many people lived in East by the time they studied, where Russian was taught.

As for the variable number 3 dealing with whether doctors had stayed in a foreign country for professional reasons, only two values were taken into account, i.e. 0 if doctors had never worked in a foreign country and 1 if they had been abroad. Although they specified the country they had been to in the survey, as already explained further above this information was not included in the programme, though results were commented anyway afterwards.

In section number 3, the answers were simplified. Most variables were only given two values “0” if they answered “yes”

and “1” if the answer was “no”. This was the case of variables: 5, 9, 10, 11, 12, and 13.

Variable number 5 labelled “Writing English” stands for the active participation of the doctor in the writing of an English paper. The label “translation consequences” in variable 9 relates to whether doctors think that the translation into English could be affected by the first draft in their own language and thus, the consequences it may have.

The next variable is closely related to this one, namely, “Saving time and money”, which is the name given to variable 10. It tries to find out doctor’s awareness on the fact that both time and money would be saved if they used easy structures in their own language that were easily translated. Also in accordance with the last two variables, variable 11, “Facilitating Translation”, represents the agreement or disagreement on the part of the doctors on the convenience of teaching doctors the mentioned easy structures so as to facilitate the coming translation. Number 12, labelled “Rejected Articles”, aims at finding out whether doctors have ever had an article rejected or not by a referee.

The last variable belongs to the last section of the survey. It is number 13, “Language Center”, and it assesses the willingness or lack of willingness on the part of the doctors to send their articles to a language center if they had one of these in their institutions.

The next group of variables does not present a two -way value answer consisting of “yes” or “no”. Variable number 6 corresponds to section number 3. This variable is related to

number 5 where the intention was to find out whether doctors actively participated in the writing task of an article. Similarly to variable number 5, with variable 6, also a clear cut dividing doctors into two groups is expected. It is labelled “First Writing Language” and it expects the doctors to inform on the language they normally use when they start to write an article. Consequently, two values are given, number 1 for doctor’s writing first in their mother tongue and number 2 for those who write directly in English. Variables 7 and 8 are only addressed to those who write their draft first in the own language and need a posterior translation. In these variables, different values are given as possible answers. In variable 7, “Translation Task”, 3 values are needed to indicate who the person is carrying out the translation task. The first value is a translation company, the second a freelance and the last one the doctor himself. Although in the survey doctors were asked to give more details, only these 3 options were analyzed with the programme. In fact, they were asked to specify if the freelance translator was a professional or just a native speaker and in case they were professionals whether they were a native or a non-native speaker. The same inquiry was addressed to those who do the translation themselves with respect to the person who reviews the articles. Although initially all the information seemed attractive for analysis, it was then discarded later on because the data were not considered relevant for the actual results expected from the survey.

The last variable to comment is variable 8 “Translation Misunderstandings”. The idea was to find out how often

misunderstandings occur due to a misinterpretation of the original text on the part of the translator. This may happen due to the translator's lack of knowledge or command in the field of medicine. 5 different values were given to this variable: frequently, which was value number 0, quite frequently, which corresponds with value number 2, occasionally, with variable 3, seldom, reflected with variable 2 and variable 1, which shows those doctors who never came across such a problem.

3.2.6 Results of the survey

In order to calculate the results Student's test was used to assess the mean differences in continuous variables between independent groups. A chi square test was used for dichotomized variables, i.e. those where answers could only be "yes" or "not", in order to compare differences in percentages. Those variables in which several answers were possible (called polychotomic answers) as well as those were a graduated scale was given (either by means of numbers or words), were calculated in percentages. A bilateral P value < 0.05 was considered statistically significant. All analyses were calculated using the Statistical Package for Social Sciences (SPSS, version 14 for Windows).

Table 3.17 shows the results concerning the number of Spanish and German participants according to their corresponding medical specialties as well as their age and sex.

	SP n=67	GER n=44	Total
Paediatrics	40	18	58
Cardiology	21	16	37
Psichiiatry	6	10	16
<i>Age (X± SD)</i>	43±14	44±11	
Men	41	34	75
Women	26	40	66

Table 3.17: Spanish and German doctors belonging to each specialty. Number, mean age and sex

The total number of surveys for analysis was 111, out of which, 67 were from Spain and 44 from Germany. The highest collaboration in this survey was attained thanks to Spanish paediatrics, who answered a total of 40 surveys, which corresponds to a 44.4% of the total collaboration. Also German paediatrics' participation was the highest with a total of 18 doctors, which corresponds to an 11%. The least collaboration was that on the part of psychiatrists, who were only 6 in Spain representing a 6.7% and 10 in Germany representing a 5.6%. As for cardiology, 21 Spanish cardiologists were interviewed; a 24.4% and only 16 German cardiologists meaning a 7.8% answered the survey. This result is a consequence of the different representation of doctors working in the different specialties. Thus, the number of paediatrics was higher than the number of cardiologists and the latter higher than that of psychiatrists.

Table 3.17 also shows data on age. Spanish doctors were aged 43 +- 14 years and German 44 +- 11. An additional Table, Table 3.18 was added to give more detailed data, which show that the greatest group of Spanish doctors was over 50 years old; in fact this group represented 42.6%. The second group in terms

of size was that of doctors between 31 and 40 years old, who represented 27.9%. A total of 23.5% indicated those whose age was between 20 and 30, and finally, the smallest group with only 5.9% was that of doctors between 41 and 50 years old. These data imply an advantage in relation to the information searched with the survey in the sense that, most Spanish doctors surveyed were over 30. This fact makes it more likely for them to have published in contrast to those who are in their twenties and graduated recently. As for the German doctors surveyed, the greatest representation was that of doctors between 31 and 40 with a 36.4% of them followed by those over 50 with 31.8%. As for the group of doctors between 41 and 50, there was a higher representation than in Spanish doctors, namely 27.3%. Finally, the smallest group showed doctors between 20 and 30 with only a 4.5% representation. Therefore, data indicate that the vast majority of German doctors were over 30 as well.

Age	20-30		31-40		41-50		>50	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
SP	16	23.5	19	27.9	4	5.9	29	42.6
GER	1	4.5	8	36.4	6	27.3	7	31.8

Table 3.18: Spanish and German doctors' age according to decades.

Regarding the second variable analyzed, “Second language learnt at school and high school”, as expected, Table 3.19 shows considerable differences between German and Spanish doctors.

	English		French		Non		French and another	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
SP	29	42,6	25,0	36,8	2,0	2,9	0,0	0,0
GER	2	4,5	0,0	0,0	0,0	0,0	2,0	4,5
	English/French		English/Another		English/Other			
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%		
	9,0	13,2	3,0	4,4	0,0	0,0		
	0,0	0,0	4,0	9,1	36,0	81,8		

Table 3.19: Second Language learnt at school and high-school by Spanish and German doctors

In Table 3.20 it is shown that most Spanish doctors studied either English (42.6%) or French (36.8%) at school, and there are even some of them who did not study a foreign language (2.9%).

This contrasts considerably with German doctors, the vast majority (90.9%) of whom studied an L2 (91%) and even an L3 (81.8%) at school, that is, not just English but they also studied Russian, Chinese, Latin or Spanish. On the other hand, only 12% of the Spanish doctors studied another language apart from English. Thus, Spanish doctors who did not study English at school have consequently a lower command of the language at the present time. Data show that 39.7% of the Spanish doctors did not study English at school, while only 4.5% of the German doctors did not study English. Moreover, this 4.5% of German doctors studied Latin and another language other than English, whereas most of the Spanish doctors representing the above 39.7% studied French or even no other language. So it can be concluded that Spanish doctors have more difficulties in the learning process of the English language because learning a second language at an early age is much more successful than learning it as an adult (Kennedy, 1988).

In Table 3.20 percentages regarding doctors' stay abroad is displayed.

	YES		NO		chi- square
	<i>n</i>	%	<i>n</i>	%	
SP	21	30.9	47	69.1	p< 0,006
GER	28	63.6	16	36.4	

Table 3.20: Spanish and German doctors' stay abroad for working reasons

The third variable examined (see Table 3.20), indicates that 63.6% of the German doctors have been working abroad but only 30.9% of the Spanish doctors have. Additionally, another aspect that might have played a role in the unequal command of the English language is the different destinations they chose for their stay. To illustrate this with a clear example, although data were not shown in the Table, the following list shows the different countries where German doctors have stayed to work for a period of time: The U.S.A. is placed at the top of the list, but also other places were mentioned such as Hong Kong, Ireland, Netherlands, England, France, South Africa, Hungary, Brazil, Italy, and Sweden. It is worth mentioning that the U.S.A. was indicated in 50% of the cases and that also other English speaking countries such as England, South Africa and Ireland were included. As for Spanish doctors, the places they have been to in order of frequency are: France 24% (where doctors only stayed a few months) the U.S.A 19% (where people stayed for periods of one week, 1, 3, or 6 months), Germany (14%), England 14% (with stays of 3 months) Belgium 9.5%, Italy, Sweden, Panama, Costa Rica, Canada and Holland. These results show a tendency on the part of Spanish doctors to travel

to France in the first place but also the USA, Germany and England show a relevant representation. However, what actually makes a difference is the fact that only 9.5% of Spanish doctors interviewed worked in the foreign country for a period of one year or more. All the rest only stayed for a few months or just for a few weeks. On the other hand, 50% of German doctors stayed for one year or even longer. These figures give evidence that the latter have been much longer in contact with other languages.

Regarding the way doctors get updated in their specialties, results reveal the hypothesis that the high command of the English language on the part of German doctors favors their consultation of English journals. In Table 3.21 it is observed that German doctors consult English journals as often as German journals because it does not imply an extra effort for them.

	No consulting no level		No consulting more effort	
	<i>n</i>	%	<i>n</i>	%
SP	3	4.4	6,0	8.8
GER	0	0,0	2,0	4.5
	Consult despite investing more time		Consult because no effort	
	<i>n</i>	%	<i>n</i>	%
SP	33,0	48.5	26,0	38.2
GER	12,0	27.3	30,0	68.2

Table 3.21: Consultation of English journals by Spanish and German doctors to get updated in their specialties

They affirm that they do not need to devote more time to the reading of the English journal in 68.2 % of the cases. Only 27.3% consult them despite spending more time than when they read them in their own language. Thus, 95.5% read English journals to be updated and only 4.5 % report not reading them

because they would invest more time. Data give clear evidence that Spanish doctors admit having more difficulties when reading English journals. In fact 4.4% of them say they cannot read them due to their lack of knowledge of the language and 8.8% do not read them because they would need more time than when reading in the L1. The most outstanding group, almost half of the Spanish doctors, read them in spite of investing more time and effort (48.5%) and only 38.2% consider reading in English and Spanish equivalent in terms of difficulty.

Variable number 5 is the starting point of the third section of the questions in the survey. In this part of the survey, variables 5 and 6 give rise to a change in the percentages to be analyzed because they divide the total of doctors interviewed into groups. Variable 5 made a distinction between those who had published in English and those who had not. Those doctors who had not published did not answer the following questions in the survey and had only to answer two more of the variables, namely variable 12 and 13. This implied a smaller representation of one of the groups for questions 7 to 13, but this reduction was accentuated with variable 6. In this one, doctors who belong to the English publishing group were again divided into two. One group corresponded to those who first write in their L1 and then translate into English, and group two was made up of those who directly write their articles in English. Therefore, an important aspect in terms of statistics has to be taken into account: with these two variables, 5 and 6 the German sample decreased with respect to the Spanish, since from this moment on only 9.1% of German doctors in comparison to

58.8% of the Spanish could continue with the questions of section number 3.

	YES		NO		chi- square
	<i>n</i>	%	<i>n</i>	%	
SP	28	41.2	40	58.8	p< 0,001
GER	40	90.9	4	2.1	

Table 3.22: Active collaboration of Spanish and German doctors in the writing of an English paper

To illustrate this with actual data, Table 3.22 (variable number 5), with the label “Writing English”, shows that 90.9% of German doctors have actively taken part in the writing process of a paper, whereas only 41.2% of the Spanish doctors have. According to these results German doctors have a much higher representation in the medical discourse community than Spanish doctors. This fact could be explained on the grounds that German doctors find it easier to write English papers.

	Sp/Ger		English		chi- square
	<i>n</i>	%	<i>n</i>	%	
SP	25	89.3	3	10.7	p< 0.001
GER	4	10	36	90	

Table 3.23: Spanish and German doctors’ first writing language when publishing an English paper

This hypothesis is clearly evidenced in the next variable, number 6, which is labelled “First Writing Language”. Table 3.23 above provides us with the information that 90% of the German doctors write their papers directly in English whereas only 10, 7 of the Spanish do. They feel undoubtedly more comfortable writing first in their L1. The results obtained in this variable, motivates the elaboration of the present research work in the sense that Spanish doctors could be definitely supported

with recurrent structures that would help them write English conclusions in papers more easily.

Regarding Table 3.24, the differences in percentages between both groups were highly significant. Most of the German doctors, i.e. 90%, affirm that when writing a medical R.A. they write it directly in English. Since variables 7 to 12 were only addressed to those doctors who write first in their mother tongue, the representation of the two countries for these questions was of nearly 90 % for Spanish doctors but only of 10% for German doctors. For this reason, it is important to bear in mind for variables 7 to 11 that the German representation cannot actually be considered representative of the whole group of German doctors because percentages refer to a minority standing just for 10%.

	English Professional		The doctor him/herself		chi- square
	n	%	n	%	
SP	14	50	14	50	n.s
GER	0	0	4	100	

Table 3.24: Person in charge of the translation into English when papers have been written first in Spanish or German

The results obtained in variable 7, with the abbreviation “Translation Task” are especially relevant. In Table 3.24 above, it can be observed that from the little percentage of German doctors who write first in their mother tongue, all of them carry out the translation themselves. By contrast, in the case of Spanish doctors who write first in Spanish, only half of them perform the translation themselves. The rest have the paper translated by a professional translator. So, once again results give evidence that Spanish doctors need more support than German doctors and resort more easily to translation assistance.

In Table 3.25, data introduced in variable called “Translation Misunderstandings” indicate that misunderstandings on the part of the translator are not so common, although they occasionally occur as reported mainly by Spanish doctors.

	frequent		rather frequent		occasional		rare		never	
	n	%	n	%	n	%	n	%	n	%
SP	3	10.7	6.0	21.4	5.0	17.9	1.0	3.6	13.0	46.4
GER	0	0.0	0.0	0.0	2.0	33.3	0.0	0.0	4.0	66.7

Table 3.25: Doctors' opinion regarding frequency of misunderstandings after the professional's translation into English

In the Table above it is observed that 32.1% of Spanish doctors affirm that they have experienced it frequently or quite frequently, 21.4% say it has happened only occasionally or rarely and 46.4% affirms this was never the case. As for German doctors, the small minority assisted by a translator say that misunderstandings happen only occasionally (50%), in fact they normally do not take place according to them (50%).

	YES		NO		chi- square.
	n	%	n	%	
SP	22	88	3	12	n.s
GER	4	100	0	0	

Table 3.26: Spanish and German doctor's awareness of the influence of their original written Spanish/German text on the targeted translation

Table 3.26, abbreviated with “Translation consequences”, shows that most doctors seem aware of the fact that the original text in the mother-tongue affects the following translation. This fact was 100% clear on the part of German doctors but only 88% of the Spanish doctors stated their agreement. As a consequence, some of them do not even think

of this possibility and they probably write without translation in mind, which makes the translator’s task more difficult.

	YES		NO		chi- square.
	<i>n</i>	%	<i>n</i>	%	
SP	20	80	5	20	n.s
GER	4	100	0	0	

Table 3.27: Spanish and German doctor’s awareness regarding economical aspects and time saving when they write with simple structures in the mother-tongue.

Similarly, and as a consequence, in variable 10, “Saving time and money” (Table 3.27), it is observed that all German doctors affirm that they are aware that writing in a simpler way with simple structures would imply spending less money on a translator since it would imply less working time for him or her. Once more, Spanish doctors share this point except for a 20% who do not. This again underlines the fact that some Spanish doctors do not write with translation in mind and do not take into consideration the translation task.

	YES		NO		chi- square.
	<i>n</i>	%	<i>n</i>	%	
SP	23	92	2	8	n.s
GER	4	100	0	0	

Table 3.28: Spanish and German doctors’ interest in learning how to write with “translation in mind” facilitating the translation task

Finally, variable 11, “Facilitating translation”, shows (in Table 3.28) that all German doctors and almost all Spanish doctors (92%) support the idea that they would consider it useful if professionals of their field were taught how to use a simple way of writing that would facilitate translation tasks and save them money as a result.

	YES		NO		chi- square
	<i>n</i>	%	<i>n</i>	%	
SP	8	28.5	20	71.5	n.s
GER	12	30	28	70	

Table 3.29: Spanish and German doctors' percentage of rejected articles by editorial boards on linguistic grounds

An additional way to determine what the situation is like regarding the command of English on the part of doctors according to their needs is elicited with variable number 12. With this variable “Rejected Articles”, it was expected to determine the reality regarding the success of the papers according to the editorial boards. In table 3.29 it is observed that a higher percentage than expected, i.e. 30% of German doctors have been rejected a paper by the referees because of poor English. Thus, it is also noticeable, that some of them feel more confident in English than they actually are. As for Spanish doctors, 28.5% have been rejected a paper and asked for revision by a native speaker, thus the end product is very similar in both cultures and puts German doctors in a position that claims they should sometimes be assisted more frequently than they actually are. Considering that 50% of Spanish doctors carry out the translation themselves it is not surprising that nearly a 30% are rejected their papers. However as German doctors mainly translate the papers themselves their responsibility for the rejection is evident.

At last, it was interesting to find out by question number 13 in the survey that there are no linguistic centers to support doctors' publications in their place of work, although some German doctors mentioned other centers in charge of this

outside their institution. However, they did not seem to trust this center.

3.3 Lexical Analysis

As far as the size of the Conclusion Sections is concerned, the number of Conclusion Sections was 408 in the English corpus and 311 in the Spanish corpus. In general terms it should be noted that the size of the Conclusion Sections was very inconsistent in both languages, some were very long and others very short. By means of WordSmith™ it was possible to obtain detailed information regarding the size of the two corpora. For instance, despite the number of conclusions being higher in English, the number of words in the corpora was higher in Spanish (104,194 words) than in English (58,865 words).

3.3.1 Results

With regard to the difference in the number of sentences in each language, it was not so noticeable, i.e. in the English corpus 3,279 sentences were recorded and in Spanish 3,905. The above information accounts for the fact that the longest conclusions were found in Spanish R.As., in which not only the number of words was higher but also the length of the sentences, despite the number of conclusions being lower. Furthermore, the Spanish language uses a wider range of vocabulary. This is

evidenced in the number of different words in the corpora. Types were 10, 817 in Spanish versus 6, 653 in English.

An example of the average length of the Spanish sentences is shown in the following sentence:

En esto nos ha de llamar la atención no sólo sobre los beneficios físicos y psicológicos que resultan de una ejecución física adecuada a las necesidades individualizadas; sino de la posibilidad de que el mismo ejercicio se convierta en reducto y camino para la búsqueda de soluciones a otros problemas de salud (v.g. trastornos de la alimentación, trastornos de la imagen corporal, etc.) o se convierta en sí mismo en un problema de graves implicaciones clínicas, donde hablaríamos entonces de la dependencia al ejercicio.

The information in this chapter is focused on the word level with the objective of comparing the most recurrent nouns, adjectives, verbs, adverbs and conjunctions in the medical corpora of the two languages and see the closest terms surrounding them, i.e. to analyze collocations.

Prior to this analysis it is relevant to clarify the term “collocation”. From the translation-oriented corpus linguistics approach, Baker (1992:47) defines *collocation* as “the tendency of certain words to co-occur regularly in a given language”. She explains that sometimes this tendency is due to the meaning of the words (e.g. “cheque” has a high probability of co-occurring with “bank” or “money”). However, meaning cannot always account for collocations. If this was the case, then the noun

“visit” would be likely to collocate with verbs such as “pay”, “make”, “carry out”, “undertake” or “perform”. Nevertheless, the noun “visit” collocates with the first and the second verbs aforementioned, but not with the rest of them. Thus, collocations occur randomly and do not follow any rules. They are formed according to typicality. For instance, the nouns “rules” and “regulations” are synonyms, but the verb “to break” only collocates with the former.

Sometimes collocations are restricted to some morphological forms of a word family (Baker, 1992:48), i.e., the verb “to achieve” collocates with the noun “aim”, but other terms belonging to the same word family as “to achieve” also collocate such as for instance the adjective “achievable” or the noun “achievement” (e.g. “achieving aims”, “achievable aims” and “the achievement of an aim”, etc.). However, the noun “rules” collocates with the verb “to bend” but not with the adjective “unbendable”. It is more common to describe rules as “inflexible”.

The mismatch that takes place within a language regarding synonyms that do not collocate with the same words can also be observed across languages (Baker, 1992:49). The reason for these mismatches is due to the fact that languages are a reflection of the culture in which they are spoken. Therefore, the value that English place on the concept of order is reflected in the collocation “law and order”, which would correspond to “law and tradition/convention” in Arabic (“al-qunuun wa al-taqaalid”), since Arabs place more value on tradition.

It is also noteworthy that the more specific a word is, the narrower its collocational range (Baker, 1992:50). For instance, the verb “to bury”, which is more general than “to inter” collocates with “people”, “treasure”, “feelings”, etc., whereas only “people” can be interred. The likeliness of broader collocational frames based on specificity provoked a high expectancy that restricted collocational patterns occurred in medicine.

In order to study the range of patterns with which the terms in the medical corpora were compatible, the first step was the creation of 10 lists: 5 lists in English and 5 in Spanish. These lists contained 5 different morphological categories (i.e. nouns, adjectives, verbs, adverbs and conjunctions) in a descending order according to their rate of appearance in the corpora (see appendix VI). Next, these 5 lists in each language were reduced and only the 10 most recurrent words corresponding to the 5 morphological categories were included in the tables displayed in this chapter. In these tables, words present in both languages are marked in bold. These words in bold were then further analyzed with WordSmith™ to check the immediate words surrounding them to establish the most significant collocations (see appendix III).

The analysis carried out was cumbersome because sometimes a term can have different grammatical functions. For this reason on many occasions it was necessary to analyze each occurrence of a certain term in order to determine its function. For instance, “increase” or “need” can function as a noun or as a verb. Another example is the term “increased”, which can

function either as a past participle (e.g. “*The results seem to support the view that the prevalence of bipolar depression does not appear to be as strongly affected by social and cultural changes, although the prevalence of unipolar depression has increased since the introduction of industrialization.*”), as a simple past (e.g. “*In addition, we demonstrated that the rates of in-hospital mortality rose rapidly as the number of risk factors increased.*”) or as an adjective (e.g. “*Taken together, low dietary folate may be associated with an increased risk of severe depression, at least in middle-aged men living in eastern Finland.*”) etc. Therefore, the rate of appearance of an initially apparent recurrent term was often dramatically reduced after detailed analysis. For this reason, finding the 10 most recurrent grammatical categories was not a simple task.

Sometimes the meaning of a word changes depending on the following word. For example, the verb “*tener*” is equivalent to the full verb “to have” in English. However, it changes its original meaning when it is followed by “*en cuenta*” and becomes “to take into account”.

An additional problem was to find exact equivalences in English and Spanish because sometimes terms do not have a unique equivalence in the other language. For instance, the English verb “to be” corresponds to “*ser*” and “*estar*” in Spanish. Similarly, the verb “*poder*” corresponds to “*may*” but also to “*can*”.

3.3.1.1. Nouns

Table 3.30 represents the most recurrent nouns that come up in both languages with their corresponding rate of appearance. This information was obtained with the help of WordSmith™. Nouns were searched in plural and singular and then summed up to calculate their total appearance. Among these 10 nouns, 5 of them are indicated in bold because they coincide in both languages: “paciente/s”, “estudio/s”, “tratamiento/s”, “resultado/s”, “grupo/s” in Spanish and “patient/s”, “study/ies”, “treatment/s”, “result/s” and “group/s” in English.

paciente/s	461	patient/s	663
estudio/s	427	Study/ies	349
tratamiento/s	372	treatment/s	185
resultado/s	307	child/children	177
relación/es	236	result/s	147
grupo/s	228	group/s	137
caso/s	200	symptom/s	126
trastorno/s	178	therapy/ies	123
variable/s	171	risk/s	120
años	170	effect/s	109

Table 3.30: Most recurrent nouns in the Spanish and English corpora

Table 3.30 shows that the order in which the first 4 nouns are displayed according to their frequency in the corpora coincides in both languages (disregarding the noun “children” for being specific of the specialty of paediatrics). It is not surprising that the most recurrent nouns in both medical corpora

are “patient/s” and “paciente/s”. However, the rate of appearance of “patient/s” (1.13%) is higher than that of “paciente/s” (0.44%).

A frequent structure with the aforementioned nouns that differs in form in the two languages is “patients presenting with”, “patients who presented with” and “patients undergoing” that corresponds to “los pacientes que presentan/presentaban” or “pacientes con”.

The adjectives “selected” and “unselected” collocate with “patient/s” (e.g. “*However, certain antiarrhythmic agents may be beneficial for suppressing arrhythmia recurrences in selected patients.*”) Similarly, the noun “paciente/s” collocates with the adjective “seleccionado”, and the negation of the same adjective “no seleccionado” also occurs in the corpus (e.g. “*La TRC es un tratamiento eficaz en pacientes seleccionados con insuficiencia cardíaca avanzada*”). Other ways to refer to a group of patients is by means of the nouns “número” or “number” in “número de pacientes”/ “number of patients” and “grupo” or “group” forming “grupo de pacientes”/ “group of patients”. In Spanish the collocation “porcentaje de pacientes” was registered as well.

The collocation “population of patients” is more common in English than the Spanish “población de pacientes”. In fact, an alternative to the above combination coming up in Spanish is the collocation “población de sujetos”. However, the most common collocation with “población” is “de alto riesgo”.

The noun “paciente/s” collocates with “riesgo” as well, i.e. “pacientes de alto riesgo”. In English, “patients” collocates

with “risk” in the following forms: “patients at risk for”, “high risk patients” and “patients who have risk factor for”.

Another typical collocation is that formed with the past participles “treated” and “tratados”, which come up repeatedly with “patient/s” and “paciente/s”.

All the above collocations with the most recurrent noun in both corpora, i.e. “patient” and “paciente” can be summarized as follows:

English	Spanish
a) - <i>patients presenting with</i>	- <i>los pacientes que presentan/presentaban</i>
- <i>patients who presented with</i>	- <i>pacientes con</i>
- <i>patients undergoing</i>	
b) <i>selected/unselected patients</i>	- <i>paciente(s) seleccionado or no seleccionado</i>
c) - <i>número de pacientes</i>	- <i>number of patients</i>
d) - <i>group of patients</i>	- <i>grupo de pacientes and porcentaje de pacientes</i>
e) - <i>population of patients</i>	- <i>población de pacientes</i> - <i>población de sujetos</i> - <i>población de alto riesgo</i>
f) - <i>patients at risk for</i>	- <i>pacientes de alto riesgo</i>
- <i>high risk patients</i>	
- <i>patients who have risk factor for</i>	
g) <i>patients treated</i>	- <i>pacientes tratados</i>

Regarding the second noun in Table 3.30, the rate of appearance of “study/ies” (0.59%) is higher than that of “estudio/s” (0.41%). “Study” collocates with the adjectives “present”, “further”, “future”, “previous”, “current” and “prospective” in the stated order in terms of frequency. In Spanish the main adjectives that collocate with “estudio/s” in a descending order according to frequency are “presente”, “previo” “futuro” and “anterior”. It is also worth mentioning the fact that the noun “trabajo/s”, which is a synonym of “estudio/s”, has a considerable lower rate of appearance in the corpus (0.16%) and it also collocates with the adjective “presente”.

The main noun collocating with “study/ies” is “results” in “the results of this study” and the next in terms of frequency is “findings” mainly present in the structure “the findings of this study”. In Spanish the main noun collocating with “estudio/s” is “resultados” in the structure “los resultados de este/nuestro estudio”. However, although the structure “los hallazgos de este estudio” is registered in the Spanish corpus, it only occurs once.

As far as verbs are concerned, the main verbs collocating with “study/ies” are the verb “to need” in a passive structure “studies are needed” and the verb “to demonstrate” in the simple present tense “studies demonstrate”. The verbs “to suggest”, “to show” and “to indicate” collocate with “study/ies” mainly in the present simple tense, although the latter verb is less frequent. On the other hand, the main verbs that collocate with “estudio/s” are the participles “realizado” and “controlado”. The above information on the most common collocations with “study” and

“estudio” considering the different morphological categories could be summarized as follows:

a) Adjectives

English: *present, further future, previous, current and prospective (study)*

Spanish: *presente, previo, futuro and anterior (estudio)*

b) Nouns

English: *the results, the findings (of this study)*

Spanish: *Los resultados, los hallazgos (de este estudio)*

c) Verbs

English: *Studies are needed, Studies demonstrate, suggest, show*

Spanish: *estudios realizados, controlados*

Regarding the third noun in Table 3.30, “treatment/s”, the most outstanding adjectives next to it are “effective”, “surgical” and “combined” (e.g. “*Combined treatment is not as good as the individual therapies alone*”). In Spanish the rate of appearance of the noun “tratamiento/s” is slightly higher, 0.36% versus 0.31% in English. This might account for the wider range of adjectives collocating with “tratamiento”. The most typical ones are “tradicional” and “combinado” (e.g. “*Los pacientes del tratamiento combinado consiguieron remitir varios síntomas negativos de la enfermedad*”). The former adjective

corresponding to “traditional” in English does not collocate with “treatment”, however the latter adjective, which corresponds to “combined” in English, functions as a collocation of “treatment” as well. Other adjectives that also collocate with “tratamiento” are “específico”, “eficaz”, and “quirúrgico”. The corresponding adjective “surgical” collocates with “treatment” in English as well, “surgical treatment”. Finally, it is not surprising that the past participle “recibido” collocates with the noun “tratamiento” in the Spanish corpus forming “tratamiento recibido”.

As for verbs collocating with the noun “tratamiento”, the verb “finalizar” must be pointed out. A similar collocation used to refer to the end of a treatment is the noun “retirada” plus the preposition “de” forming: “la retirada del tratamiento”. With respect to nouns, another noun that collocates with “tratamiento” is “eficacia” followed by the preposition “de” resulting in “eficacia del tratamiento”. From the same word family the adjective “eficaz” collocates with “tratamiento” as well. Finally, the noun “reassignación” collocates with “tratamiento” forming “tratamiento de reassignación”. This last term is especially used in Psychiatry preceding the adjective “sexual” (e.g. “*Es importante tener en cuenta que el motivo principal que refieren estos pacientes para realizar el tratamiento de reassignación sexual es el ser miembro del género contrario*”). Regarding the Spanish collocations “retirada del tratamiento” and “finalizar el tratamiento”, they can be referred to as “zero correspondence”. This is the term used by Stig (2007: 25) when there is a lack of a clear counterpart in the other language. For Stig (2007: 27),

“zero correspondence” takes place “when there is no natural match across languages and particularly in the case of forms expressing textual (rather than ideational) meaning”. The main collocations of the nouns “treatment” and “tratamiento” can be summarized as follows:

a) Adjectives

English: *effective, surgical, combined*

Spanish: *tradicional, combinado, específico, eficaz, quirúrgico, recibido*

b) Nouns

Spanish: *eficacia, retirada, resignación*

c) Verbs

Spanish: *finalizar*

The fourth element in Table 3.30 is the noun “result/s” with a rate of appearance of 0.25%. Although it is preceded by the noun “children” with 0.30% of appearance, it can be considered, that “result/s” occupies the fourth place in the list because the term “children” corresponds specifically to the specialty of Paediatrics, and only general medical English is taken into account for the analysis of this research work. Thus, “result/s” coincides in the fourth place with “resultado/s”, which is represented with an appearance of 0.29%.

Regarding verbs, the noun “results” collocates with “to suggest” and “to indicate” in the present tense. “Resultados” collocates as well with the present tense of the verb “indicar”,

“sugerir” “mostrar” and “confirmar”. The main difference between the two languages is the same as with the nouns “study/ies” and “estudio/s”; unlike the English noun “result/s”, “resultado/s” collocates with a past participle, which is “obtenidos” in this case, forming “resultados obtenidos”.

Regarding adjectives, the collocation “mejores resultados” (0.08%) is more often used than the English combination “better results”, although neither of them are very often used. The adjectives “diferentes” and “similares” collocate with “resultados”, but the collocation “similar results” was only registered once in the corpus.

Longer collocations implying more than two words were also in both corpora like the structures: “los resultados de este/nuestro estudio” and “the results of this/our study”. The most relevant collocations of the nouns “results” and “resultados” can be summarized as follows:

a) Verbs

English: *suggest, indicate*

Spanish: *indicar, sugerir, mostrar, confirmar*

b) Adjectives:

English: *better*

Spanish: *obtenidos, mejores, diferentes, similares*

The fifth noun in the list in terms of occurrence is “group/s” with 0.23% and “grupo/s” with 0.22%. The most recurrent combination is “control group” and “grupo control” in

both languages. However other adjectives like “experimental”, “ambos” and “both” also collocate with this term in both languages. Another adjective collocating with “group/s” is “clinical”. The collocation “clinical group” is more common than “grupo clínico”.

As with other nouns in the list, “grupo/s” is sometimes accompanied by a past participle forming in this case “grupo tratado”. This time the English language also presents a collocation by means of a participle, although a different one: “the exposed group”.

As already stated above, the noun “group/s” collocates with “patients” in both languages resulting in: “grupo de pacientes” and “group of patients”. The most common collocations of the nouns “group” and “grupo” are shown below:

a) Nouns

English: *control (group), (group of) patients*

Spanish: *(grupo) control, (grupo de) pacientes*

b) Adjectives

English: *experimental, both, clinical, exposed*

Spanish: *experimental, ambos, clínico, tratado*

3.3.1.2 Adjectives

Table 3.31 and Table 3.32 show the first 10 most recurrent adjectives in English and Spanish:

WORD	FREQUENCY	% OF R.W
clinical	149	0.23
high/er/est	127	0.2
long/er/est	98	0.15
significant	95	0.15
early/earlier/earliest	93	0.14
far/further (-23adv)	88	0.14
surgical	84	0.13
low/er	81	0.13
good/better/best	74	0.12
important	65	0.1

Table 3.31: Most recurrent adjectives in the English corpus

WORD	FREQUENCY	% OF R. W.
mayor	339	0.33
social	182	0.17
sexual	172	0.17
clínica	114	0.11
significativa	127	0.12
importante	116	0.11
alta	103	0.1
baja	93	0.09
menor	91	0.09
posible	92	0.09

Table 3.32: Most recurrent adjectives in the Spanish corpus

Considering that Spanish adjectives present different morphemes according to the genre and number of the noun they refer to, in order to obtain the information in Table 3.31 it was necessary to search all the possible forms of the adjectives

separately by means of WordSmith™ tool. The number represented is therefore the sum of all the possibilities offered by the adjectives.

Tables 3.31 and 3.32 show that the order of appearance of the most recurrent adjectives in the two corpora (those in bold selected for further analysis) does not coincide (e.g. “important” occupies the last place in the English list, whereas in Spanish the equivalent adjective “importante” is in the fourth place). Additionally, the number of equivalent adjectives indicated with bold letters is not the same in both languages. In English 5 adjectives were selected for further analysis: “clinical”, “high”, “significant”, “low” and “important”. In Spanish, however, 7 adjectives were selected: “mayor” “clínica”, “significativa”, “importante”, “alta”, “baja” and “menor”. The reason for this mismatch is that sometimes more than one Spanish adjective is equivalent to the same English adjective. For instance, the adjective “high” in English can be equivalent to “alto” but its comparative form “higher” corresponds to a different Spanish adjective, i.e. “mayor”.

In terms of recurrence, the adjective “mayor” is on top of the list whereas “alto/a” is in the seventh place. The adjective “high” occupies the second place in the English list.

The most frequent nouns collocating with the adjective “high” are “risk” and “level/s”, the latter mainly when the adjective is in a comparative or superlative form: “higher/highest levels”. The comparative form of the adjective collocates with the noun “incidence” as well, forming “higher incidence”. This comparative form is also frequently found with

the adverb “significantly” as well; “significantly higher”. The comparative form “higher” appears almost as many times as the form “high”. It has a high rate of appearance (0.10%). The equivalent Spanish structure formed by the adverb “más” preceding the adjective “alto/a” is registered as well (e.g. “*Se plantea la necesidad de evaluar... especialmente en la población de más alto riesgo...*”) but its frequency in the corpus is only 0.01%. Nevertheless, the adjective “mayor”, equivalent to the comparative “higher” has a rate of appearance of 0.28%.

The comparative adjective “mayor” has a very wide range of collocations thanks to its high rate of appearance in the corpus (0.28%). The main nouns that collocate with “mayor” in order of frequency are “nivel”, “número”, “porcentaje”, “parte”, “grado”, “medida”, “riesgo”, “frecuencia” and “gravedad”. Regarding other morphological categories, the adverb “significativamente” collocates with it forming “significativamente mayor” (equivalent to “significantly higher” above).

Thus, some combinations such as “higher risk” or “higher level” have two possible corresponding adjectives in Spanish depending on the sentence. For instance, “higher level” can be translated sometimes into “mayor nivel” or into “nivel más alto” For instance, in the following sentence “higher” can be translated into Spanish as “mayor”: “*However, we believe that our results point to the possibility of a much higher proportion of depression among mothers both in...*”, “una proporción de depresión mucho mayor”. Another example in which “higher” must be translated by “mayor” is: “*One can*

speculate that inhibition deficit in TD is more evident when measured by tasks with a higher demand on prefrontal networks". By contrast, in the following example "higher" is equivalent to "más alto/elevado": "*Patients with dissociate symptoms were at increased risk of subsequent PTSD and also exhibited higher levels of arousal*", i.e. "niveles más altos/elevados de aparición". This last example shows that on many occasions both options are possible in Spanish, since "mayores niveles de aparición" is a possible translation for the last exemplifying sentence above as well.

It is therefore vital for researchers to know and respect patterns of collocation that might be different and more or less restrictive in English than in their own language. Regarding the comparative adjectives "mayor" and the comparative form "alto" preceded by the adverb "más", Spanish researchers have little difficulty since they can translate both with the comparative "higher".

Other nouns that collocate with "high" are "rate" and "percentage". The adjective "high" is very often followed by a noun plus the preposition "of" forming what Luzón (2000) calls a "triplet", e.g. "high rate of", "high prevalence of" or "high risk of" (e.g. "*...the AAS users reported a high rate of adverse physical and psychological effects of AAS*"). These triplets coincide with Luzón's study (2000:69), who underlines the high frequency of "the risk of" in his analysis of medical papers. However, this study is focused only on the Conclusion Section of medical papers and additionally in this study the adjective

“high” (“high risk of”) replaces the article “the” in Luzón’s triplet: “the risk of”.

The main nouns collocating with the adjective “alto/a” in the corpus are: “riesgo”, “porcentaje”, “prevalencia” and “nivel/es”. The latter noun is used mainly in the plural form and either preceding or following the adjective (i.e. “niveles altos” or “altos niveles”). Other nouns, less common in terms of frequency but that can also be considered collocations are “grado”, “número” and “índice”. Thus, it can be stated that most collocations with the adjectives “high” and “alto” coincide in both languages, the most frequent one being “high risk” and “alto riesgo”. The other nouns forming the collocation coincide (i.e. “high level” and “alto nivel” or “high rate” and “alto índice”) although they do not follow the same order in terms of frequency, i.e. “high level” appears as the second most commonly used collocation in the English corpus but “alto nivel” appears as the fourth most spread collocation preceded by others like “alto porcentaje”. However, it can be concluded that most collocations with the adjective “high” coincide in both languages despite frequency differences (in fact, examples of “high grade” or “alta incidencia” are less frequent but also registered in the corpora) and should therefore not constitute a problem for Spanish researchers when writing English papers.

Adjective	Nouns	Adverb
high (er/est)	risk level incidence rate percentage prevalence grade	significantly
mayor	nivel numero porcentaje parte grado medida frecuencia gravedad	significativamente
Alto	riesgo porcentaje prevalencia nivel grado número índice incidencia	

Table 3.33: Collocation with the adjectives “high”, “mayor” and “alto”.

Table 3.33 shows the most frequent collocations in terms of frequency with the adjective “high” and the corresponding Spanish adjectives “mayor” and “alto”. As it is observed, there are some cases of “zero correspondence”. For instance, the noun “parte” has no correspondence in English; since the noun “parte” in the combination “mayor parte” would only correspond to the adjective “most” in English (e.g. “*La mayor*

parte de la información la reciben de la matrona” would be translated into “*most information*”).

Also the Spanish combination “*mayor medida*” has a “zero correspondence” because the adjective “greater” is used in English instead (e.g. “...*interventions would be more suitable for schizophrenic patients, whereas the treatment of cognitive impairment in bipolar patients would depend to a greater extent on their clinical status ...*). However, the technical form “*a high OP measurement on lumbar puncture*” takes place in the English corpus. Other examples of “zero correspondence” are “*mayor gravedad*”. To express the idea of “*gravedad*” the adjective “serious” is mainly used in English (e.g. “*The risk of serious cardiovascular events in AMI increased with plasma NT-proBNP...*”).

The position of the adjectives “low”, “bajo” and “menor” is similar in both languages, since “low” is the adjective number 8 in the list and “bajo” and “menor” occupy the places 8 and 9. Despite being some of the last adjectives in the list, it is sensible to describe them after “high”, “alto” and “mayor” due to their similitude. The adjective “low” is equivalent to “bajo”. However, in its comparative form, “lower” it is equivalent to “menor” but also to “más bajo” depending on the sentence.

Considering that the adjective “low” is an antonym of “high”, it is not surprising that many collocations for this adjective match with those registered for “high”. The adjective “low” collocates mainly with the noun “grade”, although it also

collocates with “level”, “risk”, “number”, “rate” and “incidence”.

As for the adjective “bajo”, it collocates mainly with the noun “nivel”, but also with other nouns in the corpus such as: “peso”, “rendimiento” and “grado” (e.g. “...*parece que PTEN no participa en la génesis de astrocitomas pediátricos de bajo grado mediante mutaciones puntuales...*”).

The most common use of this adjective is in the form of a triplet consisting of the adjective “bajo” plus a noun plus the preposition “de” as in “bajo índice de” and “bajo nivel de”.

Regarding comparisons, the adverb “más” is used with the adjective “bajo/a” (e.g. “*Nuestros datos indican que el valor óptimo de corte es más eficaz para la confirmación del diagnóstico de ICC (valor predictivo positivo muy elevado) que para su exclusión (valor predictivo negativo más bajo), al contrario que lo señalado en estudios previos.*”). As far as the comparative adjective “menor” is concerned, it collocates with “número”, and “grado”, (e.g. “...*porque no necesita de controles hemáticos periódicos y por su menor grado de interacciones con otros fármacos*”), although it also appears with other nouns such as “nivel” and “medida”.

According to the above collocations, the nouns “grado” and “nivel” collocate with both adjectives “bajo” and “menor”. In English, however, the adjective “low” and its comparative form “lower” would be used in both cases. Therefore, “lower grade” would correspond to “grado más bajo” or “menor grado” depending on the context.

In the following exemplifying sentence the adjective “lower” would correspond to “menor” or “más baja”: *“more refined surgical techniques and more effective diagnostic examinations have resulted in a lower incidence and more appropriate treatment of complications”*.

As it was the case with “high”, the English use of this adjective is wide and the Spanish researcher will find no difficulty in learning collocation patterns with this adjective in medicine.

The adjective “clinical” with 0.23% representation in English is more widely used than “clínico/a” in Spanish, which has a rate of appearance of 0.11% in the corpus. The instances where “clínica/o” was a noun were taken into account and disregarded when calculating its appearance as an adjective. The main nouns collocating with “clinical” are “practice” and “course”, but also “trials”, “symptoms” and “studies” collocate with “clinical”. The main collocation in Spanish coincides with English, i.e. “práctica clínica”/“clinical practice”. Other nouns that also collocate are “características”, “grupos”, “utilidad”, “datos”, “diagnósticos” and “casos”. In the case of “clinical”, except for the collocation “clinical practice”/“práctica clínica”, medical researchers have to be aware of the different use of the adjective in both languages.

The adjectives “significant” and “important” and the corresponding Spanish terms “significativa/o” and “imporante” are synonyms. However, it is noteworthy that despite being synonyms, the rate of frequency differs significantly in English. In fact, the adjectives are distant regarding their position in the

list. “Significant” is in the fourth place in the list whereas “important” is in place number 10, which means that the former is more widely used in a medical context. However, in Spanish, “significativa” is in the fifth place and “importante” the sixth adjective in the list, thus the frequency of use is similar in Spanish.

The main nouns that collocate with the adjective “significant” are “difference”, “increase”, “impact”, “relationship”, “effect” and “morbidity” in the stated order according to frequency.

As far as adverbs are concerned, “statistically” collocates with “significant” and the adverb “no” frequently precedes this adjective. The most common structure registered is the adjective “significant” followed by a noun and the preposition “in” (e.g. “*significant decrease in morbidity*”, “*significant difference in*”).

The adjective “significativa” appears mainly in feminine and it collocates with the noun “diferencia” mainly in the plural and followed by the preposition “en” (“*diferencias significativas en*”), however it collocates also with “manera”, and “forma”. The masculine form of the adjective collocates with the noun “progresos” (“*progresos significativos*”).

Regarding adverbs, “significativa” collocates with “estadísticamente” (e.g. “*No obstante, dicha relación tan sólo se muestra estadísticamente significativa para la muestra global.*”)

The most common collocation in both languages is “diferencias significativas” and “significant differences”. A second collocation also present in both medical corpora is

“estadísticamente significativas”/ “statistically significant”. However, many collocations in the corpora with the adjectives “significant” and “significativas” are different and thus Spanish doctors need to take this into account when writing an English paper.

As for the adjective “important”, it collocates with “role” and with the adverb “most” in the superlative form, i.e. “the most important”.

In Spanish the noun “papel” also collocates with the adjective “importante”. The adverb “más” as in English is used to establish comparisons; “más importante”. The use of “important” is very similar in both languages.

3.3.1.3. *Adverbs*

Table 3.34 and Table 3.35 show the 10 most frequent adverbs in each language. The adverbs that coincide in both languages are 7 in English: “not”, “more”, “most”, “also”, “only”, “thus” and “very” and 6 in Spanish: “no”, “más”, “también”, “así”, “muy”, “solo”/ “sólo”.

WORD	FREQUENCY	% OF R.W.
not	271	0.42
more	168	0.26
also	109	0.17
only	82	0.14
most	54	0.8
even	62	0.1
therefore	47	0.07
thus	45	0.07
significantly	32	0.05
very	31	0.05

Table 3.34: Most recurrent adverbs in the English corpus

WORD	FREQUENCY	% OF R. W.
no	828	0.75
más	637	0.58
como	606	0.55
también	159	0.14
así	113	0.1
muy	106	0.1
además	105	0.1
solo/sólo	95	0.09
ya	60	0.05
donde	56	0.05

Table 3.35: Most recurrent adverbs in the Spanish corpus

The lack of correspondence between the number of selected English and Spanish adverbs is due to the fact that in Spanish the adverb “más” is used both for comparative and superlative structures whereas in English two different adverbs are used, i.e. “more” for the comparative and “most” for the superlative.

Instances in which “more” was not an adverb but rather an adjective (25 cases) or a pronoun (4 cases) are not represented in the percentage displayed in Table 3.34. In the same way, the instances of “most” as an adjective (7) are not represented either.

Regarding the order of the adverbs in the list, the first two Spanish adverbs, i.e. “no” and “más”, coincide with the first two English adverbs “not” and “more”.

The adverb “más” as well as “more” and “most” accounts for the high tendency to make comparisons in medical corpora. The adverb “más” has a higher percentage of representation than the adverbs “more” and “most”, i.e. 0.58%

versus 0.34%, respectively. However, while the Spanish language needs the adverb “más” to make a comparison, the English language does not always need the adverbs “more” and “most”. In fact, one-syllable adjectives and some two-syllable adjectives form the comparative and superlative with the addition of the suffixes “-er” for the comparative and “-est” for the superlative. For this reason, it can only be concluded that comparisons play an important role in both languages.

The adverb “also” has a similar position in terms of recurrence in both languages, representing the third and fourth place of the list in English and Spanish, respectively. The third place in Spanish is the adverb “como” (e.g. “*Su pérdida debe ser catalogada como uno de los efectos adversos que más van a influir en su evolución.*”)

The adverb “only” is more recurrent in the English corpus, where it occurs in position 4 with a representation of 0.14%. The adverb “sólo” is in place number 8 with a representation of 0.09%.

“Thus” can be sometimes considered equivalent in meaning to “así”. However, there is no unique correspondence between these two adverbs; in fact, other words or combination of words could be possible alternative translations to these adverbs.

“Thus” is translated in the Collins Spanish dictionary (1994) as “así”, “así es que”, “de este modo”. The Collins Cobuild English Dictionary (2000) gives two possible meanings for the adverb “thus”: a) you use “thus” to show that what you are about to mention is the result or consequence of something

else that you have just mentioned. As an equivalent for this first meaning of “thus” the Cobuild English Dictionary suggests “hence” and “therefore”. b) “If you say that something is “thus” or happens “thus”, you mean that it is, or happens as you have just described or as you are just about to describe”; a formal use. The corresponding meaning to this second definition would be the one proposed in the Collins Spanish Dictionary, i.e. “así” or “de este modo”. However, the first definition given by the Cobuild English Dictionary, in which “thus” is equivalent to “therefore” or “hence”, does not correspond to the Spanish adverb “así” but rather to other expressions such as “por lo tanto”, “por consiguiente”, “consecuentemente”, “en consecuencia”, “como consecuencia”, “por eso”, “por este motivo”, “por esta razón” (e.g. “*Thus, it seems that TT is feasible among patients who require concomitant warfarin treatment or high-risk ACS patients*”). Another example in which “por lo tanto” could be the translation for “thus”: “*We suggest that when encountering the alexithymic patient, verbally expressed acceptance and empathy from the physician has a significant impact on patient satisfaction with the consultation, and thus also on the creation of the therapeutic alliance*”.

Adverbs expressing a consequence are widely used in English medical papers; in fact, the adverb “therefore”, which is even more commonly used than “thus” has a percentage of 0.07% in the corpus.

It is important to underline in most instances of the corpus the adverb “thus” expresses a consequence and is equivalent in meaning to “therefore”.

The Spanish expressions conveying consequence (e.g. “por lo tanto”, “por consiguiente”, etc) represent 0.04% of the corpus. The adverb “así” with a representation of 0.1% in the corpus is not only used to express consequence, but is also used to express other things (see paragraph below). Thus, it can be concluded that taking into account that there are other possible words in English expressing consequence (e.g. “consequently” or “as a consequence”), there is a high probability that it is more common in English papers to use words expressing consequence than in Spanish because the adverbs “thus” and “therefore” represent 0.14% of the corpus. However, a more detailed analysis would be needed to make this statement.

The Spanish adverb “así” also has different equivalences in English as well. According to the Collins Diccionario Inglés (1998) the adverb has two possible equivalences in the English language: a) “so”, “in this way”, “thus”, “by this means”, “thereby”, b) in comparisons “así” corresponds to “both”, “as well as” and “so” (e.g. “así de pobre que”- “so poor that”). The meaning of “así” in “a” is equivalent to the English “thus”, however, the second meaning in “b” of “así” is not.

In the Spanish corpus there are 48 instances of “así como” representing 0.05 of the corpus. “Así como” corresponds to “as well as” (e.g. “*El malestar psicológico global, así como el grado de ajuste social expresados por los adolescentes de nuestra muestra en YSR, es equivalente para varones y mujeres*” or “*Varios síntomas relacionados con la competencia social se mantuvieron estables, así como algunos síntomas generales de la enfermedad.*”). However, other cases of “así” also occur. For

example “así” appears in combination with “pues”, which corresponds to the English “thus”, as shown in the following sentence: *“Esta circunstancia puede explicarse, en parte, por el hecho de ser una muestra hospitalaria en la que la gravedad clínica condiciona muy estrechamente la necesidad de ingreso. Así pues, aunque los trastornos adaptativos son muy frecuentes en los pacientes de edad avanzada, no requerirían habitualmente tratamiento en régimen de ingreso hospitalario”*.

These examples of the text account for the fact that it can only be stated that “thus” and “así” have a high rate of appearance in both corpora despite not being always equivalent in meaning. In the case of “thus”, it can be concluded, though, that its most common use is that of expressing consequence.

The adverb “muy”, which comes immediately after “así”, is in the 6th place of the list. It is more frequently used than “very”, which appears as the last element among the 10 more commonly used adverbs.

3.3.1.4 Conjunctions

Table 3.36 and 3.37 show the most frequent conjunctions. 7 English conjunctions are marked in bold in Table 3.36 (“and”, “or”, “however”, “but”, “when”, “if” and “although”) because they coincide with 8 of the Spanish conjunctions in Table 3.37: “y”/ “e”, “o”, “si”, “aunque”, “pero”, “cuando”, “sin embargo”. The Spanish conjunction “e” is equivalent to “y” and both of them correspond to “and” in English. (The use of “e” is due to euphonic reasons. According

to the DRAE (2001, 858) it is used in front of a word starting with the vowel “i” to avoid the hiatus).

WORD	FREQUENCY	% OF R.W.
and	1713	2.68
or	268	0.42
however	111	0.18
but	103	0.16
when	74	0.12
if	65	0.1
while	46	0.07
although	45	0.07

Table 3.36: Most recurrent conjunctions in the English corpus

WORD	FREQUENCY	% OF R.W.
y/e	2972	2.84
o	481	0.44
si	171	0.16
aunque	131	0.12
pero	124	0.11
cuando	118	0.11
sin embargo	95	0.09
ya que	72	0.08
a pesar	63	0.06

Table 3.37: Most recurrent conjunctions in the Spanish corpus

In terms of frequency, the two most widely used conjunctions in both languages are “and” and “or” in the English corpus and “y/e” and “o” in the Spanish. Although the difference in terms of frequency is not very significant, the

conjunction “y/e” has a higher rate of appearance, i.e. 2.84% vs. 2.68%). This contributes to account for the fact that Spanish sentences are longer than English since these prepositions are used to connect clauses.

There is no coincidence regarding the third place in the list. In English the conjunction “however” comes in this place, whereas the corresponding “sin embargo” is placed the seventh conjunction in the Spanish list. (Table 3.36 includes only the instances in which “however” functioned as a conjunction. Its function as an adverb was disregarded).

The high difference with respect to the percentage of appearance between “however” (0.18) and “sin embargo” (0.09) is reduced when taking into account the presence of the conjunction “no obstante” in the Spanish corpus which has a representation of 0.03%. The latter conjunction is a synonym of “sin embargo” and its rate of appearance together with that of “sin embargo” reaches 0.12%.

The presence of “however”, followed by “but” in the list and the appearance of “although” in Table 3.36 reflects that it is very common to establish contrast in medical conclusions, since this is the main function of the 3 conjunctions. This statement is applicable to the Spanish corpus as well, where the same 3 corresponding conjunctions appear (see Table 3.37).

The Spanish conjunction “aunque”, which is placed number 4 in the list with 0.12% representation in the corpus, indicates that its use is wider than “although”, which is the last of the 10 most recurrent conjunctions in English with 0.07% representation. Additionally, the conjunction “a pesar” can be

considered a synonym of “aunque” and it is placed as number 10 in the Spanish list. The sum of the two conjunctions results in a percentage of 0.18. Despite the two conjunctions being used differently in the sentence obeying to the rules of the language they are very similar in meaning and have a significant representation (e.g. “*En resumen, a pesar de las limitaciones de este trabajo (en especial el bajo número de pacientes con TS utilizado), creemos que los resultados no apoyan la utilización del instrumento...*” and “*En una población no seleccionada de pacientes con IAM con y sin elevación del ST, la prescripción de BB e IECA en el momento del alta hospitalaria muestra efectos aditivos sobre la supervivencia al año, lo que apoya su prescripción a todos los pacientes postinfarto sin contraindicación, aunque su efecto es más notorio en el subgrupo de alto riesgo.*”)

An English equivalent to the Spanish preposition phrase “a pesar de que” is the preposition “despite” in spite of the different grammatical category. Its representation in the corpus is of 0.06%, the same as “a pesar”. Thus the representation of “although” plus “despite” is 0.13%. As with “a pesar”, the use of “despite” differs from “although” in the way they are used in the sentence. “Despite” is normally followed by a noun or by a gerund (e.g. “*Despite acknowledged shortcomings, this study introduces the idea of utilizing a solution as an adjunct to the surgical technique*”), whereas “although” is followed by a sentence (e.g. “*Although some findings are tentative and methodological short-comings preclude drawing causal inferences, we will sum up results that are inconsistent with the*”)

main trend of findings from short-term studies). In any case, these data support the importance of establishing contrasts in medical Conclusions.

The conjunction “but” is placed the fourth in the English list, whereas “pero” occupies place number 5. Thus, “but” is more frequently used in English than in Spanish with a percentage of appearance of 0.16% vs. 0.11%, respectively.

The conjunction “when” with a representation of 0.12% is very similar in terms of recurrence to the Spanish “cuando” with 0.11%. (The instances of “cuando” with a different function [adverbial locution: 4 instances] were not included in table 3.37).

However, Tables 3.36 and 3.37 show that the conditional conjunction “si” is more widely used in Spanish medical conclusions than the English conjunction “if” (i.e. 0.16% versus 0.1%). Although the difference is not extreme, it is again a fact that accounts for the longer sentences of the Spanish corpus (e.g. *“Posiblemente, en el momento actual, los sujetos que pudieran ser catalogados como negativistas, están siendo diagnosticados como antisociales (si predominan los comportamientos agresivos, impulsivos, dominantes y disruptivos), como límites (si predominan o se suman problemas de la esfera emocional), o a ambos (alimentando una posible comorbilidad espuria) o bien son incluidos en la categoría, difícilmente accesible a la investigación, de trastorno de personalidad no especificado.”*).

Nevertheless, the appearance of other equivalent English terms to “if” such as “provided” and “whether” must be taken into account. The first element does not appear as a synonym of

“if” but only as a past participle of the verb “to provide” (e.g. “*Thus, the response provided to Stroop CW stimulation should be taken as a genuine marker of their psychological reactivity*”). On the other hand, the second word, “whether” is a synonym of the conjunction “if” and it appears frequently preceded by “to” plus a verb in the infinitive, i.e. “to investigate whether”, “to assess whether”, etc. (e.g. “*There are inadequate data to determine whether selected low-risk patients should be offered neoadjuvant chemotherapy followed by surgical resection of residual disease*”). The word “whether” appears 32 times and implies a 0.05%. This would consequently make the Spanish rate of appearance of the Spanish conjunction “si” (0.16%) almost equivalent to the English sum of the terms “if” and “whether” (0.15%).

3.3.1.5. Verbs

As expected, the most recurrent verbs in the corpora are those that can function both as auxiliary and as full verbs, i.e. “to be” and “to have” in English and “ser” and “haber” in Spanish. Table 3.38 shows the verbs in the corpora ordered in terms of frequency. Data in Table 3.38 include all the occurrences of the verbs, i.e. different tenses and mode and different functions (verbs with an auxiliary function and with a full verb function). Semantic equivalences are displayed by means of a letter next to the verbs in bold. Sometimes a verb in a language has two possible corresponding forms in the other language. To reflect this fact, numbers have been added to the

letters (e.g. the English verb “to be” is labelled with “a” and the corresponding Spanish verbs “ser” and “estar” are labelled with “a1” and “a2” respectively). On the other hand, there are some verbs that do not have an equivalent in the other language, in the sense that the corresponding verb is not among the ten verbs most frequently used, and they are therefore labelled with a different letter (e.g. “considerar”, which is labelled with “i” or “need”, which is labelled with “l”).

Spanish verbs in Table 3.38 have been labelled with an “S” in front of them, standing for Spanish. Similarly, an “E” standing for “English” precedes the English verbs. Semantic equivalences have been established starting by the most frequent verbs in each language. As the rate of frequency of the verbs varies from one language to another the order in which the verbs are commented on does not correspond to the order they show in Table 3.38. The order in which the analysis of the verbs has been carried out is based on frequency parameters and semantic correspondences. Those verbs which had to be described with more detail have been presented under an individual heading in the analysis.

SPANISH			ENGLISH				
	WORD	FREQUENCY	% OF R.W.		WORD	FREQUENCY	% OF R.W.
Sa ₁)	ser	1.680	1.61	Ea)	be (aux+full)	2225	3.78
Sa ₂)	estar	245	0.23				
Sa)	ser+estar	1.925	1.85	Ed ₁)	may	348	0.59
Sb)	haber	648	0.62	Ed ₂)	can	265	0.45
Sd)	poder	520	0.51	Eb)	have (aux)	235	0.39
Sc)	tener	271	0.26	Ef)	should	184	0.31
Se)	realizar	218	0.21	Ee)	do (aux+full)	162	0.27
Sf)	deber	205	0.2	Ec)	have (full)	143	0.24
Sg ₁)	presentar	156	0.14	Eg)	show	119	0.2
Sh)	permitir	114	0.11	Ej)	suggest	118	0.2
i)	considerar	111	0.11	Ek)	will (aux)	77	0.13
g ₂)	mostrar	110	0.11	El)	need	75	0.13

Table 3.38: Most recurrent verbs in the Spanish and English corpora

Ea) The verb “to be”

If the verb “to be” is broken down into its different tenses, there are 749 instances of the present simple form “is” (1.17%). The plural form “are” appears 371 times (0.58 %). The past form “was” is documented on 154 occasions (0.24%) and the corresponding plural form 136 times (0. 21%). The past participle “been” appears 97 times (0.15%). The form “be” has 686 occurrences representing 1.07 %. Table 3.39 below summarizes the appearance of the verb “to be” in the Present, Past, Participle and infinitive in the corpus.

VERB	FREQUENCY	%
is	749	1.17
are	371	0.58
was	154	0.24
were	136	0.21
been	97	0.15
be	686	1.07

Table 3.39: Occurrences of the verb "to be" in the English corpus .

The form “is” was used as a full verb on 516 occasions and in 232 as an auxiliary verb, out of which it was acting on 10 occasions as part of a present continuous structure. In the other 222 cases it was forming part of a passive structure, which represents 0.38% of the corpus.

The plural form “are” also acts mainly as a full verb with 243 instances. However, it is also used in passive structures 115 times, which represents 0.20% of the corpus and on 15

occasions as an auxiliary verb forming different tenses, the most frequent being the present continuous.

In the past tense there were 154 instances in the singular (“was”) (0.24%). In 79 cases the verb functioned as a full verb and in 75 cases it was part of a passive structure (0.13%). Regarding the corresponding plural form “were”, 136 instances were found (0.21%), out of which most cases were forming a passive structure, i.e. 73 (0.12%). On 7 occasions it was an auxiliary verb and on 56 it was a full verb.

The form “been” is an auxiliary verb form that appears mainly in passive structures, concretely in 81 cases out of 97, (0.14%). Regarding the 16 instances left, “been” forms part of a present perfect tense on 14 occasions, of a present continuous structure on one occasion and of a modal structure in the present perfect tense on one occasion.

The form “be” appears a total of 686 times in the corpus. It appears mainly in combination with modal verbs and especially in a passive structure. The most recurrent combination is that of the verb “should”, which appears a total of 127 times (0.22%). Out of this 127 times, on 110 occasions it is a passive structure, i.e. “should be+ participle” (0.19%) (e.g. “*Further research should be directed toward the evaluation of a longer follow-up period than the 180 days used here...*”). Table 3.40 below summarizes the frequency of occurrence of the verb “to be” both as an auxiliary and a full verb.

WORD	FREQUENCY full verb	%	FREQUENCY auxiliary verb	%
is	516	0.88	232	0.39
are	243	0.41	128	0.21
was	79	0.13	75	0.12
were	56	0.09	80	0.13

Table 3.40: Occurrences of the verb "to be" as an auxiliary and as a full verb

As regards the form “be” in combination with modal verbs, “should” is followed very closely in terms of frequency by “may”, with a rate of appearance of 124 times (0.21%). However, “may” is more often used in an active structure than in the passive. In fact, it is registered 51 times in a passive structure (e.g. *“Taken together, low dietary folate may be associated with an increased risk of severe depression...”*), meaning only 0.09% of the corpus, versus 73 in an active structure (0.13%) (e.g. *“There is indication that this may be particularly true when assessing the effects of long-term (longer than 3 months) psychosocial interventions...”*).

The next modal verb in combination with the verb “to be” is “can”, which is mainly used in the present tense and in the passive (e.g. *“The following conclusions can be drawn: the occurrence of any single basic symptom does not show schizophrenic specificity in adolescents...”*). It appears in a passive structure in 89 sentences (0.15%) whereas only 6 instances of the active were registered. The negative form “cannot” in the passive was also significant, since it appeared on 16 occasions (0.02%). Additionally, 21 examples of the past tense “could” were also registered in the passive (0.03%) and in 8 sentences it was present in the active.

Much less representative are the modals “need” and “might”. The former appears 24 times in the form “need to be + participle” (0.04%) and only once in an active sentence. As for “might”, it only appears 8 times in a passive structure (0.01%) and 14 times in an active structure. The modal “must”, has a low representation as well since it is recorded on 16 occasions in the passive (0.02%) and on 5 in the active.

The modal “have to” is also present, but its representation is very low; it only appears 4 times in the corpus. On 3 occasions it is used in the present tense and in the passive and only one example in the active with the verb in the past tense is registered.

The data above show that the modals least frequently used in the passive are “have to” and “might”, which, like “may”, appear more in the active than in the passive. However, “may”, unlike “might”, has an important representation in the corpus in the passive (see above). Table 3.41 sums up the co-occurrence and frequency of the verb “to be” with the different modal verbs in the corpus:

VERB	FREQUENCY	%
can be	140	0.23
should be	131	0.22
may be	129	0.21
might be	24	0.04
must be	21	0.03
have to be	4	0.01

Table 3.41: Frequency of appearance of modal verbs with the verb "to be"

Apart from the modals, the form “be” is often used in combination with “will” and “would”. The latter is recorded 18 times and only on 3 occasions it is forming part of a passive structure. The latter appears 15 times, out of which on 5 occasions it is used in the passive. Thus, it can be stated that “will” is more commonly used in the passive (e.g. “*As such, patients who are more susceptible to end organ damage because of a more virulent disease will be identified and more aggressively treated...*”).

Another common verb that appears in the passive in combination with “be” is the verb “to remain” in the present tense (e.g. “*It remains to be seen whether the differences in rates are a result of methodological issues or different characteristics of sample populations*”).

Other common structures containing the form “be” are those used by researchers in order to mitigate their statements. They make use of verbs such as “seem to be” (17 examples), “appear to be” (21 examples), “be likely to” (9 examples) or “to be shown to” (5 examples) (e.g. “*In particular, both marital distress, in terms of losses, divorces, absence of the husband or lack of a supportive relationship, and occupational reputation seem to be factors related to depression*”).

All the instances in which the verb “to be” was used in a passive structure represent over 1.93% of the corpus. Taking into account that the representation of the verb “to be” is 3.47% of the whole English corpus, it can be observed that more than

half of the instances in which it occurs it is used as an auxiliary verb in a passive structure.

Sa₁) The verb “ser” and a.2 the verb “estar”

The verb “ser” in the present simple form “es” is documented 758 times. This form of the verb is not used as an auxiliary unlike in English, but it is always a full verb. This form is often accompanied by adjectives forming “es necesario” (31 examples), “es posible”, with the same amount of instances and “es preciso” (12 times). The plural form “son” appears as a full verb on 333 occasions. It also appears accompanied by the adjectives such as “similares” and followed by the preposition “a” resulting in “son similares a”. Sometimes the verb “ser” is accompanied by two adverbs as in “son significativamente más”.

The verb “estar” has a representation of 0.24% and it corresponds to the English “to be”, as well. The present tense of this verb (“está”) appears 88 times in the singular and 65 times in the plural. These forms collocate with the participle “relacionado/s” mainly in the feminine and plural, i.e. “están/n relacionadas” (e.g. “*Las diferencias obtenidas en los diferentes grupos de edad están relacionadas directamente con el nivel evolutivo, la mayor relación de dependencia de los padres, así como...*”). It is worth mentioning that the terms “relacionar” and “asociar” are very widely used in the Spanish corpus. In fact, with WordSmith™ it was possible to calculate the word family of “asociar”, which appears 127 times (0.12%) and the word

family of “relacionar”, which appears 188 times (i.e. 0.18%). Similarly, in English the word family of “associate” appears 121 times (0.21%) and concretely the form “is associated with” appears 15 times. The word family of “relation” is recorded 136 times (0.23%). These data indicate that it is very common in medical corpora to establish associations by means of these terms in which the form “be” is present.

Regarding other tenses of the verb “estar”, it is observed that the perfect tense “ha estado” occurs only once, the so called *pretérito imperfecto* (“estaba/n”) occurs in 10 examples and the conditional (“estaría/n”) appears 8 times. (The form “estado” appears mainly as a noun and not as a participle). There is only one case registered of the future (“estará”), 10 cases of the present tense in the subjunctive (“esté/n”) and only two instances of the *imperfecto* in the subjunctive (“estuviera/n”). Thus, it can be concluded that the main use of the verb “estar” is the present simple. In Table 3.42 all the instances of the verb “estar” in the corpus are displayed:

MODE AND TENSE	VERB	FREQUENCY	%
Indicative			
Present tense	está (n)/estamos	167	0.16
Imperfect	estaba (n)	10	0.01
Conditional	estaría (n)	8	0.01
Present Perfect	ha (n) estado	1	0.001
Future	estará (n)	1	0.001
Subjunctive			
Present	esté (n)	10	0.01
Imperfect	estuviera (n)	2	0.002
Infinitive	estar	46	0.04
Total		245	0.24

Table 3.42: Different tenses of the verb “estar” in the corpus

Regarding the verb “ser”, there are 60 examples of the *present perfect* in plural and singular “ha sido/han sido”. On 53 occasions this verb tense with the verb “ser” forms part of an active structure (e.g. “*el análisis de las intervenciones efectuadas muestra que ambas han sido efectivas en las medidas llevadas a cabo en la fase de postratamiento y en los seguimientos...*”) and only on 17 occasions it is forming part of a passive structure (e.g. “*Estos resultados han sido replicados en el presente estudio*”). Despite the lower use of the passive in this tense, it is one of the tenses more often used in Spanish for passive structures.

In the past form, 51 instances of “fue” were registered, out of which it was acting as a full verb 36 times. There was only one case of “fue” functioning as the past of the verb “ir” and it has already been excluded. On 14 occasions, the form “fue” was as an auxiliary mainly forming a passive structure (13 examples). The plural form “fueron” was present 29 times but only 3 cases of this verb acting as an auxiliary of a passive structure were registered. The *pretérito imperfecto* of this verb appeared 32 times in singular and only on two occasions acting as an auxiliary (e.g. “era controlado”, “era planificado”). Its plural form (“eran”) appears 10 times in two cases acting as an auxiliary of a passive structure.

The future tense of the verb “ser” appears a total of 15 times in its singular and plural forms “sera/serán” and it only forms part of a passive structure on two occasions.

Finally, adding up the plural and singular instances of the conditional form, 73 examples are found, only 3 of them

forming a passive structure: “serían debidos”, “serían reflejados”, “serían observados”. Only one instance of the pluperfect tense was found and no instances of the compound tenses in indicative.

In the subjunctive mode the present simple tense was present 83 times (“sea/sean”), out of which on 11 occasions it was an auxiliary verb forming part of a passive structure: “sean citadas”, “sea debida”, etc.

Other tenses in the subjunctive were registered as well. For instance, the *imperfecto* in plural and singular (“fuera/n, fuese/n”) was used a total of 22 times, being only used twice as an auxiliary in a passive: “fueran debidos”, “fuera considerado”. As for the pluperfect, only two instances of the tense were found and in both of them the verb “ser” was used as a full verb: “hubiera sido de gran ayuda” and “hubiese sido mayor”. Table 3.43 below collects all the tenses and instances of the verb “ser” in the corpus:

MODE AND TENSE	VERB	FREQUENCY	%
Indicative			
Present	es/son	1091	
Present Perfect	ha (n) sido	60	0.06
Past	fue (ron)	80	0.08
Imperfect	era (n)	42	0.04
Future	será (n)	15	0.014
Conditional	sería (n)	73	0.07
Pluperfect	había sido	1	0
Subjunctive			
Present	sea (n)	80	0.08
Imperfect	fuera (n)	22	0.02
Pluperfect	hubiera (n) sido	2	0
Infinitive	ser	214	0.73
Total		1680	1.09

Table 3.43: Representation of the different tenses of the verb "ser" in the corpus

The infinitive “ser” is also mainly used (as in English) in combination with periphrasis, which correspond to English modal verbs such as “poder”. This verbal form appears mainly in the present tense “puede/n ser”, although the conditional is present as well (“podría ser”). However, out of the 44 sentences containing “puede/n ser” in the present tense only 17 were forming part of a passive structure. The conditional “podría” appears 4 times in a passive structure and 8 times in the active. Thus “poder ser” represents a total of 0.02% times in the passive in the whole corpus and it represents 0.03% in an active structure.

The combination of the infinitive “ser” and the verb “deber” occurs in the present tense (“debe/n ser”) on 22 occasions. All these instances form a passive structure, representing 0.02% of the Spanish corpus (e.g. “*Las malformaciones venosas puras deben ser tratadas por un equipo multidisciplinar en el cual juega un papel fundamental el cirujano pediátrico, ya que es en la edad pediátrica cuando deben ser diagnosticadas y tratadas,...*”). Less frequently, i.e. on 12 occasions, the combination of “deber” with the verb “ser” is part of an active structure (e.g. “*El seguimiento debe ser a largo plazo, pues se trata de lesiones no involutivas y con un comportamiento impredecible.*”)

The conditional form of the verb “deber” also occurs with the infinitive “ser”. It is used both in the passive (on 7 occasions) and in the active (in 8 sentences) (e.g. “*Por tanto, lo que debería ser neuroprotector para el cerebro, se convierte en un agresor externo.*”)

The future form “deberá/n ser” occurs in 7 instances and only in two of them in the passive.

Finally, the verb “parecer” is commonly used with the infinitive form “ser”. Researchers resort to this combination to make statements softer not only in Spanish but also in English. There are 12 instances in the corpus in which the form “parece/n ser” is registered. This structure is equivalent to the English “seem to be”, “appear to be”, “are likely to be”, etc. The difference between the expressions in the two languages is that all the Spanish examples were in active structures (e.g. “*Parece ser que también en esta variable la edad pudiera ser un factor decisivo a la hora de provocar cambios.*”) whereas in English both active and passive structures occurred (e.g. “*The results seem to support the view that the prevalence of bipolar depression does not appear to be as strongly affected by social and cultural changes...*”).

In Table 3.44 the most typical periphrasis in the Corpus with the verb “ser” are displayed:

VERB	FREQUENCY	%
poder ser	59	0.6
deber ser	44	0.4
parecer ser	12	0.1

Table 3.44: Most recurrent periphrasis formed with the Spanish infinitive “ser”

After the above detailed analysis of the appearance of the verb “to be” and “ser”, it can be concluded that the verb “to be” is a total of 2, 225 times (3.78%) in a corpus of 58, 865 words

and the verb “ser” is a total of 1, 680 times (1.61%) in a corpus of 104, 194. Obviously the presence of this verb in English is much higher than in Spanish due to the fact that it is much more widely used as an auxiliary verb, especially in passive structures. In fact, taking into account the instances in the analysis above in which the verb “ser” is used in a passive structure only 0.11% of the corpus is represented. In comparison, data on English passive account for the dramatic higher use of the passive voice in English were it represented over 1.93% of the corpus.

Eb) The verbs “to have”

The verb “to have” appears a total of 394 times. It was presented in different forms; 201 instances of the form “have” were registered, 143 cases of “has” and 42 sentences of the past form “had”.

This verb appears as a modal verb, as a full verb and as an auxiliary verb. The modal “have to” appears 8 times in different tenses. Regarding the appearance of the verb “to have” as a full verb, it was present on 143 occasions, i.e. there are 75 examples with the present tense in the plural (“have”) (e.g. *“These findings have several important implications.”*), 39 examples of the singular (“has”) (e.g. *“Hypnotic medication has beneficial short-term effects, but insomnia symptoms reappear once treatment is withdrawn.”*) and 29 of the past form “had” (e.g. *“Both fewer revisions and fewer overdrainage complications occurred in children who had a medium- or high-*

pressure valve in place during the first 6 months after shunt insertion.”). The verb “to have” acts as an auxiliary verb in 235 instances and (0.40%) but has a representation of 0.24% as a full verb. Table 3.45 below summarizes the different functions of the verb “to have” in the corpus with its corresponding rate of appearance:

FUNCTION	FREQUENCY	%	TENSE	VERB	FREQUENCY	%
modal	8	0.01		have to	10	0.02
full verb	143	0.24	Present singular	has	41	0.07
			Present plural	have	77	0.13
			Past	had	31	0.05
auxiliary verb	235	0.40		have	235	0.40

Table 3.45: Representation of the verb "to have" in the corpus as a modal, full and auxiliary verb

Sc) The Spanish verb “tener”

The corresponding verb “tener” in Spanish occupies the fourth place in the list with a representation of 0.32% in the corpus. Its representation is higher than in English. The main forms of this verb are the present tense in the third person plural (“tienen”), the third person singular (“tiene”) and the first person plural (“tenemos”) with 79, 59 and 14 examples in the corpus, respectively. The infinitive form was present on 61 occasions, and the gerund on 25 occasions. The next tenses in the corpus in terms of frequency are the present simple tense in the subjunctive mode (“tenga/n”), from which 24 instances were registered, the *pretérito imperfecto* (“tenía/n”) with 22 examples and the participle (“tenido”). The *pretérito indefinido* (“tuvo/tuvieron”), conditional (“tendría/n”) and future

(“tendrá/n”) were also present but with a much lower representation, i.e, with 11, 8 and 7 examples each, respectively.

The verb “tener” was followed on 58 occasions by the preposition “en” and the noun “cuenta”. In this combination the verb “tener” changes its original meaning, i.e. “bear in mind” or “take into account” (e.g. *“Otro aspecto que se debe tener en cuenta es la conveniencia de incluir en futuros trabajos un grupo control para comparar los resultados obtenidos en el grupo de tratamiento”*). Thus, the originally 329 occurrences of the verb “tener” must be reduced to 271 (0.25%). Table 3.46 shows the different tenses of the verb “tener” in the corpus excluding the instances in which the verb is followed by “en cuenta”:

MODE AND TENSE	VERB	FREQUENCY	%
Indicative			
Present	3rd pers plu: tienen	79	0.08
	3rd pers singular: tiene	59	0.06
	1st pers plural: tenemos	14	0.01
Present Perfect	ha (n) tenido	9	0.01
Preterito Indefinido	tuvo/tuvieron	11	0.01
Conditional	tendría (n)	8	0.01
Future	tendrá (n)	7	0.01
Infinitive	tener	5	0
Gerund	teniendo	5	0
Imperfect	tenía (n)	22	0.02
Subjunctive			
Present	tenga (n)	24	0.02
Present Perfect	haya (n) tendio	2	0.002
Total		245	

Table 3.46: Representation of the verb "tener" in the corpus

Eb) The auxiliary verb “to have”

The form “have” acts as an auxiliary forming the present perfect tense and passive structures in the perfect tense in 101 cases in the corpus. Out of the 101 instances in which “have” is an auxiliary, the main combination of the auxiliary “have” is the

participle “been” either in the present perfect tense, were 5 examples were recorded (e.g. “*Furthermore, there have been reports that the mother’s psychiatric history or the mother’s experience of loss in childhood has an effect on the postnatal depressive symptoms*”) or in a passive perfect structure, where 29 instances were recorded (e.g. “*Involvement of the superolateral orbit and extension to the anterior or middle cranial fossa or laterally to the extracranial region have been demonstrated*”). An example of the present perfect continuous was also in the corpus (e.g. “*Lately, we have been undertaking medial advancement of both hemi-orbits, which provides adequate correction of the hypertelorism*”). However, apart from the participle “been” other verbs that often appear in combination with the auxiliary to form the present perfect are “shown”, “found” “reported” and “demonstrated” (67 examples).

Finally, the last structure in which the verb “to have” was registered is in combination with “may” forming “may have”, which was used on 5 occasions in the corpus.

Regarding other tenses such as the conditional and the future, there is almost no representation of this verb since only one instance of the form “would have” and 3 of “will have” come up in the corpus.

The form “has” functions mainly as an auxiliary forming the present perfect tense. In fact, 104 cases of this tense were registered (e.g. “*The need to include consideration of function in daily life, productivity, performance of social roles, intellectual capacity, emotional stability and wellbeing, has emerged as a*

crucial part of clinical investigation and patient care”). It is also common to find it forming a present perfect passive structure.

The most recurrent combination of the form “has” in the perfect tense is when it is followed by the participle “been”. It is recorded on 65 occasions out of which only on 12 occasions the participle “been” functions as a full verb keeping its semantic meaning (e.g. “*High-level penicillin resistance of the streptococci responsible for endocarditis has been increasingly common and medical therapy with vancomycin had a good response.*”). In the other 53 instances left the participle “been” is forming a passive structure and it combines especially with the participle “shown” (e.g. “*Partly because it has been shown that the duration of the prodromal phase and the DUP have a negative impact on improvement rates during inpatient treatment [3, 24], partly because...*”).

On the 39 occasions left, the verb “has” forms part of a present perfect tense different from “been”, being the most common “become” and “performed” (e.g. “*...the diagnosis of subdural effusion, subdural empyema, brain abscess and hydrocephalus has become easy.*”). In Table 3.47 below all the instances of the auxiliary verb “to have” have been summarized:

AUXILIAR + PARTICIPLE	FREQUENCY	VERB	VOICE	FREQUENCY
have+participle	101	been	active structure	5
			passive structure	29
		shown found reported		
		demonstrated		
has+p participle	104	been	active structure	12
			passive structure	53
		shown become performed		

Table 3.47: The verb "to have" as an auxiliary verb in the Present Perfect and the verbs that occur with it

Sb) The Spanish auxiliary “haber”

The Spanish auxiliary “haber” cannot function as a full verb, but only as an auxiliary forming the present perfect tense, pluperfect, future perfect and conditional and the corresponding forms in the subjunctive. Thus it is equivalent to the verb “to have” in its function as an auxiliary verb. The representation of “haber” in the corpus is 0.62%, thus superior to the representation of the auxiliary “have”, which only reached 0.40%.

The most common tense with the auxiliary “haber” is the present perfect, which appears 256 times in the singular (“ha”) followed by a participle of the verb “ser”, i.e. “ha sido” (e.g. “*Con el tipo de stent utilizado ha sido posible intentar el implante directo en todos los casos, sin riesgo de pérdida o desplazamiento...*”). This tense form occurs 28 times, out of which only on 6 occasions it is followed by a second participle forming a passive structure (e.g. “*Si hasta ahora, el embarazo no planificado ha sido asociado a un mayor riesgo en el control, evolución y resultado de la gestación, se convierte además en un indicador de riesgo psicopatológico en el puerperio*”).

The participle of the verb “ser” is the most commonly used to form the present perfect tense; however, other very recurrent verbs in the corpus in this tense are “mostrar” (e.g. “*El índice blush ha mostrado su utilidad en múltiples estudios multicéntricos para una mejor definición del estado de la perfusión.*”), “encontrar” (e.g. “*Shea y col. [1990] afirman que no se ha encontrado una asociación entre personalidad*”).

patológica y respuesta desfavorable a la TCC.”), “demostrar” and “observar”.

The perfect tense in the third person plural is recorded 197 times, being the most recurrent the verb “ser” again (“han sido”). The corpus contains 32 instances of “han sido”, out of which on 13 occasions the verb “ser” is followed by a second participle forming a passive structure (e.g. “*Desde diciembre de 1987 todos los casos han sido registrados y participan de los estudios de la Sociedad Internacional de Oncología Pediátrica,...*”). From the above data it can be concluded that the most frequent tense with the verb “haber” is the present perfect representing 0.43% of the corpus.

The *Pretérito Imperfecto* occurs on 3 occasions in the so called “impersonal form” (e.g. “*En dos de los casos había una patología de base atópica que junto a la existencia de serotipos especialmente agresivos de...*”) and the pluperfect appears in 11 cases (e.g. “*Esto mismo señala el estudio de Diéguez et al. [1999], donde el 82,8% de los sujetos con una edad media de 19,8 años había empleado el preservativo en su primer coito,...*”). In the corpus there are 17 instances of the pluperfect in the third person plural (i.e. “habían” + participle) (e.g. “*Estos resultados son consistentes con la mayor parte de las investigaciones que habían medido la alexitimia en el trastorno de pánico*”).

The future perfect was located on 6 occasions, and 14 cases of the conditional and *condicional perfecto* were registered (e.g. “*Por otra parte, habrían de realizarse estudios*”).

epidemiológicos, tanto sobre la población general como sobre...”).

Regarding the subjunctive mode, the present tense was registered on one occasion and present perfect on 8 occasions in the singular and 6 in the plural (e.g. “*De hecho, hasta el momento actual no hemos hallado estudios que hayan tenido como objeto el análisis de la relación entre la depresión posparto y la maternidad no planificada que hayan empleado diagnóstico clínico según criterios diagnósticos DSM-IV*”).

The pluperfect (“hubiera+ participle”) was registered in singular and plural 8 times (“*Por otra parte, hubiera sido de gran ayuda realizar sesiones conjuntas con el cónyuge...*”). The same tense with the form “hubiese” + “participle” was only twice in the corpus.

The sum of all the instances of the verb “haber” forming a different tense from the present perfect (e.g. *Imperfecto*, pluperfect, etc) is 37 examples, 0.04% of the corpus. Thus, it can be concluded that the present perfect tense is the most widely used in both languages with 0.35% (205 instances) in English and a slightly higher percentage in Spanish, i.e. 0.43% (453 instances).

However, regarding the presence of passive structures in the present perfect tense, the percentage is lower in Spanish since it appears less than half as much as the English, i.e. a rate of 0.08% of the corpus vs. less than 0.03% in Spanish.

Sd) The verb “poder”

After the analysis of the main auxiliary verbs in both languages, modal verbs come as the next most recurrent category of verbs. In Spanish the verb “poder” is listed in the fourth place of the list with a rate of appearance of 0.50%. In English two modal verbs equivalent to “poder” come up, “may” (0.55%) and “can” (0.41%).

The verb “poder” is documented 520 times. The highest percentage of representation of this verb corresponds to the present tense. The third person singular (“puede”) is registered 171 times and the third person plural (“pueden”) 104 times. The first person in the plural (“podemos”) is the next in terms of frequency with 50 instances in the corpus. The two aforementioned verbal forms collocate mainly with the verb “ser” in the form “puede/pueden ser”, but the singular form of the verb “poder” (“puede”) also collocates in the corpus with other infinitives such as “tener”, “afirmar”, “resultar” and “concluir” (e.g. “*Respecto a los resultados obtenidos, se puede concluir que:...*”).

The first person in the plural (“podemos”) also collocates with other infinitives such as “concluir”, “afirmar” and “decir” (e.g. “*Con respecto a las características de la muestra de familiares y usuarios estudiadas, no podemos afirmar que alguna de ellas sea relevante en cuanto a la modificación de los conocimientos, actitudes y expectativas*”).

The next verbal form in terms of frequency is the conditional (“podría”), from which 69 instances were recorded.

This form collocates mainly with the verb “ser” and also with “explicar” (e.g. “*Este incremento podría ser producido por cascadas neuroquímicas, de las cuales, el circuito de los aminoácidos excitatorios (Glutamato) y sus receptores (NMDA), hipertrofiados tras superar el nivel defensivo, serían excitotóxicos para las neuronas centrales (y periféricas)*”). The third person in the plural of the same tense (“podrían”), is also frequent, it occurs 25 times.

Regarding the subjunctive mode, the present simple (“pueda/puedan”) appears 34 times. Next in terms of frequency is the infinitive “poder” with 27 examples. The least relevant forms according to the rate of appearance are the *imperfecto de subjuntivo* “pudiera/n” (15), the participle “podido” (15), “podrá/n (5) “pudo”/”pudieron” (3) and “podía/n” (2).

In sum, it can be concluded that the verb “poder” appears in the Spanish corpus mainly in the present tense or in the conditional and it collocates with the verb “ser” or other verbs introducing a statement like “afirmar”, “decir”, “explicar” or “concluir”. A summary of the appearances of the verb poder and the most common verbs that appear with it in the different tenses is displayed in Table 3.48 below:

TENSE	VERB	FREQUENCY	VERB
Present			
3rd person singular	puede	171	ser
3rd person plural	pueden	104	tener
1st person plural	podemos	50	afirmar
			resultar
			concluir
			decir
Conditional	podría/n	69	explicar
Present Subjunctive	pueda/n	34	
Infinitive	poder	27	

Table 3.48: Most common tenses in the corpus of the verb "poder" and the verbs collocating with it

Ed₁) The modal verb “may”

In English, the verb “may” is the most common modal with a representation of 0.59% in the corpus. According to Sánchez (2003:231) this modal expresses possibility. It is used by researchers to express the aforementioned idea especially collocating with the verb “to be”, which is used on 121 occasions (e.g. “*This may be important in patients where the third ventricular floor is not completely translucent together with a difficult or unexpected location of the distal BA complex*”). However it also collocates with other infinitives like “suggest”, “lead”, “represent”, “help”, “benefit”, “occur”, etc. The corresponding translations suggested by Sánchez (2003:231) are “puede ser” and “pudiera ser”. Thus, the last sample sentence in brackets could be translated into “*Esto puede ser importante en pacientes en los que...*”

Although the modal “might” does not appear in Table 3.38, its representation in the corpus (despite being lower than that of other modals) is still relevant, i.e. 0.10% of the corpus. This modal is used either to express permission (i.e. when used as the past of the modal “may”) or it can also express possibility (Sánchez, 2003:205). In the examples in the corpus it is the second variant that appears and thus it would be translated with the verb “poder” (e.g. “*Hence, it might be possible that the content of the ST we used could explain our outcomes.*”)

Among the different options of translation given by Sánchez (Sánchez, 2003:205), the subjunctive or conditional seems to work in all the instances of the corpus (“podría” or

“pudiera”) as in the following example, which could be translated as “uno podría especular que...”: *“One might speculate that psychotherapeutic treatment could help offspring to cope better with the parental disease”*.

After the analysis of the sentences containing “may” in the corpus it can be stated that most of the times the equivalent Spanish verb form is either “puede” or “podría”(e.g. *“This may help the patient seek new ways of adaptation to psychological disease, enhancing cooperation with the goals of therapy and the therapist”*).

Ed₂) The modal verb “can”

The modal “can” (0.41%) is mainly used in affirmative sentences. It occurs 166 times in the corpus, however, the negative form “cannot” also appears 16 times (e.g. *“Our study design cannot rule out whether amalgam patients only react more sensitively to their body sensations or are more sensitive to smallest amounts of mercury which are well tolerated by other persons.”*). The form “can” collocates like the verb “may”, with the infinitive “be” (91 instances) (e.g. *“Myocardial channels can be created by endocardial insertion of an ablation electrode followed by temperature- regulated HF energy delivery using a catheter-based system”*). The same happens with the past tense of the verb (“could”), which occurs 77 times and collocates with the verb “to be”, on 19 occasions (e.g. *“A tendency for a decrease in heart rate variability in CABG patients could well be explained by an increased heart rate”*).

The use of the verb “can” in the corpus expresses in all cases possibility. As Sánchez (2003:2001) points out, the difference between “may” and “can” in affirmative sentences is that “*may* indica la posibilidad de que algo ocurra y *can* la posibilidad de hacer algo”. Thus in the last sentence in brackets (“*Myocardial channels can be created by endocardial insertion of...*”), the author is referring to the possibility of creating such channels, whereas in the sentence “*This may help the patient seek new ways of adaptation to psychological disease*” the writer refers to the possibility that the patient be actually assisted in the near future. Quirk (1972:97) further clarifies the difference between the two modals by means of the sentences “*The road can be blocked*”, meaning, “it is possible to block the road”, and “*The road may be blocked*” meaning, “it is possible that the road is blocked”. The corresponding Spanish translations for these two modals would be in both cases “poder”, since the way to distinguish this nuance in Spanish is by means of the subjunctive and indicative. Therefore, sentences with “can” would use the verb “poder” in the indicative (e.g. “*pueden ser creados con...*”) and those with “may” the same verb in the subjunctive mode (“*esto puede que ayude al paciente a...*”). For this reason the Spanish researcher must pay attention to the different use of the modals in English and the wider range of possibilities and correspondences.

Regarding the past form “could”, Sánchez (2003:201-203) differentiates several uses of this modal. The ones registered in the corpus correspond to two of his propositions. On the one hand, the translation of “could” for the so called in

Spanish *pretérito indefinido* or *pretérito imperfecto* (“pudo” or “podía”) (e.g. “*In another analysis of our samples, we could show that 66% of inpatients present with depersonalization, originating in age-specific identity problems*”). In the sentence in brackets “could show” would correspond to “pudimos mostrar” or “podíamos mostrar”. Another example in which the verb “could” would also be translated in the same way is the following: “*First, no one could prove how these symptoms represented depression.*” On the other hand, Sánchez suggests the translation of “could” for the conditional “podría” (e.g. “*This strategy could potentially improve local tumor control while reducing the morbidity of more diffuse external therapy.*”). For instance, in the last exemplifying sentence “could improve” would correspond to “podría mejorar”. In general terms and according to the data above it is clear that English medical texts have a higher tendency to use modals and to use a wider variety of modals than Spanish medical texts, where the subjunctive is used instead. In fact, for the verb “poder” with a representation of 0.50% in English two possibilities take place in the corpus; “may” and “can” with a representation of 0.96%.

Ef) The modal “should”

The next representative modal verb in the corpus in terms of frequency is “should”, which has two main uses (Sánchez, 2003:221): a) it is used as an auxiliary to form the conditional equivalent to “would” and b) it is used to express convenience, advice or moral duty. This second meaning is the

one used in the medical corpus. The possible translations Sánchez suggests are present simple, *pretérito imperfecto* or conditional (i.e. “debes”, “debías” “deberías”). However, he points out that he prefers the simple present for the modal “must” which is more categoric (e.g. “*You must respect the law*”- “*Debes respetar la ley*” vs. “*You should be more sympathetic*”- “*Deberías/Debías ser más compasivo*”).

The verb “should” collocates in the corpus with the verb “to be”. They co-occur 123 times mainly forming a passive structure. The most recurrent participles in this passive structure are: “considered” (e.g. “*Diagnosis of an NE cyst should be considered even if there are no segments of spinal dysraphism or even if the cysts are located dorsally.*”), “used” (e.g. “*Transplantation should be used only when and if AVR starts to fail or is not feasible in these patients.*”) and “made” (e.g. “*Therefore, specific attempts should be made to identify the location of a sample when acquiring human heart tissue, and comparisons should be made between similar regions of the heart.*”). In all the examples in brackets, the translation for “debería” based on Sánchez above would work. In fact, this is the most common equivalence of this modal in the corpus. Table 3.49 below summarizes the most common verbs collocating with the most recurrent English modals in the corpus:

MODAL VERB	VERB
may	be suggest represent help benefit occur
can	be be used
should	be be considered be made

Table 3.49: Verbs collocating with the modals "may", "can" and "should"

Sf) The verb “deber”

In Spanish, the corresponding verb to the English “should” is the verb “deber”, which occurs 223 times. However, the participle “debido/s” occurs 38 times followed by the preposition “a” and forming “debido a”. The latter expression corresponds to the English “due to”, in which the meaning of the verb “deber” as “should” is no longer recognised. This expression does not only come up in the form “es debido a”, which is the most recurrent, but it also appears in the present tense, both in indicative and subjunctive 4 times (“se debe a” and “se deban a”) (e.g. “...*aspecto que puede indicar que las diferencias en el nivel de conocimientos terapéuticos adquiridos se deban a la evaluación-intervención particular de cada caso*”). Thus, the actual representation of “deber” with an

equivalent meaning to that of the modal “should” is 205 times (1.19%).

The modal “deber” appears mainly in the third person singular of the present simple both in singular and plural. The singular form (“debe”) appears 73 times and the plural form (“deben”) was recorded in 40 instances.

The present subjunctive (“deba”) was registered only on 5 occasions.

As already stated above, Sánchez points out that “should” and “must” both correspond to “deber” in Spanish. He prefers “must” to be translated by the present simple. For instance, the modal “must” could be considered equivalent to “debe” in the following exemplifying sentence: “*Los hallazgos encontrados con respecto a la etapa de proceso de cambio en la que se encuentran los pacientes con TCA, y..., nos señalan que ésta es una cuestión específica en que debe trabajarse desde el primer momento*”. Other examples of the present tense of “deber” in the corpus are “debe tenerse en cuenta”, “se debe ser muy cauto” or the complete sentences: “*En el niño el tratamiento agresivo precoz no quirúrgico debe ser la primera opción terapéutica en la PE*” or “*La traqueostomía como opción terapéutica debe desterrarse en el manejo de estos enfermos*”. According to Sánchez (2003:221) all the examples above would be better translated by the auxiliary “must”, which also corresponds to the Spanish “deber”.

The modal “must” has a representation of 0.05% (see appendix III), which is actually a low percentage in comparison to the other modals. It is therefore important to underline that,

although the present of the verb “deber” often corresponds to the English “must” (e.g. “*El tratamiento del angioma subglótico debe ser multidisciplinar. Las múltiples opciones terapéuticas deben utilizarse personalizadas para cada paciente*”), on some occasions “should” seems to be another possible equivalent for the present tense of “deber” in the corpus. In fact, Sánchez explains their difference is based on how categorical the modal is, but it is difficult to assess how “categorical” the Spanish verb “deber” is in some examples (e.g. “*Se debe recordar que un factor protector es aquél que disminuye la probabilidad de que una enfermedad aparezca, y es fundamental identificarlos*”). In any case, the present of “deber” corresponding to “must” is twice as frequent in Spanish as in English with a percentage of 0.11%.

The conditional (“debería”) appears 69 times, which corresponds to 0.07% of the corpus. The third person singular (“debería”) and plural (“deberían”) and occasionally the first in the plural (“deberíamos”) were in the corpus. The conditional “debería” is equivalent to “should” in the English corpus (“*Debería de ser valorada la posibilidad de prolongar la edad de asistencia pediátrica más allá de los 14 años así como el establecimiento...*”).

Other tenses registered are the future and *imperfecto de subjuntivo* and the gerund and infinitive, but all of these tenses only sum up to 18 examples all together, thus these tenses are not relevant.

Ee) The verb “to do” and the verb “realizar”

The next verb in Table 3.38 is “to do”, which acts mainly as an auxiliary verb. In over 90% of the cases it is followed by the adverb “not” forming a negative sentence. However, there are a few exceptions; on very few occasions it functions as a full verb (e.g. “*We were not able to do this in the current study.*”). The form “did” is used on 4 occasions as a means to emphasize (“...*whereas the earlier study of HIV-positive gay men carried out in The Netherlands [9] did find that a GI resulted in greater distress reduction than a waiting list control condition*”). Consequently, the verb “to do” must be mainly considered an auxiliary in its English function in the corpus and not a synonym of the verbs “realizar” and “hacer” in the Spanish corpus.

It is additionally impossible to establish a unique correspondig verb in English in terms of semantics to the verb “realizar” because this verb collocates with many different nouns such as “estudio”, “selección”, “seguimiento”, “evaluación”, “control”, “tratamiento”, “visita”, “análisis”, “trabajo”, “prueba”, “sugerencia”, “actividades”, etc. The corresponding English translations would imply the need of different verbs and even of completely different structures of the language.

The verb “realizar” appears 218 times (0.21%). Together with the verb “tener” it is one of the most recurrent full verbs. In fact, they are even more recurrent than some modals. The verb “realizar” appears mainly in the present simple and in the third person singular (“realiza”) and plural (“realizan”). A total of 40

instances were recorded of the latter two forms. The infinitive “realizar” is documented 73 times in the corpus and it collocates with the noun “estudios”. But the most recurrent verbal form is the participle “realizado/a” both in singular and plural, which appears 99 times. The participle collocates with “estudio/s” on 23 occasions, with “investigación” 12 times and with “trabajo” 3 times.

Other verbal forms like the *pretérito imperfecto* (“realizaba/n”), the future (“realizará”) , the *Indefinido* (“realizó”), the present (realice/n) and *imperfecto* in the subjunctive mode (“realizara”) and the gerund also occur but they only come up on 15 occasions all together.

The verb “hacer” is a synonym of “realizar”. It is less widely used in the corpus, and, although an instance of “hacer estudios” appears, it is more common to use “realizar”. In fact, the verb “hacer” is used with a more general sense.

The most common forms of this verb are the present simple in the third person singular and plural (“hace” and “hacen) and the infinitive (“hacer”). On 14 examples “hacer” collocates with “referencia”. This shows the tendency to establish relations when writing medical texts and at the same time it gives rise to a different meaning. The combination “hacer referencia” corresponds to “make reference” or “regarding” rather than to the verbs “do” or “make” (e.g. “*Por lo que hace referencia a los hábitos tóxicos, casi el 44% se ha emborrachado alguna vez y el 21,2% declara haber consumido drogas ilegales, durante el último año*”). Thus, although initially 118 forms of this verb were registered, it actually has a

representation of 0.10% with 104 instances in which the lexical meaning of “hacer” is recognised.

Sg₁) The verb “presentar”

The verb “presentar” is recorded 156 times and represents 0.14% of the corpus. Together with “realizar” it is one of the most recurrent verbs in the corpus after auxiliaries and modals. It mainly appears in the present tense. In fact, 56 instances of the third person plural were recorded (“presentan”) and 36 instances of the first person in the plural (“presentamos”) were in the corpus. However, only one instance of the first person in the plural was recorded (“presentamos”). As already observed with other Spanish verbs, the first person in the plural is the least common of the 3 possibilities stated above. All the instances in the present tense in the indicative sum 94 instances versus only 8 in the subjunctive mode, in the third person plural (“presenten”) (e.g. “*El hecho de que las personas con TAG presenten un nivel significativamente menor de alexitimia que los pacientes con pánico parece especialmente relevante*”). Although the singular form “presente” appeared 49 times, it only functioned as a verb in the third person singular on two occasions. In fact, it mainly functioned as an adjective, collocating with “estudio” on 17 occasions (e.g. “*Algunas de las limitaciones del presente estudio serían las siguientes:..*”), and with the noun “trabajo” on 15 occasions (e.g. “*En el presente trabajo vemos como el 39,2% de los sujetos con una edad media de 17,06 años...*”).

Other tenses of the verb “presentar” appeared much less in comparison with the present tense. For instance, the infinitive “presentar” appears 15 times, the participle in all its forms, i.e. singular, plural and feminine and masculine 13 times and the gerund 7 times. The least recurrent tenses were the *indefinido* with 6 instances, the *imperfecto de subjuntivo* with two. Additionally, differently to what happened with other verbs, the conditional was seldom registered, i.e. it is recorded on 3 occasions.

The verb “to present” comes as a synonym of the verb “to show” in the *Collins English Thesaurus* (1995). Similarly, the verb “presentar” appears as a synonym of the verb “mostrar” in the *Diccionario razonado de sinónimos y contrarios* (1989). *The Collins Spanish Dictionary* (1994) presents “to show” as a translation for both “mostrar” and “presentar” (1994). The equivalence in meaning between these two verbs can be observed in the following sentence in which “shows” could replace “presents”: “*This paper presents the management of 23 children with congenital AAD with severe neurological deficits and the several problems encountered during the care of such patients...*” The same could be affirmed of the following Spanish sentence containing the verb “muestra” in which “presenta” could have been used instead: “*En cualquier caso, hay que indicar que el estudio muestra cierta coherencia, ya que los criterios utilizados para determinar los problemas psicológicos u objetivos de intervención han sido aplicados para estimar el cambio o la mejora de los mismos.*”

The English verb “to present” appears in place fifteen in the list with a representation of 0.12%. Thus, its frequency is similar to the Spanish “presentar” (0.14%). However, “to present” is much less commonly used than “to show” (0.19%). The Spanish “mostrar”, has the lowest representation of all (i.e. 0.11%).

Sg₂) The verb “mostrar”

The verb “mostrar” appears mainly in the present tense. A total of 42 instances of the third person plural (“muestran”) were registered and 142 instances of the form “muestra” appeared. However, only 20 out of the 142 were acting as the third person singular of the verb “mostrar” (e.g. “*Nuestro trabajo muestra que algunas variables presentes en las primeras 24 h del ingreso, como la edad, la diabetes mellitus, la historia de infarto de miocardio previo y...*”). In fact, the most common function of “muestra” in the corpus is a noun (e.g. “*La principal limitación de nuestro estudio es que el tamaño de la muestra no nos permite realizar un análisis en los distintos grupos de edad,...*”). Nevertheless, a total of 62 instances of the verb “mostrar” are registered in the present tense.

The participle “mostrado” is also relevant in terms of frequency, with 19 instances, and few examples of other tenses occurred in the corpus.

Although the rate of appearance of the verb “mostrar” is comparatively much lower than that of “show”, the former (unlike “to present”) forms part of the 10 most recurrent verbs

and is therefore displayed in Table 3.38. In general terms it can be stated that “show” is the most used verb of all 4. The sum of the two English verbs frequently synonyms “to show” and “to present” is 0.31% versus the sum of the Spanish “presentar” and “mostrar” which results in 25%.

Eg) The verb “to show”

Regarding the tenses of the verb “to show”, as every other verb in the corpus it mainly appears in the present simple, where 41 occurrences were registered. However, on 35 occasions it is used as a participle, since this verb is very frequently used in the present perfect (e.g. “*Some authors have demonstrated that systemic morphine can suppress noxious stimulus [16, 17], while others have shown an increased dynorphin production after opioid administration...*”) and especially in passive structures (e.g. “*UF has been shown to be successful in such patients when urine output is less than 1,000 ml/day, in relieving symptoms of pulmonary edema,...*”).

The past tense of this verb is also relevant in terms of frequency, for 36 examples occur in the corpus. It is common to see the form “showed” followed by the adjective “significant” (e.g. “*At week 12, SST showed significant between-group effects on the behavioral and QOL measures, which dramatically enhanced those of the CT phase*”).

Sh) The verb “permitir”

The last two verbs forming part of Table 3.38 in Spanish are “permitir” and “considerar”. Of the 111 instances registered in the corpus (0.11%), there are 92 examples of the present tense of the verb “permitir”. Out of these, 61 are in the third person in indicative (“permite/n”) (e.g. “*La vigilancia periódica permite la detección y tratamiento temprano de problemas añadidos como la escoliosis*”) and 31 in the same person in the subjunctive mode (e.g. “*Además, se necesitan estudios longitudinales y diseños metodológicos complejos que nos permitan comprender los procesos de influencia mutua entre estas variables*”). The rest of the tenses are therefore irrelevant in terms of frequency. Interestingly, despite the verb “to allow” coming up in English (e.g. “*This method may allow for risk stratification of patients with coronary risk factors, as some of the examinees showed less impaired response despite the presence of well-established risk factors.*”), it has a very low rate of representation, i.e. 0.02% of the corpus. Other possible equivalences to the Spanish verb “permitir” like “let” or “permit” are almost inexistent in English.

Si) The verb “considerar”

Regarding the verb “considerar”, its rate of appearance is similar to that of the English verb “consider”, i.e. 0.11% and 0.1%, respectively. However, the Spanish verb forms part of the 10 most used verbs in the Spanish corpus, whereas the verb “to

consider” does not. In fact, other verbs like “to improve” or to support” precede this verb in terms of frequency.

Surprisingly, the present tense is not the most usual verbal tense of the verb “to consider”. This tense was registered on 19 occasions. Of the past participle, for instance, 38 examples were collected. In fact, there are 28 instances of the form “be considered” collocating mainly with modal verbs and especially with the modal “should” in the form “should be considered” (e.g. “*When vital structures are involved or threatened, total or subtotal resection should be considered*”). This verb again shows that the passive is widely used in English.

In Spanish the most recurrent verbal form is the present simple but, unlike what happened with other verbs, the first person plural (“consideramos”) is the most recurrent form. “Consideramos” is often followed by “que”, concretely, in 21 instances (e.g. “*Por tanto, consideramos que se ha de refinar aún más el conocimiento que se tiene de la patología*”). The third person plural (“consideran”) is recorded 12 times, making the representation of the present tense of “considerar” 0.03% of the corpus.

Another recurrent verbal form is the infinitive “considerar” with 31 examples (e.g. “*Es lógico, pues, considerar que el propio trastorno psicopatológico es una fuente de distorsión de la validez de las encuestas en este tipo de población*.”). Thus, it can be stated that, whereas in English the participle of the verb “to consider” is used especially in passive structures, in Spanish the verb “considerar” is used in the present simple and in the infinitive.

Ej) The verb “to suggest”

The verb “to suggest” has a representation of 0.18% in the corpus and occupies the eighth place in the list. It appears on 91 occasions in the present tense, mainly in the third person plural (“suggest”). The verb is very often followed by “that” (e.g. *“The results of our study suggest that there were only small differences between men and women in the use of effective cardiac medications in patients with AMI.”*) and it collocates with the nouns “study”, “results” “data” and “findings”, from which 13, 10, 8 and 6 instances were registered, respectively. The participle “suggested”, present in 19 examples, is often embedded in a passive structure (*“Furthermore, it is suggested that TD is related to hypofrontality and to hypodopaminergic state in prefrontal areas.”*).

A possible equivalent verb in Spanish is “indicar”, which also collocates with “study” and “data”, but its rate of appearance is much lower, i.e. 0.07% of the corpus. However, “indicar” is actually equivalent to the verb “to indicate”, which also collocates with “findings” (e.g. *“The findings of this study indicate that the extent of burns is not so important when compared to the possibility of disfigurement from the point of risk of developing a serious psychiatric disorder.”*), “study”, and “results” in the stated order regarding frequency, and has 0.11% percentage in the corpus.

The verb “to increase” appeared initially to be one of the most recurrent. However, “increase” functions as a noun on 30 occasions and collocates with the adjective “significant” very

often (e.g. “*Hakkola [1] showed a significant increase in cases of depression, fatigue and aggression in drivers of tankers containing hydrocarbons.*”). Additionally, the form “increased” is registered 59 times, but the past participle “increased” only occurs on 13 occasions. In the 46 examples left, “increased” was preceding a noun and thus functioning as an adjective. With this latter function it collocates on 10 occasions with the noun “risk” forming “increased risk” (e.g. “*In a patient population with AVS, it is possible to identify a subgroup of patients with mixed nodular and diffuse sclerosis, who are at increased risk for CAD including multivessel disease.*”). Therefore, the total number of instances registered where “increase” functions as a verb is only 37 and thus it only represents 0.06% of the corpus.

Ek) The auxiliary verb “will”

The ninth place in Table 3.38 corresponds to the auxiliary “will”. It represents a verbal tense and has no semantic meaning, but since its use is relevant it is worth devoting it some attention. Logically, in conclusion sections researchers tend to express how their findings could be improved in the future or what action should be taken in the future. This accounts for the high rate of appearance of this auxiliary that represents 0.13% of the corpus. It collocates with the verb “to be” most of the times followed by an adjective (e.g. “*It will be important to monitor the ICD death categories of undetermined and accidental deaths in addition to suicides, particularly in relation to drug overdoses*”). The auxiliary also collocates with the noun

“studies” to express the idea stated above and thus combines with verbs like “to direct”, “to determine”, “to need”, “to demonstrate” and the like (e.g. “*Additional studies will be directed toward the comparison in the diagnostic and predictive power between this wavelet method and other conventional measurement of HRV*”). The verb “to provide” is also registered after “will” to express a future hope for better results (e.g. “*Such studies will provide an important framework for treatment protocols involving patients with poor LVEF in the future*”). The future in Spanish is much less frequent in general terms.

E1) The verb “to need”

The last verb in English forming part of the most recurrent verbs is “to need”. This is not surprising, since one of the most recurrent actions of researchers in their conclusions is that of underlying the need to carry out research in a certain direction. This verb is not only used as a verb but it is also registered as a noun conveying the same idea (e.g. “*The need of new approaches to ambulatory care, encompassing the biological, psychological and social aspects of human disease, has emerged*”). In fact, the form “need/s” is documented 46 times in the corpus as a noun both in the singular and plural, and it only appears 34 times as a verb in the present tense (e.g. “*However, our study suggests that these interventions need to be focused on the couple, because both male and female might benefit from the psychosocial support and hence...*”).

The most repetitive structure with the verb “to need” is: “need”+ “to be” + “participle”, as in the example in brackets above. In fact, it appears 24 times.

The most typical structure is the passive, since out of the 39 occurrences of “needed”, 30 come after the verb “to be” in the form “is/are needed” (e.g. “*Longitudinal research is needed to understand clearly the interactions between normal developmental variations in children and the risk factors of mental health problems...* or “*Prospective studies are needed to determine the predictive value of the ECG strain pattern in native Africans*”). Interestingly, the verb “necesitar” in Spanish has a very low rate of appearance, i.e. 0.03% of the corpus because Spanish uses the noun “necesidad” and the adjective “necesario” in structures like “es necesario” instead.

After this analysis, it can be concluded that medical English Conclusions are dramatically less wordy than medical Spanish Conclusions. The higher rate of appearance of the Spanish conjunctions “y/e” and “si” accounts for this fact.

The most common nouns used in medical texts are the same in both languages (“patient/s” and “paciente/s”, “study/ies” and “estudio/s”, “treatment/s” and “tratamiento/s” and “result/s” and “resultado/s”). Exact collocations in both languages are quite frequent (e.g. “high risk patients” and “pacientes de alto riesgo” or “grupo control” and “control group”). However, “zero correspondence” is also common, i.e. a clear lack of counterpart in the other language (e.g: “retirada del tratamiento” or “finalizar el tratamiento”).

In the case of adjectives, Spanish researchers have to pay attention to English collocations, especially with the adjectives “clinical” and “clínico/s”, which are combined differently in the two languages, except for the corresponding collocations “clinical practice” and “práctica clínica”

Sometimes, one term rather than another is used, for instance, in Spanish the noun “trabajo” is not so frequently used as “estudio” despite being synonyms. Similarly the equivalent collocations “results of the study” and “resultados del estudio” were registered while the Spanish collocation “hallazgos del estudio” is much less frequent.

Likewise, some adjectives are more frequently used than others in the same language. This is the case of “significant”, which is more commonly found in medical texts than “important”.

Another relevant aspect is the collocation of Spanish nouns with participles in Spanish (e.g. “estudios realizados”) versus the use in English of the present simple and passive structures (e.g. “studies are needed” and “studies show”).

The verb “to be” is used as an auxiliary verb in passive structures in more than half of the instances in which the verb “ser” occurs in Spanish. In Spanish, on the other hand, both “ser” and “estar” are used in the active voice, in the present tense and mainly combined with adjectives (e.g. “es necesario”, “es preciso”, “es posible”).

The verbs “to be” and “ser” are also widely used in both languages to establish comparisons in combination with “relacionado” and “associated”. They are also used as a means

to express “hedges” both in English (e.g. “appear to be”, “seem to be”, “be likely to”, “be shown to”) and Spanish (e.g. “parecer”).

Finally, it can be concluded that English medical texts have a higher tendency to use modals and to use a wider variety of them.

3.3.2 Linguistic Analysis of the Moves

As already explained in previous chapters, one of the aims of this research work is to find the recurrent lexicogrammatical structures representing the different moves. This was done in English and Spanish with the idea of establishing similarities and differences of the language used and to create a list with possible equivalences. With this purpose, medical Conclusion Sections in the current research work were divided into 4 moves: *Background*, *Summarizing the study*, *Evaluating the study*: indicating *Disadvantage / Limitation* and *Further Research*.

In order to contrast the most recurrent expressions used in the two languages for each move, 8 lists corresponding to the 4 moves in both languages were created (see appendix III). The linguistic analysis was performed on the basis of the results of the aforementioned lists in which simple words or expressions conveying the same idea were collected under a heading that was given the same or an equivalent name in the two languages and the same number. Thus, for instance, in the move *Background*, the first heading in both languages represents time

expressions present in the corpora. The heading's name is "time expressions" in English and "Expresiones de tiempo" in Spanish, and both are numbered with number one in the list.

At this point it is important to clarify that the complete reference of the exemplifying sentences given to illustrate the linguistic analysis are not given. Thus, the discipline, the name of the journal and the volume and pages from which the example has been taken is not written in brackets after the exemplifying sentence. The main reason that explains this lack of information is that specialties are not differentiated in the analysis; it is general medical English that has been analyzed. In fact, specific words belonging to a certain medical field that could have altered our results regarding frequency have been rejected, for this reason specifying the journal was not so relevant. For instance, in the study the noun "children" was not taken into account when analyzing the most recurrent nouns because it belongs to the field of Paediatrics rather than to general medical English.

3.3.2.1 *Analysis of the move Background*

As it was explained previously in this research work, there are normally references in the Conclusion Section of Research Articles to the work performed by other colleagues. For researchers comparing and contrasting their results with that of others is extremely important and this is what the move *Background* reflects. When comparing English and Spanish, the word fields common to both languages were given a number

from 1 to 4. These 4 common semantic fields in the move Background are displayed in Table 3.50 below:

1	Time Expressions/Expresiones de tiempo
2	Comparing with previous studies/Comparación con estudios previos
	2.1 General Structures
	2.2 Analysis of the nouns, adjectives and adverbs
3	Literature/Literatura
	3.1 Neutral literature
	3.2 Comparing literature
4	Have demonstrated-shown/Han demostrado-señalan-afirman

Table 3.50: Common word fields in the move Background in English and Spanish

1. Time Expressions-Expresiones de tiempo

The most relevant aspect in the move *Background* is that it presents a wide variety of tenses to refer to the past. These tenses occur most of the times in combination with time expressions. Although the most recurrent tense in this move is the present perfect in English and the corresponding *pretérito perfecto* in Spanish, the present tense appears in both languages, and also the simple past in English and the corresponding *pasado simple* and *pretérito imperfecto* in Spanish.

Researchers frequently want to express the state of the art in order to introduce and justify their research work. In this revision of what has been done by others they many times refer to an up-to-now existing lack, deficit or problem that they later on will explore in their paper. At this point it is not surprising that certain overlap takes place sometimes between the move *Background* and the move *Limitation*.

a) In the past - Últimos años, durante los últimos años

It is therefore common for them to use words such as “limit” in combination with time expressions such as “últimos años” to describe what the situation has been in the past and refer to the contribution they make to that situation (e.g. “*Asimismo, los resultados tienen implicaciones para la investigación cognitiva de los últimos años que ha limitado la generalización de Devine [198]), apoyando ...*”). The expression “últimos años” in Spanish is very often preceded by the preposition “en”; “en los últimos años” or by the preposition “durante”: “durante los últimos años” (e.g. “*Durante los últimos años hemos incorporado la resincronización a nuestro arsenal terapéutico para el tratamiento de pacientes con insuficiencia cardíaca*”).

The corresponding literal equivalence to “últimos años” in English is “past years” (e.g. “*In the past 60 years, psychosomatic medicine has addressed...*”). Also a much less defined expression, “in the past”, occurs in English (e.g. “*Findings in the past were put to use in critically reviewing the ICD-10 and DSM-IV classification, in which compulsive acts and obsessive thoughts were placed at the same level in diagnosis*”). This more general reference to the past is present in the Spanish corpus as well: “en el pasado” (e.g. “*La retracción de la zona del plano nasal para cicatrización por segunda intención, que fue descrita en el pasado como un inconveniente de la técnica, no afectó la correcta colocación del esfínter velofaríngeo en la gran mayoría de los casos*”)

b) Decade-Década

Regarding the construction “últimos años”, it is noteworthy that, while it is widely used in the Spanish corpus, in English the noun “decade” is normally preferred (e.g. “*Over the past decade, substantial progress has been made in this field*”). The preposition “over” present in the last exemplifying sentence would be the literal equivalent to the Spanish “durante”. “Over” is sometimes combined with the noun “period” (Eg. “*Over a 10-year period, there was no change in ...*”). As for the literal equivalent noun to “decade”, “década”, there are also some instances registered in the Spanish corpus (e.g. “*En la última década se ha puesto énfasis en la investigación de los fenómenos...*”). Many different words appear in the Spanish corpus determining the word “década” either preceding the noun: “en las dos últimas”, “en la última década”, “a finales de la década de los 80” or following it “en la década de los noventa”. However, only one example of the word defining “decade” coming after the noun was registered in English, namely, “two decades ago”.

c) Ago- Hace

As the last example shows, another time expression in English in consonance with the noun “decade” is “ago” (e.g. “*Two decades ago, Reiser anticipated the success of ...*”). “Ago” is often used in the corpus to refer to the past (e.g. “*A characteristic complication of a transmural myocardial*

infarction was described for the first time 48 years ago”). The literal correspondence in Spanish, i.e. the preposition “hace” is also present in the Spanish corpus, although it is not used frequently (e.g. “*Los IAC constituyen en el momento actual una opción farmacológica interesante que hubiese resultado insospechada hace sólo diez años*”).

d) Now, up to now, to date - desde, hasta, hasta la fecha, revisión

The preposition “desde” is commonly used in the Spanish corpus to point out a date in the past when something started (e.g. “*Desde el año 1993 con la reintroducción de la clozapina y la aparición de risperidona en 1994 se inició el tratamiento con antipsicóticos de...*”). Interestingly, no instances with the literal equivalent “since” were found in the English corpus.

A very common structure to express limitation regarding previous research is by means of the preposition “hasta” in combination with the noun “momento” forming “hasta el momento” or, alternatively, the same preposition together with the adverb “ahora” in “hasta ahora” (e.g. “*De hecho, hasta el momento actual no hemos hallado estudios que hayan tenido como objeto de análisis...*” or “*Si hasta ahora, el embarazo no planificado ha sido asociado a un mayor riesgo en el control, evolución y resultado de la gestación, se convierte además en un indicador de riesgo psicopatológico en...*”). When researchers use the above time expressions, the limitation in the move is

most of the times attained by means of words with an inherent negative connotation such as the adverb “no” in the sentence in brackets, i.e. “*hasta el momento actual no hemos hallado*” or in “*no ha sido descubierto hasta el momento*”. Other words co-occurring with these expressions that help express the limitation are the adjective “*escasas*” combined with “*evidencias existentes*” or the verb “*limitar*” (e.g. “*En general, la literatura sugiere que..., como consecuencia de las escasas evidencias existents hasta el momento sobre...*”). The expression that is found in the corpus with the same meaning in English is “up to now” (e.g. “*Up to now, it is known that inhibition function involved in working memory...*”). Alternatively, the same meaning can be conveyed by means of the noun “date” together with the preposition “to” resulting in the structure “to date” (e.g. “*Despite the abundance of positive reports published to date, the association between SNPs and CAD has been inconsistent, with the exception of...*”). The literal equivalent in Spanish was also present in the corpus with the noun “*fecha*” and the preposition “*hasta*” i.e. “*hasta la fecha*”, which affects “*trabajos realizados*” in the following sentence: “*La revisión de los trabajos epidemiológicos de TCA realizados en España hasta la fecha, ha puesto de manifiesto una serie de limitaciones metodológicas*”. This example shows the presence of the noun “*revisión*” that collocates with the expression “*hasta la fecha*” (e.g. “*Recientemente, la revisión de Fazio y Olson [2003] ha llegado a una conclusión semejante, pero...*”). The equivalent English term “revision” is not used in this move, although it is present in the corpus. In English, it does not collocate with time

expressions but with other technical medical words (e.g. “*That VPS underdrainage in a patient with communicating hydrocephalus with a functional LPS resulted in ventriculomegaly and revision of the VPS resulted in resolution of symptoms, likely indicates that lumboperitoneal shunting...*”).

e) Recently- reciente, recientemente (estudios realizados /publicados- report)

A very typical adjective that researchers use to refer to studies that have been carried out but not so much time ago is “recent”. The adjective collocates with the nouns referring to investigation such as “trial” or “study” (e.g. “*A recent double-blinded randomized controlled trial showed no benefit...*” or “*Recent studies have reported...*”). In Spanish, the most recurrent combination with “reciente” is the noun “estudio”. The collocation is frequently completed by the past participle of the verb “realizar”. The presence of the preposition “en” and the determinant “un” in the same collocation give as a result: “en un estudio reciente realizado” (e.g. “*En un estudio reciente realizado con muestra española...*”). However, the following example shows that the collocation “estudio realizado” is independent of the adjective “reciente”. In this case, a different element previously mentioned takes place, “revisión”: “... *afirman Aloisi et al (1995), de una revisión de resultados de estudios italianos realizados...*”. Clearly, the participle “realizados” plays an important role in the move *Background*. It combines with any word related to previous research, thus not

only with “estudio” but also with “trabajos” (e.g. “*La revisión de los trabajos realizados en España hasta la fecha...*”). In English, the presence of an equivalent verb to this participle was “carry out”, but its rate of appearance was very low in comparison to that of “realizados” in English. In fact, only 8 examples of the verb were registered. Additionally, although the verb collocates with the noun “study” (e.g. “...*whereas the earlier study of HIV-positive gay men carried out in The Netherlands [9] did find that a GI resulted in greater distress reduction than a waiting list control condition.*” or “*However, considering the difficulty in carrying out studies in similar rural settings and using more sophisticated methods...*”) it combines with other nouns like “work” and “test” (e.g. “*A patch test should only be carried out if clinical symptoms such as lichenoid lesions of the mucosa exist along with the unspecific symptoms.*”), etc.

In Spanish, it is observed that in terms of structure, other equivalences to “realizados” containing the word “estudio” were found. For instance, an alternative participle with the same function as “realizados” that combines with the noun “estudio” in the move *Background* is “publicados” (e.g. “*Aunque la mayoría de los estudios publicados hasta la fecha indican que...*”). There is an equivalent structure consisting of a noun and a participle in English. Nevertheless, the equivalent literal participle “published” does not collocate with the noun “studies” but with “reports” (e.g. “*as compared to reports published in the seventies, which led to the abandonment of the technique*”).

On some occasions, it is the noun “advances” that collocates with the adjective “recent” (e. g. “*In spite of recent surgical and radiologic advances, the indications for surgery remain controversial*”). Although the noun “avance” comes up in the Spanish corpus, no collocation with the adjective “recent” was registered.

Occasionally, the adverb “recientemente” rather than the adjective “reciente” introduces a sentence to refer to previous research (e.g. “*Recientemente, la revisión de Fazio y Olson [2003] ha llegado a una conclusión semejante, pero...*”). The Spanish adverb shows many different options regarding its place in the sentence. The use of the English adverb “recently” is widespread in this move as well, but, although it occasionally comes up at the beginning of a sentence (e.g. “*Recently, the safety and efficacy of targeted catheter-mediated myocardial gene therapy have been reported, which may have important implications for the widespread applicability of this technique*”), it is more common to find it embedded in the complex form of a verb tense like “has only recently been assessed” or “has recently been used” (e.g. “*TVI is a promising method for quantifying LV asynchronous contraction in HF patients with bundle brunch block, and has recently been used to elucidate and quantify ...*”). The position of the adverb does not obey to special rules of medical English but corresponds to the different use of the adverbs in the two languages, however as many sentences contained this adverb it has been underlined.

2. Comparing with previous studies-Comparación con estudios previos

2.1 General structures

The most typical step in this move is the comparison with other author's results. These results can be just mentioned to display the state of the art as it has been demonstrated above by means of time expressions. Alternatively, results can be compared to establish differences and similarities.

Researchers tend to compare their results with that of others to give their investigations support. The 4 most typical nouns used in these comparisons are "estudios", "datos", "resultados", "hallazgos" and "investigaciones" in Spanish and the corresponding "studies", "data", "results", "findings" and "research" in English. Those with a lower rate of appearance are the noun "hallazgos", which is less commonly used than "findings" in English, and the noun "research", which is comparatively less widely used in English than "investigaciones" in Spanish. Other nouns that occasionally take place are "cifras" and "afirmaciones" in Spanish and "definition" and "observation" in English.

a) Participles: obtenidos, realizados, publicado, expresado, encontrado- reported, obtained

Other features present throughout the move *Background* is the presence of participles, especially "obtenidos" referring to

“resultados”, giving rise to the collocation “resultados obtenidos”. Other two participles that have already been mentioned are “realizados” and “publicados”. “Realizados” combines with “trabajos” and “estudios”, (i.e. “estudios realizados” and “trabajos realizados”) and “publicados” collocates mainly with “estudios”. Finally, the participle “expresado” is also present although not so much as the other two participles mentioned above. It appears in front of the preposition “por” indicating the authors that made the statement (e. g. “*Coincidimos plenamente con lo expresado por...*”). Occasionally, the participle “encontrados” refers to “resultados” (*Los resultados son semejantes a los encontrados por Planes y Gras [2002] en estudiantes universitarios de...*), although, the verb “encontrar” is also used in the present tense.

The presence of participles in the English corpus was insignificant in comparison to the Spanish corpus. The most similar participle found with the same function as that of the Spanish participles is “reported” (e.g. “*The incidence of spina bifida in the city of Al-Madinah Al-Munawarah was similar to those reported from the Eastern Province of Saudi Arabia*”). and occasionally “obtained”, but not necessarily referring to “results”. For example, in the following sentence, “obtained” is complementing the noun “coefficients”: “*These lower internal consistency coefficients were not the same as those obtained in another study on the....*” The participle “obtained” collocates with “results”, but the collocation does not take place in the move *Background*. Normally, the “results obtained” refer to those that the researcher has obtained with his/her own research

and many times, these “results” are referred to with other nouns such as “effects”, “analysis” or “data” (e.g. “*According to the data obtained from this study, the existence of RBBB in ECG may reflect severe MS and widespread calcification over the valve*”).)

b) Expressions of “similarity”

At the beginning of this section it was explained that in RA one often comes across expressions conveying similarities or differences regarding previous studies. Next, the most common structures used to express a similarity in English are displayed in the list below:

- *to be in line with*
- *to be in keeping with*
- *to be in agreement with*
- *to be consonant with*
- *to accord with*
- *to be similar to*
- *to be the same as*
- *the adjective “comparable” + studies/ to compare*
- *to join with*

As for Spanish, there are even more possibilities to express similarity with previous studies:

- *estar en línea con*
- *concordar con/ estar en concordancia con*

- *estar en consonancia con/ ser consonante con*
- *coincidir con/ resultar coincidente con*
- *ser coherente con/ ser consistente con*
- *encajar con*
- *similar a/ De manera similar/ser similar a*
- *al igual que*
- *constatar conjuntamente con*
- *ser equiparable a*
- *semejante (a)*
- *repetirse*
- *ya observado en anteriores estudios*

The above expressions have been analysed paying attention to semantic criteria. If their semantic value is equivalent in both languages or very close, expressions have been presented under the same heading. When the meaning was similar in both languages but presented a different semantic root expressions have been presented under a different heading.

b₁) Estar en línea con - To be in line with

According to the list, a few more options take place in the Spanish corpus to express similarity. Some structures are literally equivalent and some are not. For instance, the English structure “to be in line with” is literally equivalent to “estar en línea con”(e.g. “*Our findings are thus in line with the studies of Moritz et al. (35) who elegantly applied NP tasks adapted to schizophrenia research*” and “*Estos resultados están en línea*

con los aportados por otros estudios referidos a poblaciones de otros países”.) As the Spanish example in brackets shows, it is frequent to find this structure in combination with the adjective “otros” referring to authors or studies “en línea con otros estudios” and “en línea con otros autores”. Similarly, in English, the reference to “other authors” and “studies” takes place, but in combination with the structure “in line with” the adjective “several” is registered in the corpus (e.g. “*Our results are in line with those of several authors, showing moderate agreement between patients’ and informant’s reports.*”).

A clear difference among the examples in the corpus is that all sentences containing the structure “in line with” in English start with a possessive pronoun in front of the main noun of the sentence: “our results”, “our definition”, “our findings”, whereas in Spanish the main noun is very often preceded by the demonstrative “estos” (e.g. “*Estos resultados estarían en línea con las sugerencias de otros autores de que lo importante...*”). In this case, it is the Spanish expression that makes it sound as if the researcher was more distant from his/her statement and results. The search for lack of commitment on the part of the researcher, which was already mentioned in the introduction of this research work, is underlined in this exemplifying sentence by the conditional tense “estaría”. This characteristic of the language is more typical of the English language in general. This effect is mainly attained by means of the passive voice as it will be shown further in the analysis.

A very slight variation of “to be in line with” that also comes up in the English corpus is “to be in keeping with” (e.g.

“*The results of our study are in keeping with previous data from epidemiological and...*”). In this exemplifying sentence, the adjective “previous” preceding the noun “data” underlines the move *Background* making clear there is a reference to studies carried out in the past.

b₂) In agreement with, in accord with

An equivalent structure to “in line with” in which the same meaning is conveyed is obtained by means of the preposition “in”+the noun “agreement”+preposition “with”: “in agreement with” (e.g. *These data are in agreement with previous studies that report...*). Again, the sentence contains the adjective “previous” that seems to co-occur frequently with these structures expressing concordance. Interestingly, an equivalent expression takes place in the Spanish corpus; however, it is used to convey the opposite meaning. The noun “desacuerdo” appears instead of “acuerdo”, “agreement”. Despite the fact that the meaning of the sentence is just contrary to “in agreement with”, the structure is equivalent and contains the same lexemes and the opposite lexical items. It consists of the preposition “en”+ the noun “desacuerdo”+ the preposition “con” (e.g. “... *en desacuerdo con los resultados por Pantula y Pollock...*”). The verb “to accord” is very similar in meaning to the noun “agreement” and it plays the same role in “accord with” (e.g. “*The present findings accord with the demands of other authors that a diagnostic distinction between... should be made*”).

b₃) Be consonant with - ser consonante con, estar en consonancia con

Although the structure in Spanish is slightly different, a comparable expression for the English “accord with” is “ser consonante con”. The adjective “consonante” is used with the same lexical value as the verb “to accord”, so they could be used in the same context (e.g. “*Este resultado es consonante con la evidencia obtenida en el ámbito de estudio del estrés familiar...*”).

The actual English equivalent structure in terms of syntax and lexis to “ser consonante con” consists of the verb “to be”+the adjective “consonant”+ the preposition “with” resulting in: “to be consonant with” (e.g. “*This is consonant with early clinical observations made by Meyer [31] showing that ERP could modify...*”). In Spanish, a little variation takes place in the expression “estar en consonancia con” or simply “en consonancia con” which despite being a different grammatical structure is practically the same as “ser consonante con” (e.g. “*Estos resultados están en consonancia con las afirmaciones de Skinner [1957]...*” and “*...en un estudio reciente realizado con muestra española, en consonancia con los resultados obtenidos en investigaciones de otros países...*”).

b₄) Concordar, estar en concordancia con

Regarding “concordar”, the term presents several syntactical possibilities. The most recurrent appearance is in the

form of a verb in the present simple tense, “concuerta” followed by the preposition “con” (e.g. “*El método de autoagresión utilizado de forma preferente por los sujetos de nuestra muestra también concuerda con los datos de la literatura consultada que...*”). However, many instances were registered of the structure formed by the verb “estar”+ the preposition “en”+ the noun “concordancia”+ the preposition “con” combining in “estar en concordancia con” or simply, “en concordancia con”, resulting in parallel structures to the ones mentioned in the above paragraph, that is, “estar en consonancia con” and “en consonancia con” (“*Nuestros resultados están en concordancia con los obtenidos en otros estudios, utilizando los mismos instrumentos de medida...*” and “*Observamos, en concordancia con los resultados de...*”). Finally, in relation with the term “concordancia”, one instance of the adjective “concordantes” was registered in the corpus (e.g. “*Estos hallazgos son concordantes con las sociedades científicas 2, 8, que aconsejan...*”)

b₅) Ser coherente con

The collocation formed by the noun “resultados” and the participle “obtenidos”, present in many of the afore mentioned exemplifying sentences, is also present in the instances of the next structure used to express similarity in Spanish, i.e. “ser coherente con” (e.g. “*Los resultados obtenidos en el estudio... son coherentes con la investigación preexistente en el campo de...*”). In the last example it is worth pointing out the presence

of the adjective “preexistente”, which clearly makes reference to the past. The combination of “ser coherente” with the adjective “preexistente” is perfectly exchangeable in terms of meaning with “estar en consonancia con”, despite the difference in the syntactical structures.

A very similar construction is “ser consistente con” (e.g. “*Los resultados no son consistentes con lo que esperábamos ni con la literatura [Patience, 1995] que afirma que...*”).

b₆) Coincidir con, concordar con

In terms of meaning, it can also be pointed out that the verbs “coincidir con” and “concordar con” are equivalent to the previous structures, besides being simpler from a syntactic perspective. The collocation “resultado obtenido” emerges again with the verb “coincidir con” and its variant “resultar coincidente con” (e.g. “*Estos resultados coinciden con los obtenidos por Toro-Álvarez [1991] donde no se manifiesta evidencia acerca de la relación...*” and “*Este resultado resulta coincidente con el obtenido por las...*”). The verb “coincidir” is present in different forms in the corpus: sometimes it is used in the present tense: “*Coincidimos plenamente con lo expresado por...*” sometimes the gerund form is used: “*McKelvey et al. (1999) encontraron, coincidiendo en parte con nuestros hallazgos y otros mencionados anteriormente, en...*” and sometimes the past simple tense: “*Nuestros resultados coincidieron con revisiones anteriores que señalan el entrenamiento...*”.

Many of the Spanish structures presented, that is, those with verbs “concordar” and “coincidir” and that with the adjective “coherente” or other belonging to the same word field (“concordar”, “estar en concordancia con”, “coincidir”, “resultar coincidente con” and “ser coherente con”) do not have a literal equivalent in English, but they could actually be considered equivalent in meaning to any of the aforementioned English structures containing the preposition “with” such as “to be in line with”, “to be in agreement with”, etc.

b7) Encajar con, repetirse

Two additional examples registered in the Spanish corpus are “encajar con” and “repetirse”. The verb “encajar” is practically the same as “coincidir”, “repetirse”. Although “encajar” refers rather to the fact that the same data or results take place in the different papers, it is very similar in meaning to “coincidir” and “repetirse”. For this reason, these verbs could be also replaced by “encajar con” and consequently it could be used for the translation of any of the aforementioned English structures (e.g. “*En concreto, la línea de evaluación intervención funcional diseñada para este caso encaja con la utilizada en otros trabajos*” and “*Tanto la mayor evitación del daño como la baja auto-dirección son datos que se repiten en todos los estudios....*”).

b₈) Similar- similar, semejante

An equivalent word in terms of semantics in both languages to establish comparison is “similar”. In English, its main use regarding its syntactical function and place in the sentence is the combination of the adjective “similar” with the verb “to be” in the structure: “to be similar to” (e.g. “*These findings are similar to, but less robust than, those described in previous studies of AAS effects in men*”). The equivalent Spanish structure “ser similar a” also occurs (e.g. “*Concomitantemente se observa que el porcentaje de sujetos de nuestra muestra que se autolesiona [24’3%] es similar al encontrado en un estudio de Blaauw y Kerkhof [1998]...*” or “*En nuestro análisis, estos resultados son similares, aunque algo diferentes, a otros estudios que señalan...*”), but the term “similar” often appears in the prepositional phrase with adverbial function “de manera similar” or “en un sentido similar” (e.g. “*De manera similar a trabajos previos, este estudio adolece de algunas limitaciones metodológicas, que pueden...*” or “*En un sentido similar, Nicholas y Drrheim [1996] manifiestan que...*”). Sometimes, the adjective “semejante” is used as a synonym of “similar” (e.g. “*Los resultados son semejantes a los encontrados por Planes y Gras [2002] en estudiantes universitarios de...*”).

No literal equivalent has been registered, but as “semejante” is a synonym of “similar”, it could be translated into English using the structure “to be similar to”.

The adjective “parecido”, which has a very close meaning to that of “similar”, has not been analysed for its low

rate of appearance; in fact, it was only registered on 3 occasions.

b) Al igual que- like, as

“Al igual que” is a Spanish structure with no equivalent literal structure in English. It co-occurs mainly with the adjective “otros” referring to other studies or authors (e.g. *“Teniendo en cuenta el sexo del cuidador, constatamos, al igual que otros autores, que son las mujeres las que presentan más trastornos...”*). Additionally, it also collocates with the verb “encontrar” in the past tense: *“También en esta nueva situación dependen, al igual que encontraron los autores anteriormente referenciados...”*.

The preposition “like” was found in English as an equivalent to “al igual que”, however, it did not collocate with the noun “authors” to refer to literature. For instance, in the following example “like” refers to “studies” (e.g. *“Although it is based on a retrospective design like most of the prodromal studies, it provides some important vistas on the pathogenetic process of incipient psychosis in younger age groups”*). Additionally, its rate of appearance was not very significant, i.e. only 0.04%.

Alternatively it could be resorted to the structures “to be in agreement with” and “to be in line with” as an option to translate Spanish sentences containing this “al igual que”. Then again, this structure is comparable to the English conjunction “as” in the sentence: *“As in other studies comparing therapies,*

evaluator double-blindness was not possible". Thus, "al igual que" could be translated into English using "as" and "like" depending on the sentence, but the difference between the two languages is that in English most of the times they collocate with "studies" whereas in Spanish it is more common to refer to "otros autores".

**b₁₀) To compare- equiparable, the same as-
comparar, equiparable a, convencional, tradicional**

The verb "to compare", especially in the past tense, "compared", and its corresponding adjective "comparable" were registered in the English corpus (e.g. "*Comparable studies, done prior to the current study, have been able to successfully demonstrate...*"). However, it is important to clarify that on most occasions this term is used to express a comparison between two groups, patients, hospitals, etc., mentioned in the paper being described rather than to establish a comparison with a previous paper. In fact, the most typical collocation in English is "compared to/with a control group". Regarding Spanish, there was only one example in which the term "comparar" was used to refer to previous research (e.g. "*Esto podría explicar, por ejemplo, porque los niveles de hostilidad han sido tan bajos en este estudio en comparación con otras investigaciones donde los sujetos...*"). However, it should be pointed out that on some occasions the term is combined with the adjectives "tradicional" and "convencional" and in these cases it compares a certain method, technique, and the like, currently being described or

investigated in the paper with a previous existing one. In these sentences the comparison is focused on the item compared and indirectly also on other findings in the past (e.g. “*Long-term follow-up demonstrates that the DHS is an effective means of CSF diversion with shunt survival comparable to that of traditional techniques*”). Also in Spanish the structure “en comparación” co-occurs with “convencional” with the same effect as in English (e.g. “*Los resultados de este trabajo indican que el tratamiento de la estenosis coronaria mediante stents recubiertos con sirolimus o paclitaxel puede reducir la tasa de revascularización hasta un 69%... en comparación con el stent convencional*”).

Despite the low rate of “comparable” to refer to previous research, the adjective “equiparable” in the structure “ser equiparable a” was registered in the Spanish corpus to refer to previous studies in combination with the collocation “resultados obtenidos”: “*Además, los resultados obtenidos en nuestra serie son equiparables a los del estudio recientemente publicado por Schlosser et al 13, realizado con...*”.

Finally, the structure “the same as” appears as a means to compare with previous studies, in fact it appears in the sentence with the adjective “previous” but its use is not very frequent (e.g. “*This is in the same range as most previous research using the same type of variables...*”). A possible equivalent to this structure could be either “equiparable” or “al igual que”, thus, sentences containing the aforementioned Spanish expressions could be translated using “the same as” as well. The latter

structure is often used to express the opposite “is not the same as”.

c) Expressing differences

However, researchers also express differences when they compare their investigation with previous research (e.g. “*These lower internal consistency coefficients were not the same as those obtained in another study on the...*”). For this purpose they can either use a negation with the aforementioned expressions showing similarity or they can also use the following structures.

- *Estar en desacuerdo con*
- *Diferir*
- *Contradecir/ al contrario que/En contra de*
- *Mostrar diferencias*

c₁) Estar en desacuerdo con, diferir, contradecir, al contrario que, contrapuestos- unlike, as opposed to

The first expression, “estar en desacuerdo con” has already been mentioned above as an opposite to the positive English structure “to be in agreement with”.

The most common verb used is “diferir”. The verb collocates mainly with the term “previo” in the form of an adjective but also with the adverb, “previamente” (e.g. “*El diseño del presente estudio difiere principalmente, de otros trabajos realizados previamente con el mismo instrumento...*” or

“Nuestros resultados difieren de los de estudios previos en una menor puntuación con significación estadística...”).

As far as “contradecir” is concerned, it mainly refers to “estudios” and “resultados”, which are often present in the same sentence (e.g. *“Nuestros presentes hallazgos también contradicen los otros estudios realizados con anterioridad en España”*). An alternative structure containing the adjective “contrario” is “al contrario que” (e.g. *“Nuestros datos indican que el valor óptimo de corte es más eficaz para la confirmación del diagnóstico de ICC que para su exclusión al contrario que lo señalado en estudios previos...”*), which can be considered equivalent to the English word “unlike” (e.g. *“Unlike what has been reported by other authors [19,21], we did not observe any episode of apnea or...”*).

A last alternative is the use of the Spanish adjective “contrapuestos” that comes after the noun “resultados” (e.g. *“En otro sentido, nuestro estudio informa, de resultados contrapuestos a los de Elloy...”*).

The English expression “as opposed to” was present in the corpus but it was not very frequent, in fact, only 2 instances were registered.

It is much more uncommon for authors to express differences with respect to other colleagues’ work than similarities in this move. Comparatively, it can be stated that English has a more limited variety of vocabulary than Spanish to express differences. In fact, most of the times English uses one of the following three words to express differences: “unlike”, “difference” and “compare”. The latter verb often co-

occurs with the noun “differences” to actually express a difference (e.g. “*In many studies that compare models on ... substantial differences and various weaknesses ... can be unveiled.*”).

2.2 Analysis of Nouns, Adjectives , verbs and Adverbs

a) The nouns: Study, Data- Estudio, Trabajo-Datos

It is necessary to point out that they co-occur with adjectives, many of which have a main characteristic: they possess an intrinsic connotation referring to the past. The most recurrent adjective in this group is the adjective “previous”/ “previo”. But other relevant adjectives with this characteristic in the corpus are: “prior”/“precedente”, “other”/“otros”, “early”/“anterior”, etc.

b) The adjective “previous”/ “previo” and the verbs “to show” and “hallar”

The main collocations that occur in combination with the adjective “previous” in the move *Background* in English are “studies” and “data” (e.g. “*Although previous studies have provided information on... there is a lack of data regarding...*”). “Data” collocates with the adverb “previously” especially in combination with the participle “presented” in the collocation “previously presented data” (e.g. “*Taking into account previously presented data from the ESBY study have found...*”).

Although different tenses were found in the constructions in which the adjective “previously” was present, the most common tense to be found is the present perfect as the examples in brackets above show. Additionally, the verb “to show” is also quite recurrent forming the combination: “previous studies have shown”.

Regarding the Spanish collocation “estudios previos”, it is also normal to find different tenses following it, but the present perfect is as common as in English, especially with the verb “hallar” (e.g. “*Se ha hallado en estudios previos que...*”). The verb “indicar” also takes frequently place in the Spanish corpus; in fact it is very common to see the structure “estudios previos indican”.

“Data” and “study” deserve especial attention. The noun “data” is very widely used in the move *Background* and in both languages. As they are so widely spread in the move, they also co-occur, that is to say, “data” often combines with “studies” embedded in different structures e.g. “to be in agreement with” (e.g. “*...these data are in agreement with previous studies that report diogin and...*”).

c) Data & other, literature and support-datos

“Data” also collocates with the adjective “other”, “other data”, and it is frequently followed by the noun “literature” as in “other data in the literature”. Another possibility for “data” in this move is the combination with the verb to “support” followed by the name of an author. This can be observed in

“While our data collected here support Klein’s theory advocating that...” As far as Spanish is concerned, the presence of the noun “datos” is present in almost all the typical structures that belong to the move *Background* as shown in Table 3.51:

“Estos datos son similares a los + participle”
“Estos datos coinciden con los encontrados por”
“Nuestros datos concuerdan con”
“Estos datos son coherentes con los+ participle”
“Estos datos estarían en línea con”
“nuestros datos difieren de”

Table 3.51: Structures with the noun “datos” in the move *Background*

As it is the case in English, “datos” as well as “resultados” (see paragraph below) is also present in many sentences containing the noun “estudios” like in: “*Estos datos coinciden con otros estudios en los que se ha encontrado...*”.

d) Otros estudios, resultados+estudios-otros estudios

The noun “estudio” also deserves especial attention since it does not only collocate with “datos”, but additionally with a relevant adjective in this move, “otros” giving rise to the collocation “otros estudios”. Apart from the noun “datos”, also “resultados” plays an important role in combination with “estudios” in this move.

A very common structure is: “Nuestros datos” or “nuestros/los resultados”+ the structure “estar en concordancia con” + los+participle + por/en otros estudios”. (“*Nuestros*

resultados están en concordancia con los obtenidos en otros estudios” and “Los datos estarían también en concordancia con los aportados por otros estudios como el de...”).

The close link between “resultados” and “otros estudios” is also reflected in “*Teniendo en cuenta nuestros resultados y los aportados por otros estudios similares sobre...*”.

The noun “studies” in English is comparatively less commonly used in combination with the adjective “other”, in fact, whereas 22 instances are registered in the Spanish corpus of “otros estudios” representing a 0.02% of the corpus, the combination “other studies” comes up only on 3 occasions in English (e.g. “*Only 23% of the hemorrhages were reflected in the patients’ charts, which may explain the low complication rates reported in other studies that did not analyze postoperative neuroradiological studies*”).

e) Other authors - otros autores

In English the adjective “other” comes up mainly preceding the word “authors” in the collocation “other authors”. The collocation appears mainly with the common structures already mentioned in this move such as “accord with”, “to be in line with”, and the like... (e.g. “*The present findings accord with the demands of other authors that diagnostic distinction between... should be made*”). An instance was registered with the verb “to join” (“*We join with other authors in calling for a ban on the paediatric use of...*”). Sometimes, “other authors” is used

with the present perfect tense in (e.g. *“This assumption has also been supported by other authors who postulated that...”*).

The same equivalent expressions co-occurring with the most common structures stated above appear in Spanish as well. Therefore, “*otros autores*” is placed after “*estar en línea con*”, “*diferir*”, “*concordar*”, etc. (e.g. *“Estos datos coinciden con los de otros autores 2, 6, 7 y contradicen la creencia del...”*). Occasionally, also the present perfect appears with this collocation: “*Los resultados muestran en gran parte lo que otros autores ya han expresado con anterioridad,...*”.

f) The adjectives “previous” and “prior” and “precedente”

As already stated, “previous” is the main adjective collocating with “studies”. But its relevance in the corpus is also due to its recurrent combination with other nouns, e.g. “research” (e.g. *“The current study builds on previous research demonstrating...”*), “reports” (e.g. *“Previous reports of ventricular endoscopic neuronavigation have used retractor-based, flexible arm systems or...”*) and “review” (e.g. *“Previous reviews indicate that only 3–5% of CPP present bilaterally...”*). The latter noun also appears together with the adjective “retrospective” (e.g. *“In this retrospective review of our 7-year experience with ... the high ...failure were similar to past decades”*). Also “retrospective studies” occurs in the corpus (e.g. *“But recent longitudinal research [64, 65] and*

retrospective studies suggest that this belief may be overly optimistic.”).

Another possible adjective combining with “studies” is “prior”, which is actually a synonym. The combination “prior studies” is also often forming part of a sentence in the present perfect and it co-occurs with the verb “to demonstrate” (e.g. *“Prior studies have demonstrated the utility of...”*). Occasionally, “prior” also combines with other nouns such as “research” (e.g. *“Much of prior research into the effect of interventions on immunity as well as the current study operated from the assumption that most breast cancer patients...”*). Similarly, in Spanish another possible adjective acting in a similar way to “previous” is “precedente”, and it collocates with “study” (e.g. *“En la mayoría de los estudios precedentes se había utilizado como grupo control...”*).

**g) The adjectives “anterior”, “reciente” and “previo”-
“early”**

Unlike in English, where “prior” also combines with “research”, no instances were registered of “precedente” with “investigación”. In fact, this adjective does not appear in the Spanish corpus and the corresponding adjective “preceding” in English is almost non-existent in the medical corpus, since only one example appears. Thus, the noun “investigation” in Spanish collocates with “anterior”, which at the same time also collocates with “estudio”. The collocation “investigaciones

anteriores” tends to appear in the plural (e.g. “*Los sujetos con adicción alta aparecen... en investigaciones anteriores*”).

The noun “investigación/es” collocates with other adjectives such as “preexistente” (e.g. “*Los resultados obtenidos en el estudio... son coherentes con la investigación preexistente en el campo de...*”) and “recientes” (e.g. “*Nuestros resultados son coincidentes con recientes investigaciones...*”).

In Spanish, apart from the noun “estudio”, another noun that collocates with “previo” in the corpus is “trabajo” (e.g. “*De manera similar a trabajos previos, este estudio adolece de algunas limitaciones metodológicas, que pueden...*”) and also “literatura” (e.g. “*Existe muy poca literatura previa al respecto, sin embargo los escasos estudios realizados han hallado que las mujeres con trastorno psiquiátrico...*”).

Regarding the collocation “anteriores estudios”, the adjective “anteriores” appears either preceding or following the noun, thus “estudios anteriores” or “anteriores estudios”. Another relevant characteristic in relation with the adjective “anteriores” is that the noun phrase it forms with the noun “autores” collocates with the verb “coincidir” (e. g. “*Contrariamente, no coincidimos con los anteriores autores, y por lo tanto confirmamos...*”) and additionally the verb “coincidir” combines with “estudios anteriores” as well (e.g. “*Es lo mismo que concluyen Jadack, Hyde y Keller [1995] ..., coincidiendo con estudios anteriores, ...*”). It is also common to find the presence of a participle with the combination “estudios anteriores” giving as a result the structures: “demostrado en estudios anteriores”, “evidenciadas en anteriores estudios” or

“observada en anteriores estudios” (e.g. “*Ésta es una situación ya observada en anteriores estudios 13, 14 y es un dato...*”). Finally, another typical combination in the corpus is that of the adjective “anteriores” with the noun “revisiones” (e.g. “*Nuestros resultados coincidieron con revisiones anteriores que señalan el entrenamiento...*”).

A possible equivalent to the Spanish adjective “anterior” with an intrinsic meaning referring to the past in English is “early”. What makes this adjective accompanying “studies” different is the fact that it frequently appears in the comparative form “earlier” and that a past participle is used (e.g. “*In earlier studies, the methods used to quantify outcome measures...*”). Sometimes, the noun “studies” is not present, but there is still a reference to previous studies that take this time the form of “clinical observations”, for instance (e.g. “*This is consonant with early clinical observations made by Meyer [31] showing that ERP could modify...*”).

h) The adjective “escasos”, “little” and “few” and the nouns “evidence” and “evidencia”

Finally, the noun “estudios” collocates with the adjective “escasos” in Spanish. This adjective does no longer belong to the group of adjectives referring to the past. The inherent meaning of “escasos” has to do with the idea of “little”, and it is used in combination with “estudios” by Spanish researchers to indicate that little research has been done on a certain topic. An

example of this collocation is “*Son escasos los estudios en los que se determina la exactitud diagnóstica...*”.

The literal English equivalent adjective “scarce” does not collocate with “studies”; thus, other options are needed for the adjective “escasos” in Spanish. A possible alternative is the adjective “little” in combination with the noun “evidence” in the structure “there is little evidence”, which clearly implies the idea of little investigation in the context of a research paper (e.g. “*There is little evidence that these types of clinical services are equally useful to all patients*”).

Another alternative is the use of the adjective “few” preceding the noun “studies” (e.g. “*In summary it can be said that there are currently few systematic studies concerning...*”). The adjective “scarce” collocates, however, with the noun “literature” which will be analyzed in the next section.

On some occasions, there is not just little research, but rather no research. The lack of studies can also be pointed out in this case negating the verb “haber” in the structure “no hay estudios” (e.g. “*Aunque no hay estudios que comparen la cirugía con las otras opciones, la mejor evidencia disponible en la actualidad indica que...*”).

The last exemplifying sentence above shows another way to express the idea that there is little research; the noun “evidencia” collocates with “escasas” in the structure “escasas evidencias existentes” (e.g. “*...como consecuencia de las escasas evidencias existentes hasta el momento sobre...*”).

The close relation between the word field of “disponer” and the noun “evidencia” mentioned above is used in the

structure “no disponemos de evidencia” to convey the lack of previous research. This is reached thanks to the adverb “no” affecting the verb “disponer” (e.g. “*Aunque no disponemos de evidencia científica que demuestre la eficacia de estos...*”).

The noun “evidencia” collocates with the adjective “acumulada” (e.g. “*Aunque el criterio actual de asincronía es..., la evidencia acumulada indica que...*”). In English the same literal equivalent collocation takes place (e.g. “*Accumulated evidence suggests that PD is a viable modality for both the short- and long-term management of patients with refractory congestive heart failure*”).

i) The adjective “recent”- “reciente”

However, the English noun “evidence” does not only collocate with the adjective “little” and with the participle “accumulated” as shown above, but also with the adjective “recent” (e.g. “*Recent evidence from the COMPANION trial suggests that it may also improve mortality*”). This adjective occurs in the move *Background* in both languages. As expected, in medical RA, “reciente” collocates with the noun “estudios”, especially accompanied by the participle “realizado/s” and the preposition “con”, resulting in the structure: “estudio reciente realizado con” (e.g. “*Sin embargo, en un estudio reciente realizado con estudiantes, vemos...*”). In English, the collocation “recent study” also takes place (e.g. “*In addition, recent studies have reported a clinically, non significant increase in the risk...*”). The adverb “recientemente” often co-occurs in Spanish

with the noun “estudio” (e.g. “*Muy recientemente, un estudio con pregabalina, un nuevo...*” or “*Además, los resultados obtenidos en nuestra serie son equiparables a los del estudio recientemente publicado por...*”). However, the adjective “recent” also collocates with other nouns in English. For instance “recent” collocates with “research” (e.g. “*But recent longitudinal research and retrospective studies suggest that...*”), with “advances” (e.g. “*In spite of recent surgical and radiologic advances, the indications for surgery remain controversial.*”) and with “trial” (e.g. “*A recent double-blinded randomized controlled trial showed...*”). In Spanish, the adjective “reciente” was also registered with the noun “editorial” (e.g. “*Fenton y Schooler [2000], en una editorial reciente de la Schizophrenia Bulletin sobre estudios basados en la evidencia, también apuntaron ventajas similares para...*”).

It is common to see the conjunction “aunque” placed at the very start of the sentence when informing about the lack or little existing investigation. This conjunction also appears in combination with the adverb “actualmente” and the verbal form “sigue siendo”, in this way, the reader is informed about research in the past (e.g. “*Aunque actualmente el gold standard para el tratamiento de la invaginación intestinal infantil de forma conservadora sigue siendo el neumoenema guiado por fluoroscopia [11-13], consideramos que la realización ...*”). The English adverb “currently” also occurs to indicate a contrast with the past; in the following example the contrast is underlined by the comparative English structure “more...than”:
“*Patients with prior... require a more aggressive approach than*

currently employed.” The combination of the adverb “currently” with the adjective “available” is very commonly used to indicate what research has attained up to the moment of the paper and especially to mention what the limitations of research are. With this purpose a connector expressing contrast such as “however” or “nevertheless” is placed immediately after the collocation “currently available” (e.g. “*Various effective surgical procedures are currently available, however ...*” or “*The tools for this type of genetic analysis are currently available*”,... *Nevertheless, ...*”). In this use, different nouns appear as the subject of the structure “currently available” such as: “procedures”, “tools”, “devices”, etc.

3. Literature- Literatura

The terms “literature” and “literatura” are synonyms of “studies” and “estudios” in medical corpora. It is very common to refer to the literature in the move *Background*. With this purpose, the noun “literature” and “literatura” are mainly used in two different ways. On the one hand, they are just used as a means to inform what has been done up to the moment of the research paper. On the other hand researchers compare their findings to previous results obtained in the literature, that is, in the past. The first use of the noun “literature”/ “literatura” is here called a “neutral or informing literature” and the second is referred to as “comparing literature”.

a) Neutral literature

The most common way to inform on what the state of the art is like in English is by means of verbs such as “to suggest” and “to show” (e.g. “*Scientific literature suggests that there are gender differences in the experiences of infertility*” or “*The literature has shown that many of these injuries are unpreventable...*”). As both examples in brackets show the present simple and the past perfect are the most common tenses in combination with “literature”. In Spanish, the literal equivalent verb “sugerir” occurs in the corpus (e.g. “*En general, la literatura sugiere que deberían ofrecerse tratamientos unimodales antes de los tratamientos combinados, como consecuencia de las escasas...*”).

A term that combines with the Spanish “literatura” is “revisar” either used as a participle or as a verb in the present tense (e.g. “*...El maltrato es una variable influyente en la literatura revisada*” and “*De hecho, si revisamos la literatura sobre tratamientos psicológicos...*”).

It is common on the part of researchers to mention that there is not sufficient research as a means to justify their investigation. The words “hay alguna literatura” and “pero no concluyente” underline this idea in the example “*... hay alguna literatura revisada pero no concluyente*”. Another example showing this reference to scarce literature is “*Existe muy poca literatura previa al respecto, sin embargo los escasos estudios realizados han hallado...*”. In English, the adjective “scarce” collocates with “literature” with the same aim (e.g. “*This series*

is also meant as an addition to the scarce literature on paediatric...”).

b) Comparing literature

In order to compare with previous literature researchers use the typical structures already mentioned in this section. They mainly use among all the possibilities “ser consistente con”, “coincidir con” and “concordar con” (e.g. “*Los resultados no son consistentes con lo que esperábamos ni con la literatura [Patience, 1995] que afirma que...*” or “*Nuestra muestra también concuerda con los datos de la literatura consultada que es indicativa de...*”). Additionally, other words such as “confirmada” serve to express similitude between their own findings and that in the literature (e.g. “*Esta relación está bien confirmada por la literatura, ya que se han encontrado resultados similares...*”).

In English it is less common to find instances of the word “literature” establishing a comparison with previous studies in a strict sense, although some instances were registered. Comparisons are mainly done using comparative forms of the adjectives (e.g. “*Developmental cystic dysraphic lesions unassociated with the stigmata of spinal ... have a significantly higher incidence [28%] than reported in literature*”).

4. Have demonstrated/shown -Han demostrado/señalan/afirman

In Spanish, when talking about other author's research, there is a tendency to use either the present tense or the present perfect, although occasionally the past is also used. In English, it is the present perfect tense that is most widely used.

There are a group of verbs in the present perfect that are recurrently used. The most typical in Spanish are: "demostrar", "afirmar", "manifestar" and "señalar". Many times it is the author that acts as the subject of the sentence (e.g. "*Riley and Riley [1986], han demostrado que éstas, incluso a dosis muy bajas producen una marcada inhibición de la respuesta sexual...*"), but many others the subject of the sentence is not a person but an object, for instance, "estudios" (e.g. "*Numerosos estudios citados en la introducción han demostrado que existe relación entre la esquizofrenia y el consumo de drogas*"). The same structure occurs in the English corpus (e.g. "*Kaufman et al [21] have demonstrated that disbandment of...*").

A common combination in Spanish is the conjunction "como" together with the adverb "ya" forming "como ya" followed by a past tense (e.g. "*Como ya señalaba el estudio de...*" or "*...tal como ya había afirmado López et al. [1990]*").

Regarding English, the most common verb is "to demonstrate" (e.g. "*Some authors have demonstrated that systemic morphine can suppress noxious...*"). As the exemplifying sentence corroborates, the verb "to demonstrate" in the present perfect tense is used in active and in the passive

constructions both in English and in Spanish, despite the passive being more characteristic of the English language.

3.3.2.2. *Analysis of the move Summarizing*

The aim in this move is, as its name indicates, to summarize the findings in the paper. When comparing the two languages, the word fields common to both languages were given a number from 1 to 7 as shown in Table 3.52 below:

1	Study- Estudio
2	Results- Resultados
3	Data- Datos
4	Findings- Hallazgos
5	Paper- Trabajo
6	Series- Serie
7	Other Expression- Otras Expresiones

Table 3.52: Common word fields in the move Summarizing in English and Spanish

1. Study- Estudio

In the case of *Summarizing*, the function of the move is to summarize what has been done or achieved in a paper. This explains that the first section found in this move included all the expressions in English and Spanish containing the word “study” or “estudio”. This word represents in English the 0.59 % of the corpus and “estudio”, the 0.41% in Spanish. The importance of these words is crucial since they are the second most widely used nouns in the two corpora. The 4 main combinations with

this word in terms of frequency in English and Spanish are the following:

A	El estudio	59	This study	79
	En este estudio	20	In this study	23
	Este estudio	49		
B	En nuestro estudio	23	Our study	39
			In our study	6
C	El presente estudio	12	In the present study	4
			The present study	20

Table 3.53: Main structures with the nouns “study” and “estudio” in the move Summarizing

The start of this move was equivalent in most cases; the most common way to introduce the move in English is “this study”, which occurs 79 times. The same structure is found as the second most common combination in Spanish “este estudio”, which comes up 49 times. However, the most common structure in Spanish is “El estudio” (59 times), that is equivalent to “este estudio”. The English combination “in this study” (23) is equivalent to “en este estudio” in Spanish. Despite being very common it is more frequent to use a possessive pronoun in both corpora: “en nuestro estudio” in Spanish and “our study” (39) or “in our study” (6) in English.

It is also common in English to refer to “the present study” (20). This expression has a literal correspondence in Spanish “El presente estudio” (12), but it is less frequent than the English equivalent.

Regarding the verb following these 4 combinations, it is noteworthy that in Spanish there are two tenses used, the present simple tense and the present perfect, whereas in English besides these two tenses the simple Past tense also occurs. This tense hardly appears in Spanish. At a lexical level, many verbs were used in both languages but others were not. Thus, the 3 lists below show those verbs present in both languages and those present only in one of them. The information in brackets indicates the times that the verb appeared in the corpus and the percentage that this implies.

Verbs present in English and Spanish:

- a) *Mostrar* (0.11%) - show (0.20 %)
- b) *Encontrar* (0.07%) - find (0.11%)
- c) *Confirmar* (0.05%) - confirm (0.05%)

The verb “to show” is used much more in English than in Spanish collocating with “study”. It comes mainly in the present tense, but also in the past and present perfect. A total of 14 examples were collected, whereas in Spanish only one example was registered in which the verb was in the present tense. The verb “found” was mainly present in the structure “In this study we found...” from which two instances were collected. E.g: “*In this study, we found a high prevalence of LV diastolic dysfunction with...*” or “*Our study found*”, only in one occasion. As for Spanish, the verb was used only once in the present perfect tense: “*En este estudio hemos encontrado*”. Thus this

verb is used in the past in both languages and not very often in combination with “study”. Regarding “confirm”, it was more often found in Spanish, i.e. 3 times, “*El estudio confirma*” than in English, where only one instance was registered “*The present study confirms*”.

Verbs only present in English:

a	suggest
b	demonstrate
c	provide
d	identify
e	support
f	address
g	be based on
h	to be performed
i	substantiate
l	offer
m	reinforce
n	highlight
o	indicate

Table 3.54 Verbs only present in English in the move Summarizing

The 5 verbs on top of this list have a relevant rate of appearance. The most recurrent one is “suggest”, which does not appear in Spanish combined with the word “study”, but it does with other nouns such as “resultados”, “hallazgos”, etc: “*los resultados de este estudio sugieren...*”. Its occurrence is 118 times in the corpus which implies 0.20% of the total corpus and it is mainly found in the present tense when combined with the above structures. It is relevant that in most cases it is followed by “that”. Ex: “*This study suggests that...*” The second verb on this list, “demonstrate”, appears 59 times, that is 0.10% of the

corpus and it is also mainly found in the present tense. All of them are mainly used in this tense which seems to be the most common way to present a result of a study. “Provide” with 54 times and a 0.09% and “identify” with 45 times and 0.08% are the next in terms of recurrence; however, they only collocated once and twice respectively with “study”. Finally, the verb “support” comes up 42 times, which represents 0.7% of the corpus. “Support” is also found as a noun 30 times. The rate of appearance of the rest of the verbs in this list was not relevant, however, it is worth mentioning that the verb “to perform” is not present preceded by the 4 main structures mentioned above, but it is, however, found very often as a past participle (e.g. “*This study has been performed to answer the question of whether there is any difference in the initial phase of adolescent psychiatric problems covering ...*”).

The verb “to support” is present in the past tense (e.g. “*Our study supported the cutoff score of 44, suggested by Blanchard et al. [7], which led to greater diagnostic efficacy.*”) It is noteworthy that other verbs such as “give” or “lend” are often combined with “support” turning it into a noun. Ex: “*The present study does not lend support to the hypothesis...*”

It is probably surprising to find that the verb “to be based on” does not appear in the *Move Summarizing* in Spanish. Of course, many instances of this verb were registered in the corpus, but they did not belong to the mentioned *Move*; in fact, most of the times it is used to refer to general studies rather than to refer to the RA in which it appears (e.g. “*De manera global, los estudios basados en cómputos de citas coinciden en*”

hallar que...” or “*en una editorial reciente de la Schizophrenia Bulletin sobre estudios basados en la evidencia*”). In the latter example, the verb “basarse en” belongs to the move *Background*.

It is therefore important to make clear that some verbs have a significant rate of appearance in the English corpus, such as “focus” or “reveal”, but this does not mean that they come up in the move Summarizing. The verb “focus” occurs mainly in the move “Further Research” (e.g. “*further studies need to focus...*”).

Likewise, it might also call our attention not to find corresponding Spanish verbs to the ones that appear in the English *Move Summarising* such as “indicar”. The latter verb occurs in the corpus but again it is not relevant in the *Move Summarising* in combination with the noun “estudio/s”. Only 3 instances were registered with the noun “estudios” in which “indicar” belongs to the *Move Background* (e.g. “*Aunque los resultados de la mayoría de los estudios publicados hasta la fecha indican que la psicoeducación...*” or “*De manera similar, estudios llevados a cabo por Bukovic et al. (2000) y por Ip et al (2001), indican que los estudiantes de la Escuela Superior de Medicina...*”).

Verbs only present in Spanish:

a	Presentar
b	Observar
c	Utilizar
d	Estimar
e	Ver
f	Comprobar
g	Hallar
h	Poner en evidencia
i	Aportar

Table 3.55 Verbs only present in Spanish in the move Summarizing

The 3 main verbs at the top represent those which combine with the noun “study”. In our Spanish corpus, there is a clear tendency to use the verb “presentar”, which appears 147 times (0.14 %), and in the present tense.

The English verb “present” appears with the noun “paper” rather than with the noun “study” (e.g. “*This paper presents the management of 23 children with congenital AAD with severe neurological deficits and the several problems...*”), in fact, when it co-occurs with “study” it does not function as a verb but as an adjective (e.g. “*the present study...*”).

The next verb in terms of recurrence that combines with the 4 structures above is “observar”. It appears “48 times (0.05%). The verb “utilizar” has a higher percentage (see appendixes) than “observar”, it appears 66 times (0.06%). However, its appearance in combination with “study” is not relevant; in fact it only comes up once and in the past tense: “*En el presente estudio se utilizaron...*”

Similarly, despite the verb “observar” having a high rate of appearance it only appears once as a collocation of “study”: “*En nuestro estudio se observa...*” This is also the case of “presentar”, which only shows one case where it is clearly combined with “study”. However, paying attention to those verbs used only in one language, it is noticeable that Spanish uses a verb related to the senses, “observar” when dealing with studies, (what is more, a synonym of this verb is also present in the list, namely, “ver”), whereas, this is not the case of English. Although it is difficult to say that a certain verb collocates with these 4 structures, it is the whole list that combines. Nevertheless, it can be concluded that verbs using a “se”, that is, verbs with an impersonal form, are frequently found after these structures: “se observa”, “se ha estimado”, “se confirman”, “se ha comprobado”, “se postulaba”, etc.

It can be concluded that the 3 most common verbs used with the noun “study” in the 4 structures presented above are “show” (14), “demonstrate” (14) and “suggest” (8). The first one is mainly used in the present tense, but also in the past, the second one is also used in the present, but also instances where the subject “we” is used and passive structures “it was demonstrated” are present. As for “suggest”, also the present tense is the most important one. All the 3 verbs are usually followed by “that”, and sometimes preceded by adverbs that give emphasis such as: “clearly shows”, “strongly supports”, etc. Thus, the most recurrent structures in combination with the noun “study” in English would be:

- *This study /demonstrates / demonstrated that*
- *The present study demonstrated /has demonstrated that*
- *Our study has demonstrated that*
- *In this study it was demonstrated/ we have demonstrated that*
- *This study shows that/ showed that/ has shown*
- *The present study shows that/ has shown that*
- *Our study showed/ shows*
- *This study suggests that*
- *The present study suggests that/ has suggested that*

As for the Spanish language, there are not such clear combinations in this language. There is not a certain verb which is used more frequently than the others with the noun “study”. The most frequent is “confirmar”, but it is only used 3 times. Many different verbs are used with this noun: “encontrar”, “estimar”, “comprobar”, “hallar”, “observar”, “revelar”, “presentar”. Therefore, Spanish writers should try to limit all these variants to the most common ones in English shown above to make sure their writing sounds natural when using the noun “study”.

a) Conclusion- Objective- Conclusión/ Objetivo

Another typical combination in Spanish with the noun “estudio” is the one preceded by the preposition “de” followed by either the demonstrative “este” or the possessive “nuestro”, thus giving the combination “de este estudio” and “de nuestro

estudio”. These 3 words frequently act as a complement of two nouns “conclusion” and “objetivo”. Some examples of these are:

- *El principal objetivo de este estudio,*
- *El objetivo de nuestro estudio,...*
- *Un segundo objetivo de nuestro estudio...*
- *Algunas conclusiones de este estudio,*
- *La conclusión más clara de nuestro estudio,...*

Also in English there is a tendency of the word “study” to cooccur with “conclusion” Ex:

- *The major conclusion from this study is that...*
- *From our study, the following conclusions can be drawn*

“Objective”, by contrast, does not seem to collocate with “study”. The noun “paper” is a similar term to “study”, although unlike “study” in a “paper” results are necessarily written and published. Unlike “study”, “paper” does collocate with “objective”: “*The main objective of this paper...*”

The noun “goal”, a synonym of the noun “objective”, appears in the corpus with “study”: “*The main goal of our study was to examine...*”. Another synonym found is the word “aim”, which also collocates with “study”. However, it does not collocate when it functions as a noun but when it functions as a verb: “*Empirical studies aim to capture...*”, however, in this

case it is not used to express the summary of the study but to convey the purpose.

2. Result- Resultado

Before proceeding to the linguistic analysis of this noun it is necessary to mention its relevance in the corpus. Its representation is 0.25% in the English corpus and 0.29% in the Spanish. Therefore, these nouns occupy the fifth and fourth place on the list of the most common nouns, respectively.

“Resultado” is the most recurrent noun with the noun “study” coming up 0.01% times in the form “Los resultados de nuestro/ este estudio”. With this structure, the verb “sugerir”, also takes place (e.g. “*Los resultados de este estudio sugieren...*”)

As for English, the structures “the results of this study” and “the results of our study” appear 14 and 6 times respectively, which makes a total of 0.04%. This structure is followed by the verb “to suggest” on 10 occasions in the corpus (e.g. “*The results of this/our study suggest...*”)

The noun “resultados” collocates with the noun “trabajo” in the form: “Los resultados de este trabajo”. Naturally, the verbs that follow these two structures in Spanish coincide with many of the verbs that combined with “study” above (mostrar, confirmar), but other verbs also present are: sugerir, indicar, comparar, permitir (concluir), poner de manifiesto, ir a favor de, etc.

In English, the structure “The results of this/ our study” collocates with the verbs: “suggest”, “demonstrate” and “show”. However, other new verbs that appear with this structure are: “lead us to” (make several recommendations), “point” (to the importance of), “are in keeping with” (previous data), etc.

Also equivalent literal structures are present in both languages with the noun “results”:

- a) *Nuestros resultados*
- a) *Our results*
- b) *Los resultados /Estos resultados*
- b) *The results/these results*

Both are followed by the same verbs mentioned up to now. The new appearances with respect to the verbs seen up to now are the verbal form “can serve as”, (e.g. “*These results can serve as a hint pointing...*”) and the verb “to evidence” (e.g. “*Our results evidence that patients with...*”) in English. In Spanish there are also combinations of two verbs or expressions such as “permitir confirmar” or “dar pistas” (e.g. “*Estos resultados permiten confirmar...*”).

An interesting verb that appears with “*estos resultados*” is “demostrar”. The verb “to show” is one of the most commonly used with “study” in English. In Spanish, however the noun “estudio” and the verb “demostrar” do not collocate. Nevertheless “*resultados*” do collocate with “demostrar” (e.g. “*Estos resultados demuestran que la técnica B-O consigue...*”). An equivalent literal structure, in which the nouns “*resultados*”

and “results” are used, is “to be in line with” in English and “estar en consonancia con” in Spanish (e.g. “*Nuestros resultados están en concordancia con...*”/ “*Our results are in line with*”).

In Spanish, “los resultados” appears very often followed by a participle form such as: “obtenidos”, “alcanzados”, and “encontrados”. Although participles are also present in the English language in the same structure, they are much less common than in Spanish. E.g. “resultados obtenidos” comes up 37 times in Spanish whereas in English only two instances were found of “obtained results”.

The noun “results” is sometimes followed by the verb “to be” plus an adjective, which describes these results. The most common adjectives describing them in both languages state a clear comparison, e.g. the Spanish adjectives “coincidentes” “consistentes” “semejantes”, “coherentes”, “similares”, “mejores” and “consistent”, “similar”, “best”, “improved”, “excellent”, “superior”, “in line with”, in English.

The collocation mentioned in the section above of “study” and “conclusion” both in English and Spanish, is also present in this section with “resultados”. In fact, it comes up with the structure “A la luz de los resultados obtenidos” or “A la vista de los resultados obtenidos”. “A la luz de” and “resultados obtenidos” also collocate. Ex:

A la luz de los resultados obtenidos podemos concluir que, efectivamente...

A la vista de los resultados obtenidos podemos concluir...

Examinados los *resultados obtenidos a la luz de la* revisión efectuada podemos afirmar...

A pesar de las posibles complicaciones, según nuestros *resultados* podemos *concluir* que el empleo sistemático...

The fact that the two words belong together is also evidenced in other similar sentences containing “resultados” and “conclusiones”, but this time “conclusion” in the plural and also joined with a preposition:

Después de evaluar nuestros *resultados* con el uso de... llegamos a las siguientes *conclusiones*...

En conclusión, se muestran en este estudio *resultados* preliminares según los cuales...

Regarding English, also the structure “in the light of” was present but not in combination with “results” but rather with “experience” and “data” (e.g. “*In the light of this descriptive present data...*”). As for the expression “En conclusion”, it appears literally in English as well “in conclusion”. What calls our attention is the fact that “the/these results” and “our results” are more commonly preceded by an expression such as “in conclusion” or by linkwords in English than in Spanish. Examples of these linkwords are: “Furthermore”, “nevertheless”, “therefore”, “overall”, “however”, etc.

3. Data- Datos

The first information to be considered regarding this noun is that it is used more in Spanish than in English. “Dato” has a percentage of 0.14% in the corpus and it is the number 19 in the list in terms of frequency. “Data” has a representation of 0.11% and it occupies place 25 in the English list.

The first thing that calls our attention is that “dato” is usually followed by a complement containing a noun “los datos de la muestra estudiada”, “los datos del estudio”, “los datos de nuestro estudio”, “los datos de nuestro trabajo”, etc. Another aspect typical for “dato” is that, as with other nouns already analyzed, it is frequently followed by a participle: “los datos obtenidos”, “los datos recogidos”, “datos expuestos”. In English only one instance of “data obtained” was registered. As for “data” followed by the preposition “of”, only two examples are contained in the corpus. And 4 instances of “data followed by the preposition “from” occur in the corpus (e.g. *“Taking into account that previously presented data from the ESBY study have found SCS and CABG to be comparable in terms of symptom relief...”*).

In English, “our data” and “these data” are immediately followed by a verb, which is normally one of the ones that collocate with study: “confirm”, “suggest”, “demonstrate”, “indicate”, “identify”, “support”, etc. New verbs that were not present with the noun “study” are: “describe”, “to be in agreement with”, which is very similar with “to be in line with”, “contribute” (to the understanding of), “to lead (us to suggest)”.

Also in Spanish the combinations “Los datos”, “Estos datos” and “nuestros datos” are followed by the same verbs that collocate with “estudio”: “indicar”, “surgir”, “mostrar”, “confirmar”, etc. Other verbs that did not appear with “estudio” are: “señalar”, “revelar”, “apuntar”, “apoyar” and combinations of two verbs like: “parecer indicar”, “permitir hallar”, “venir a validar”. Paying attention to the tense, it is worth mentioning that in English the present simple is mostly used in this case, though the past occurs occasionally. Spanish also uses the present tense mainly; however the conditional in “vendría a validar” occurs as well.

The above expressions “parecer indicar”, “permitir hallar” are clear examples of the use of “hedgings” in medical texts.

A common preposition that collocates with “datos” is “según”, followed by the possessive “nuestros” or the demonstrative “estos”: “Según nuestros/estos datos...” equivalent to “De acuerdo con los datos obtenidos”. In a sense, it could also be considered a synonym of “A la vista de los datos expuestos...” which collocated with “resultados” in the section above. Similar structures in English are the expressions “according to the data obtained” and “in the light of this descriptive present data”. However, “según” has a representation of 0.06% of the corpus, twice as much as “according to “. The verb “based on” with the same meaning was only registered 5 times (0.01) but it only appears once in combination with the noun “data” (e.g. “*Based on the observed data, there is no safe way to use a trampoline*”).

“A la vista de” has a representation of 0.1%. This indicates that it is more common in Spanish to use this expression than “según”, which is much shorter. Also the frequent complementing nouns “los datos de nuestro estudio” as well as the participles accompanying the nouns, for instance “datos obtenidos”, which is not the most common thing in English, accounts for this fact. Other examples would be periphrasis such as: “venir a validar” or “poner de manifiesto”. All the examples in this paragraph account for Spanish sentences being necessarily longer than English.

4. Findings- Hallazgos

There is a significant discrepancy in terms of incidence of nouns. In English, the noun is much more recurrent, with a representation of 0.18% and in Spanish it only stands for a 0.02% of the corpus. In the case of “hallazgos” it is observed that it has a similar behavior to that of “datos”. This noun is also present in combination with a noun complement “los hallazgos de este estudio sugieren que...” and it also appears accompanied by a participle “Los hallazgos encontrados”. This again gives rise to long sentences as stated above. The presence of the present perfect contributes to the resulting long sentences in Spanish “nuestros hallazgos han confirmado que la utilización...” The verbs after this structure again coincide with the verbs mentioned in the other sections: “confirmar”, “sugerir”, “señalar”, etc.

In English, the same verbs are also used, but some recurrent new verbs or combination of words appear as well : “be surprising”, “highlight”, “should enhance”, “seem to confirm/to indicate”, “have (some important clinical implications)”, and “can be summarized (as follows)”. Unlike “data”, “findings” also appears in longer structures as it normally happens in Spanish: “The findings in this study” and “The findings of this study”.

5. Paper- Trabajo

Although “paper” corresponds to the Spanish word “artículo”, it has been decided to analyze the appearance of the word “trabajo” instead for reasons of recurrency. Although the noun “artículo” comes up in the Spanish corpus, it only appears on 6 occasions (e.g. “*Al plantear las conclusiones sobre un tema tan controvertido como es la actitud terapéutica ante el RM, hay que tener en cuenta que los artículos revisados generan poca evidencia científica ya que...*”).

The word “trabajo” is much frequently used in the corpus. In fact, there is also a difference in terms of frequency between “trabajo” and “paper”. The former noun has an occurrence of 0.16% while “paper” only has 0.02%. With this statistical information, Spanish writers should bear in mind the low frequency rate of this noun and try to use other more commonly used in English instead such as “study”, or at least they should not use it too much. Still the context for these terms is very similar when they co-occur with the noun “objective”.

“El objetivo de este trabajo ha sido estudiar...” and “The main objective of this paper has been to highlight...”. The terms are also employed with an equivalent meaning when the authors want to talk about summarizing the ideas in their paper: “Este trabajo recoge de una manera sintética...” and “This paper attempts to provide an overview of the ...” or “The focus of this paper has been...”. Typical verbs used with this noun are: “present”, “address”, “attempts” (e.g. “*This paper addresses some of the issues concerning the most common location of...*”). As shown in the example, “paper” appears mainly preceded by the demonstrative “this”, although the structure containing a possessive pronoun also takes place: “En nuestro trabajo”.

6. Serie- Series

In Spanish, this noun, which refers to a group of people, especially of patients, is used in this move preceded by the possessive “nuestra” or the article “una” only in the singular form. When referring to a group of patients, it comes up on its own, however it is also found followed by the preposition “de” and another “noun”, but with this form it no longer refers to “patients”. It is then used to design a group of things such as in “una serie de pasos”, “una serie de problemas”. The plural form, however, is almost always used with the former meaning but it is not present in this move. The percentage of “series” in Spanish with meaning is 0.02% mainly introducing the move with the form “En nuestra serie...” English structures are “in our series”, “in this series”. This word is also found with a slightly

different meaning as it was the case in Spanish: “a series of children”, but very few instances with this second meaning are found. The occurrence of the word referring to “group of patients” is higher than in Spanish, i.e. 0.05%. In English the noun also appears in combination with the possessive pronoun “our” in the move *Summarizing*: “In our series of patients...” and “In our series”.

Although less frequent, the equivalent expression “number of” is used in the *Move Background* (e.g. “*Out of a total number of 121 patients [including the authors’ series] with ISCCA, only 9 were children...*”).

7. Other expressions

A typical noun in Spanish present in this move is “muestra” (e.g. “*En nuestra muestra no es posible...*”, “*En nuestra muestra se describe...*” or “*En la muestra estudiada se observan...*”). It is interesting to underline that this word is used more often in Spanish than in English with a representation of 0.14% and 0.08% respectively. Additionally, the noun “sample” appears more often in the move *Further Research and Limitation* (e.g. “*This study is based on a small sample; therefore, the results need to be replicated with a larger group of patients.*”) rather than in this move, i.e. in the move *Summarizing*.

Occasionally the noun “análisis” was used in Spanish in the *Move Background* (e.g. “*Tras el análisis de nuestro estudio cabe enumerar las siguientes conclusiones...*”).

It is very common to find a concluding expression in front of the sentence introducing this move in both languages. Containing the word “conclusión” and “conclusion” several expressions can be observed in Spanish:

- *En conclusión,*
- *En virtud de todo lo visto la conclusión es clara*
- *Las principales conclusiones de nuestro meta-análisis son...*
- *Para concluir, y teniendo en cuenta los resultados obtenidos...*
- *En síntesis, del conjunto de nuestros datos puede concluirse...*

Some English expressions that belong to the same word family as “conclusion” comparable to the Spanish expressions above are also present in the corpus although they are not so common. Among them the following can be mentioned:

- *In conclusion,*
- *It has been concluded that,*
- *We conclude that*
- *The following conclusions were drawn from the above discussion...*

A less frequent word expressing conclusion but also present in both languages contains the semantic idea of the Move. Thus, expressions such as: “En suma”, “En resumen” or

even “En síntesis”, “Recapitulando” and “En definitiva” “De la revisión realizada...”are found in Spanish and “In sum”, “to summarize” in English. Occasionally, more complex structures containing this word take place: “The evidence presented above is an attempt to summarize...”

It is also common to use the noun “experience” in both in English and Spanish in this move (e.g. “*Our experience in 29 patients suggests or “Tras la presentación de nuestra experiencia con los tumores...”*, “*Nuestra experiencia demuestra...”*). And the same happens with “experiment” and “experimento”, which are used occasionally (e.g. “*El presente experimento pretendía...”* and “*In this experiment of neonatal...”*).

A difference that can be established is the use of the term “investigación” in Spanish: “En esta investigación se han hallado...”

Other differences to be underlined are the use of the passive in English such as in: “It has been concluded that” or “Enhanced and prolonged activation of... was observed”.

In spite of the lack of passive in Spanish, it is remarkable the fact that the presence of the authors of the paper is more often highlighted in English papers by the use of the personal pronoun “we”. Additionally, it is very common to underline the success they obtained with the verb “to be able” and especially with the ordinal “first”. Some examples are: “We were able to demonstrate...”, “We are the first to report extensively on...”, “We present the first description”, “We could clearly show” or “We found”. This is not so common in Spanish where only a

few instances were found using the verb “decir”, which sounds a bit humbler “Podemos decir que se confirma...” or “Podemos decir que a partir de esta investigación...” Interestingly, all the instances of “hemos encontrado” in Spanish were anticipated by the negative adverb “no”, thus implying a limitation rather than a success. Thus it can be concluded that in general terms the presence of the paper’s author in this move is more determining and positive in English than in Spanish.

**3.3.2.3 Analysis of the move Evaluation of the Study:
Limitations and Advantages**

This move deals with sentences that transmit the idea that there were limitations to the study or problems in the performance of the study. It also deals with any kind of problem or limitation expressed within the conclusion either in relation to the study itself or to any other study and thus this move includes examples expressing limitations in a more general sense.

In this move, 7 semantic fields or equivalent expressions/terms according to the meaning of the words were found to be used in both languages with an equivalent semantic function.

1	Time Adverbs- Adverbios de tiempo
2	Sample Size- Tamaño de la muestra (2.1 Small number- grupo/serie pequeño/reducido)
3	While and Although- Aunque
4	Limitations- Limitaciones
5	Problems/Complications/Difficult- Complicaciones/Problemas/Difícil
6	Cannot/ Not possible/Not known- No es posible/No podemos/ No sabemos
7	Incomplete- Incompleto

Table 3.56 Common word fields in the move Limitations in English and Spanish

1. Time Adverb.- Adverbios de tiempo Still-Aún/Todavía

The most commonly used time adverb in this move in English is “still” (0.05%), which has two possible literal equivalents in Spanish. Summing up the percentages of these two Spanish adverbs, they also reach the same percentage as “still”, i.e. “aún” with a representation of 0.03% in the corpus and “todavía” with 0.02%.

The first thing that calls our attention is the presence of the adverb “no” in the Spanish corpus. This adverb appears after “todavía” (e.g. “*Todavía no se ha clarificado si...*”) and either after or before the adverb “aún” (e.g. “*Al nivel más fundamental no ha sido aún definitivamente respondida la cuestión de si los tratamientos...*” or “*Esto se debe a que aún no existe una definición uniforme del concepto de calidad*”). The combination of this negative adverb with the time adverbs and with certain words with a positive connotation afterwards “clarificado”, “definición uniforme” and “respondida la cuestión” turns the aforementioned “positive” ideas into limitations.

The adverb “no” does not appear in combination with the adverbs of time in English in the move *Limitation*. To convey the same meaning, other words with an intrinsically negative meaning are used, e.g. the word “lack” could be used to express the equivalent Spanish structures “no tenemos” or “no existe”. An example in the corpus is “*We still lack any definitive tool for organizing the complex...*” This structure would be the one used to translate sentences like “*Todavía no tenemos datos estadísticos que apoyen esta afirmación*” or “*Aún no existe una*

definición uniforme del concepto...” Alternatively, an example was found in the corpus with the literal equivalent to the word “lack”, namely “carencia” (e.g. “*Esto contrasta con la carencia de estudios empíricos sobre...*”). Another example of an intrinsic negative word in English is “confusion”. The exemplifying sentence “..., *much confusion still exists with respect to...*” can be considered equivalent to other Spanish ones found in the corpus like “aún no sabemos” or “no está clara aún” (e.g. “*Aún no sabemos qué significan estas irregularidades...*” and “*No está clara aún la eficacia de...*”).

Other two examples of English words with an inherent negative meaning in a sentence containing the adverb still are the adjectives “critical” and “poor” (e.g. “..., *there are still a number of critical areas in need of exploration before such assertion can be fully validated*” and “*However, CPC still has an extremely poor prognosis*”).

Sometimes words accompanying the adverb “still” do not have a negative connotation; however, they are combined with other words that turn them into negative. For instance, the verb “to be” combined with “far from” co-occurring with the verb “to accomplish”, turns the positive intrinsic meaning of the latter verb into the opposite like the following sentence shows: “*We are still very far from accomplishing this.*” The structure “to be far from” is quite recurrent in this move. Sometimes though, it is preceded and followed by words with a negative connotation resulting in a clear limitation (e.g. “..., *but the problem of ... is still far from being disentangled*”). This structure has a literal equivalent in Spanish, “estar lejos de”,

which is followed in the corpus by a word with a positive connotation, namely “alcanzar” conveying thus a limitation. The limitation in this case is reinforced by means of the adjective “imposible” preceding the verb “alcanzar” (e.g. “*Estamos aún lejos de alcanzar el ideal de ofrecer a los pacientes un consenso... En la práctica esto es probablemente imposible de alcanzar*”)

The adjective “imposible” is another example of a word with a negative connotation co-occurring in the same sentence as the adverb “aún”. Similarly, the adjective “insuficiente” co-occurs in the next sentence with the adverb “aún” expressing a limitation: “*Por todo ello, podemos considerar que la evidencia... es aún insuficiente para extraer conclusiones relevantes que ayuden...*”.

Another similar combination of words in the corpus that gives the idea that the goal is still far from being reached is “está aún por” (e.g. “*En cualquier caso, el desarrollo de dichos sistemas está aún por hacer, precisando de más investigación...*”).

In English, neither the verb “to reach” is used in this move or a similar structure to “estar aún por” in a literal way. However, this can be considered equivalent to the recurrent English structure “still remains to be” plus a participle, which has a very similar meaning (e.g. “*One important problem still remains to be solved*” or “*The nature of vulnerability still remains to be elucidated*”). Nevertheless, an example was found in the Spanish corpus that is a bit closer to the semantics of the verb “remain” in English. It is a sentence containing the verb

“quedar” and the adjective “pendientes”: “*Por el momento, a pesar de todas las cuestiones que aún quedan pendientes de estudio...*”

Sometimes, researchers want to transmit to the reader the limitations of the study on the grounds that something has just emerged or been discovered and that this implies a limitation to different aspects. A recurrent structure containing the adverb “todavía” was found in Spanish for this goal: “todavía es temprano para” or “todavía nos encontramos en los albores de” (e.g. “*Todavía es temprano para poder establecer como indicación de la terapia...*”). Alternatively, the structure “todavía está en sus inicios” (e.g. “*La experiencia está en sus inicios, pero...*”). This idea that some kind of technique, method, strategy, therapy, etc., is brand new is also expressed in English using the expression “is still in its infancy” (e.g. “*Operative MRI is still in its infancy and its role in paediatrics is far from being defined*”).

Similarly, authors also express the idea that something needs further research and thus is not very developed transmitting the idea that there is still research to be done. The English expression found in the corpus for this is “there is still a lot to learn about” (e.g. “*There is still a lot to learn about the exact correlation between...*”). A similar meaning but with some nuances is present in the Spanish corpus. This time authors want to express the need to learn something with more detail and depth rather than to express the idea that something is new. For this purpose the expression used is “se ha de refinar más

aún” (e.g. “*Por tanto se ha de refinar aún más el conocimiento que se tiene de...*”)

a) Hasta el momento- Until

The rate frequency of “until” and “hasta el momento” is very low, in fact, “until” has a representation of 0.01% in the corpus and that of “hasta el momento” is even lower. These two adverbs present in this move can be compared, but they are not exactly used in the same way. In fact, “hasta el momento” appears in a context that overlaps with the move *Background* (e.g. “*hasta el momento, la investigación... se limita a los trabajos de...*”), whereas the adverb “until” appears in a context that could also belong to the move *Further Research* (e.g. “*Until prospective randomized studies are conducted... the neurosurgeon must rely on the literature*”). This last exemplifying sentence expresses the need that more research be done.

Although the following Spanish sentence also shows an overlap with the *Move Background*, the presence of the adverb “tampoco” later on, which has a similar meaning to the adverb “no” followed by the verb “explicar”, underlines the idea of limitation (e.g. “*...Pero el rol de este supuesto antígeno es hipotético, dado que no ha sido descubierto hasta el momento, y tampoco se puede explicar...*”). The need for further studies to overcome possible doubts, problems or to confirm certain findings in this last example is comparable to the previous English sentence stated above. In fact, this sentence is preceded

by another that corresponds to the move *Further Research* as it is observed in “*However, considering the relatively small patient number in this study, it is obvious that further studies will be necessary. Therefore, until our study is confirmed by larger studies we would still consider coronary angiography as a diagnostic tool to evaluate...*” This is why these two items present in the move *Limitation*, “*hasta el momento*” and “*until*” can be compared, but they are not really used in the same way. In fact, a possible translation for “*until*” could be “*hasta que*” rather than “*hasta el momento*” and a possible one for “*hasta el momento*” could be “*up to now*”, “*up to the present time*” rather than “*until*”.

Additionally, it is important to clarify that “*hasta el momento*” appears in the move *Background* without necessarily overlapping with the *Move Limitation* (e.g. “*En los estudios publicados hasta el momento sobre la relación entre esquizotipia y hábito de fumar, se han utilizado...*”).

b) Actual - Current

In general, “*actual*” and “*current*” are used to express limitations at the present time that take place because studies have not reached the results expected, failures took place, not enough tools, studies, strategies, etc were available, etc.

Both terms have a relevant high percentage. The adjective “*current*” appears 0.07% and the Spanish word “*actual*” appears 0.06% in its function as an adverb and adjective.

The most typical way to express the limitation using the adjective “current” is by means of words that have an intrinsic negative connotation, as it happened with the adverbs above. Some examples of these negative words are “inadequate”, “unanswered” and “imprecision”, all of them containing a negative prefix, namely “-in” and “-un”: (e.g. “*The current treatments of congestive heart failure are inadequate*” or “*While the current study leaves these questions unanswered, it will hopefully promote interest...*”). This last example again shows a clear link to *Further Research*.

The negative connotation of words accompanying the adjective “actual” also takes place in Spanish. The following example in which the adverb is present, shows the adjective “limitada”, which is intrinsically negative affecting the positive connotation of the noun “eficacia”(e.g. “*En cualquier caso, los diferentes recursos señalados tienen en su estado actual una eficacia limitada...*”). Also adjectives containing a negative prefix, “-in” combine with the adverb “actual” in the *Move Limitation*. An example would be the adjective “insuficientes” (e.g. “*Los recursos existentes en la actualidad en la consulta de Atención Primaria son insuficientes*”). The indication that “there is not enough of something” expressed in the last exemplifying sentence can be also conveyed by means of other words in Spanish. For instance, the adverb “few” plays a similar role to the adjective “insuficientes” as far as meaning is concerned in the following sentence: “*In summary, it can be said that there are currently few systematic studies concerning...*”

Limitation can also be transmitted by other adverbs implying constraint such as “only” (e.g. “*Nevertheless, eliciting a family history of CAD is currently the only available screening tool to identify...*”). The literal equivalent adverb “solo” was also present in the corpus in the same move (e.g. “...; *la tasa de implante actual en nuestro país solo llega al 1% de los posibles candidatos*”).

The presence of the adverb “no” is outstanding in both languages. In Spanish it appears immediately after the adverb “actualmente” (e.g. “*Actualmente no está resuelto si el trastorno...*”) In English, the adverb “no” comes after the time expression “at present”, which is in terms of meaning equivalent to the Spanish adverb “actualmente” (e.g. “... *at present, no significant achievements have been made in the substantiation of the diagnostic validity of mental disorders, and current categorical classifications should not be regarded as valid classifications*”).

Another aspect that can be considered a limitation is the fact that something is at the present time being studied since this implies that not enough information is available for the performance of the study. For this reason, sentences informing on other studies being carried out by the time of the study itself have been recognized as a limitation. Thus, the following two exemplifying sentences express, according to this, a limitation: “*Other devices are currently being investigated for this indication in clinical trials*” and “*Several initiatives are currently underway ...*” Equivalent structures expressing exactly the same idea were present in the Spanish corpus with

combinations such as “continúa siendo actualmente tema de estudio” and “ en la actualidad precisa seguir siendo estudiada” (e.g. “*En este sentido, desgraciadamente, la homofobia todavía es una realidad social en la actualidad que precisa seguir siendo estudiada para determinar*”).

2. Tamaño muestra- Sample size

There are many different combinations to indicate that the number of patients was not representative enough by means of the nouns “sample” and “muestra”, especially in the Spanish language. One of the most recurrent structures is “el tamaño de la muestra”.

a) The adjectives “reducido” and “pequeño” and its opposites “numerosa”, “representativa” and “mayor”

The adjective “reducido” collocates with these two words giving rise to the combination: “reducido tamaño de la muestra” (e.g. “*Debido al reducido tamaño de la muestra no se ha podido establecer...*”). The same lexical term is used with a different structure, namely as a participle, “reducida”. In this new structure it affects “muestra” rather than “tamaño”, but the past participle “reducida” has the same function as an adjective, i.e. it is describing the noun “muestra”: “*...los resultados obtenidos no deben tomarse de forma absoluta ya que la muestra analizada es muy reducida*”. A different combination of the same words also present in the corpus consists of the noun

“muestra” being converted into an adjective, “muestral” (e.g. “*Estos hallazgos, a pesar del reducido tamaño muestral, confirman...*”).

Another adjective that collocates with the noun “muestra” is “pequeño” and in the combination of these two words the small sample size is referred to as “lo pequeño de la muestra” (e.g. “*Dado lo pequeño de la muestra no es posible extraer conclusiones en términos de efectividad*”). It is also common to find the adjective “pequeña” after the noun “muestra” and followed by the complement “de pacientes” as the example shows: “*Algunas limitaciones de este estudio son que se trabajó con una muestra pequeña de pacientes...*” The adjective “pequeña” also collocates with “N” that stands for “número” meaning “number” and refers to the sample size in the structure “la pequeña N” (e.g. “*Las diferentes formas de abuso no pueden ser controladas, dada la pequeña N*”).

Adjectives that can be considered antonyms of the adjectives “pequeño” and “reducido” also collocate with “muestra”. This is the case of adjectives such as “representativa”, “numerosa” and “mayor”. Alternatively, the verb “aumentar” can also act as an antonym despite being a different morphological category. In order to make these antonyms end up expressing a limitation one option is the use of the conjunction “si” expressing the condition that something positive would be possible “si”, meaning “if”, the sample was more representative. In these sentences positive ideas are expressed by expressions such as “es posible que” or by means of the conditional tense: “probablemente alcanzaría” (e.g.

“*Consideramos que el aumento en las expectativas es también importante, y que probablemente alcanzaría significación estadística si dispusiéramos de una muestra de mayor tamaño*” or “*Es posible que las diferencias encontradas pudieran llegar a ser significativas si este estudio se realizara sobre una muestra de mayor tamaño*”). As for the other adjective functioning as an antonym of “pequeño”, i.e. “representativa” and the verb “aumentar”, expressions referring to what should be done such as “sería adecuado” or “deberían” come in front of them turning the meaning of the sentence into a limitation (e.g. “*Sería adecuado disponer de una muestra más representativa de la población general...*”).

b) The adjectives “small” and “moderate” and its opposite “larger”

The English language does not offer such a wide variety of combinations regarding the different ways to refer to the reduced size of the sample. The most common adjective present in the corpus is “small”, which is literally equivalent to “pequeño”. Most of the times it is used as an adjective preceding the noun “size” (e.g. “*The risk factors are not clear and, given the limitations of a small sample size, will likely require a...*”), but other times the adjective “small” is placed after the verb “to be” (e.g. “*First, the sample size is small and there is...*”). An alternative adjective to “small” also found in the corpus is “moderate” (e.g. “*..., which could not be statistically detected due to our moderate sample size*”).

Occasionally, the noun “size” is omitted from the structure “small sample size” (e.g. “*This study is based on a small sample, therefore...*”), and also an example is present in which “sample” is not directly affected by the adjective “small” but appears later on in the sentence. In its place the noun “study” comes up (e.g. “*This small study conducted on a hospital-based sample of...*”).

There is an instance in the corpus of the combination of “sample” with an adjective having a positive intrinsic meaning, namely, “larger”, which could also be considered an antonym of “small” (e.g. “*Further research should confirm these results in a larger sample group...*”).

As the most common word in this move is the noun “limitation”/ “limitación”, it is not surprising that “el tamaño de la muestra” and “sample size” co-occur with the aforementioned noun (e.g. “*Algunas limitaciones del presente estudio serían las siguientes: 1. el tamaño de la muestra...*” and “*The risk factors are not clear and, given the limitations of a small sample size, will likely require a ...*”).

c) The verbs “poder”, “permitir” and “can”

The noun “muestra” also co-occurs with verbs expressing the capacity of doing something such as “permitir” or “poder”. These verbs, however, are preceded by the adverb “no” and express consequently the opposite, namely, incapacity (e.g. “*La principal limitación de nuestro estudio es el tamaño de la muestra ya que no nos permite realizar un análisis en...*”).

Other structures combining with the noun “muestra” containing a verb expressing capacity and preceded by the adverb “no” are the following: “no es posible extraer”, “no pueden ser controladas” and “no se han podido establecer” (e.g. “*Debido al reducido tamaño de la muestra no se ha podido establecer...*”).

In English an equivalent structure expressing the impossibility of performing something on the grounds of the small sample size is present. This English structure contains the modal verb “could” in a negative sentence, which is equivalent to “poder” in the Spanish corpus, and combines it with “due to”. The resulting structure is “... could not be + participle + due to + small/moderate sample size” (e.g. “... *which could not be statistically detected due to our moderate sample size*” or “*The effect of ... on the outcome could not be commented upon in this study due to the small sample size*”).

The verb “allow” functions occasionally like “can”, however it is not accompanied by a negation but it appears combined with the verb “to need” to express a limitation (e.g.: “*On the one hand, this possibly indicates the need for some revisions to allow more precise diagnoses...*”).

d) The verbs “to generalize” and “generalizar”, the adverb “cautiously” and the noun “prudencia”

As a consequence of the limitation caused by the small sample size, it is very common for authors to state the impossibility of taking results as definitive and applicable to the general population. To convey this idea, the English language

frequently resorts to the verb “generalize”, which appears in the following example in the same sentence as the adjective “difficult” underlying the idea of limitation. The inherent negative connotation of “difficult” highlights even more the idea that a limitation is being expressed: “*Because of the sample size, it is difficult to be certain about the extent to which the findings can be generalized to...*” or “*For all these reasons, the prevalence rates... described for the women in our sample may not be generalized to ...*”

The presence of the negative adverb “not” is an essential element in the corpus to express limitation. In the following example the adverb “not” precedes the adjective “clear” and converts it into “unclear”: e.g. “*The risk factors are not clear and, given the limitations of a small sample size, will likely require a ... effort to be thoroughly studied*”.

In Spanish, the word used to express the idea that the limited sample does not allow generalization is “generalizar”. However, although the terms used in the two languages belong to the same word family, in English it is used mainly as a verb and in a passive structure as shown in the examples above, whereas in Spanish it is used in the form of an adjective “general” or of a noun “generalización” (e.g. “*... sería adecuado disponer de una muestra más representativa de la población general...*” and “*Nuestros resultados provienen de una muestra incidental no demasiado numerosa por lo que la generalización a otros colectivos semejantes deben realizarse con prudencia*”). An alternative way to transmit this idea is by means of an expression which is similar in meaning to

“generalize”, namely “no deben tomarse de forma absoluta” (e.g. “*Por ultimo, no cabe duda de que los resultados obtenidos no deben tomarse de forma absoluta, ya que la muestra analizada es muy reducida.*”). The last two exemplifying sentences demonstrate that the plural noun “resultados” is commonly used in combination with “muestra”. In both examples, the authors state that as a consequence of the small samples, results in the study cannot be generalized and should be taken “de forma absoluta” and “con prudencia”. It is observed then that these expressions, and also “con precaución” (e.g. “*Aunque los resultados deben ser tomados con precaución por el tamaño de la muestra...*”) co-occur in the same sentence as “resultados” and “muestra”.

In English, the adverb “cautiously” appears in the same context as in Spanish, that is, it collocates with “results” and “sample” (e.g. “*Our results must be considered cautiously. First, the sample size is small and there is no control group*”). In English, however, the passive voice appears, whereas in Spanish only one out of the 3 instances presents a passive structure, the other two are in the active voice.

e) The noun “significación” and the adjective “significativas”

Finally, the collocation of “muestra” with the noun “significación” and the adjective “significativas” should be mentioned. In Spanish, “muestra” collocates with the pair of words “significación estadística” (e.g. “*Algunas limitaciones:*

1. el tamaño de la muestra... vamos perdiendo significación estadística” or “probablemente alcanzaría significación estadística si dispusiéramos de una muestra de mayor tamaño”). An alternative way to express the same meaning is using “diferencias significativas” in combination with “muestra” as it is shown in the following example: *“Es posible que las diferencias encontradas pudieran llegar a ser significativas si este estudio se realizara sobre una muestra de mayor tamaño”.*

In English an instance was in the corpus that combined “sample” with the adverb “statistically” (e.g. “... *which could not be statistically detected due to our moderate sample size...*”).

f) Small number- Grupo/Serie and pequeño/reducido

When referring to the sample itself, in Spanish the most common combination of words is the adjective “pequeño” which is placed either in front or after the noun “grupo” (e.g. *“En nuestra muestra no es posible ... principalmente por el pequeño tamaño de este grupo”* and “... *“el grupo es demasiado pequeño para concluir...”*).

Although less common, an instance was present in the corpus of “número de sujetos” (e.g. *“El número de sujetos no es muy alto...”*). The group of the sample was also once referred to as “número de casos clínicos” (e.g. “... *La ausencia de sujetos control y el reducido número de casos clínicos presentados..., resulta un problema...*”).

In English, by contrast, the word “group” is not used, the most common words used are the adjective “small”, “number” and the noun “patient” combined differently: “small number of patients” is more common but “small patient number” is also present in the corpus (e.g. “*However, considering the relatively small patient number of this study, ...*” and “*Another limitation in our study is the small number of patients*”). “The number of patients” without the adjective “small” is the most recurrent structure, and it is frequently followed by the participle “studied” (e.g. “*Additionally, the number of patients studied was limited...*”). In this construction the noun “subjects” also takes place replacing “patients”. This can be considered an almost equivalent literal structure to the Spanish “*número de sujetos*” (vid. above) (e.g. “*The small number of subjects studied constitutes a limit in...*”). As the two last examples show, “limit” collocates with “number of subjects” and “number of patients”.

The last collocation with “number of patients” is “series” (e.g. “*As the number of patients in this series is small, further randomized studies are necessary...*”). Occasionally, it co-occurs with the preposition “despite” (e.g. “*Despite the small number of children treated in this series, we suggest...*”). Another repeated structure with “despite” in the corpus is “despite +adjective/participle+ the noun “shortcomings” (e.g. “*Despite acknowledged shortcomings, this study...*”).

In Spanish, the noun “series” does not appear in the same sentences as “*grupo pequeño*”, but its use in the sentence is similar to the noun “*muestra*” in the sense that it collocates with

the adjectives “amplias” and “reducidas” (e.g. “*De una serie tan reducida es difícil obtener conclusiones sólidas...*” or “*Habida cuenta de las limitaciones..., son necesarias series más amplias para asentar con certeza estos hallazgos*”). The combination of “series” and the adjectives above overlaps with the move *Further Research* as the examples show.

3. While-Although- Aunque

These conjunctions are used to express a contrast. For this reason, as it was expected, they join something positive and something negative regarding the subject matter of the paper. This explains that these conjunctions are surrounded by words with a clear intrinsic positive or negative meaning. This is the case in both languages, where many equivalent structures take place. For instance, the conjunction “while” is present in sentences with positive terms such as “successful”, “great help” or “consistent and clear”, “benefit” and “increase”. At the same time these sentences also contain words with an intrinsic negative or restrictive meaning such as: “only”, with a restrictive connotation “did not improve”. These latter words implying a limitation are placed next to the words with an intrinsic positive meaning. In this way, the positive words end up expressing a negative idea (e.g. “to worsen”, “complicated”, “unanswered” or “limitations”).

In the sentences containing “while”, a positive and a negative word are most of the times combined with the conjunction. The part of the sentence containing “while” does

not express the limitation or disadvantage (e.g. “... *systems while successful, are more complicated to use and provide a less rigid...*”). Exactly the same can be applied to the conjunction “although” which means the same as “while” in this move and has a very similar use (e.g. “*Although this model may serve as..., a methodological problem of early diagnosis has to be raised...*”).

Only on two occasions “though” was acting in the same way as “although” in the sentence above (e.g. “*Though the Abio- Cor has yet to receive FDA approval as an alternative to heart transplantation, the promising early results suggest that in the future...*”).

In Spanish, the conjunction “aunque” is used in exactly the same way. Some examples of positive items co-occurring with it in the corpus are: “los resultados pueden apoyar”, “parecen sugerentes”, “eficaz”. Some examples of words conveying a limitation in combination with “aunque” are: “insuficientes”, “no existe”, “no es posible”, “no disponemos”, “and “limitada”. In Spanish it is common to see the word conveying the negative information next to “aunque”. In fact, “aunque” often precedes the adjective “difícil” (e.g. “*Aunque resulta difícil comparar los resultados de este estudio,...*”). An equivalent example to the English instances combining positive and negative terms co-occurring with the conjunction “aunque” would be: “*La metodología utilizada, aunque limitada, se ha mostrado eficaz...*”

4. Limitations- Limitaciones

These two nouns are the most representative with 0.08% and 0.07% in English and Spanish respectively. It is important to clarify that the noun “limitaciones” appears mainly in its plural form rather than in the singular in Spanish.

“Limitations” and “Limitaciones” collocate with “study” and “estudio” respectively. A very recurrent pattern in Spanish is the noun “limitación”/ “limitaciones” with “del estudio” acting as a complement: “Las limitaciones del estudio” or with the demonstrative “este”, “Las limitaciones de este estudio”. This structure appears mainly with the verb “to be” in the present tense: “Las limitaciones del estudio son” (e.g. “*Algunas de las limitaciones de este estudio son...*”). As already noted, the noun “estudio” is often preceded by the adjective “presente”, and this collocation remains when combined with the noun “limitation” (e.g. “*Algunas de las limitaciones del presente estudio*”). As the two last exemplifying sentences show, the “noun” “limitación” is commonly accompanied by the determinant “algunas” for the plural and “una” and “otra” for the singular (e.g. “*Una de las limitaciones de la intervención es el empleo de la respiración controlada*”).

The same equivalent literal structures take place in the English corpus. “Limitation” collocates with “study” giving rise to “Limitation of the study”. As it happened in Spanish, the noun “study” is many times accompanied by the adjective “present”: “*Further research would be required to get clearer conclusions, due to the limitations of the present study*”. The

same equivalent structure is present in the Spanish corpus: Instead of the article “the”, the demonstrative “this” occurs, like the demonstrative “este” in Spanish, and especially, the possessive pronoun “our” (e.g. “*A first limitation of our study was the absence of comparison with...*”). The presence of the pronoun “nuestro” is also frequent in Spanish and some instances were registered in combination with “limitation” in the corpus: “*La principal limitación de nuestro estudio es que...*”

Other words that are synonyms of the noun “limitation” in the corpus are “shortcoming” and “failure”. However, the noun “shortcomings” was registered only twice in the corpus (e.g. “*Despite acknowledged shortcomings, this study introduces the idea of utilizing a solution as an adjunct to the surgical technique.*”). As for the noun “failure” it was not characteristic of the *Move Limitations* (e.g. “*This observational series suggests that patient characteristics may serve the neurosurgeon to predict the risk of shunt failure and...*”). This noun was found mainly forming two collocations: “shunt-failure” and “heart-failure”.

**a) Adjectives that collocate with “limitation”-
“limitación”**

Regarding the most common adjectives that collocate with “limitación”, as the example above shows, “principal” is one of them and also “importante” (e.g. “*Una importante limitación de este estudio...*”). These collocations are also found in the plural (e.g. “*En cuanto a las limitaciones más importantes*”).

de nuestro estudio...” and “*Las principales limitaciones metodológicas del estudio...*”).

In English, however, no instances were found in the corpus of the corresponding literal adjectives in Spanish, namely, “important” and “main”. By contrast, synonyms of “important” collocate with “limitation” like “significant” and “serious” (e.g. “*The failure of PMR to achieve the same clinical results that have been seen with TMR may be due to several significant limitations*” and “*This may be regarded as a serious limitation, reducing this study to a strict ...*”).

According to the important role that methodology plays in R.As, it is logical that a very common collocation in both languages is the adjective “methodological” and “metodológicas”, used by doctors to refer to limitations in methodology (e.g. “*Several methodological limitations of this investigation should be considered...*” and “*La revisión de los trabajos..., ha puesto de manifiesto una serie de limitaciones metodológicas*”).

**b) Nouns that collocate with the noun “limitation”-
“limitación”**

Back to the analysis of the nouns, “resultados” very often appears in the same sentence as “limitaciones” (e.g. “*Sin embargo, a pesar de estas limitaciones, nuestros resultados...*”). An additional word that adds to this collocation is the noun “generalización” giving rise to the structure “limitar + la generalización de los resultados” (e.g. “*..., los sujetos con... no*”).

representa más del 10% de una población no clínica. Esto limita la generalización de los resultados” or “... *datos insuficientes que pueden limitar la generalización y validez de los resultados*”). The equivalent structure using the same lexical terms can be reflected in the example of the English corpus: “*The generalization of these results is limited*”, however, the syntactic structure differs from the Spanish one in the use of the verb “limited” which is active in Spanish but a past participle in English.

The noun “generalización” also deserves some attention. Terms belonging to the word family of “generalización” come up very often in the form of a noun in the corpus when the noun “limitations” appears. Thus it is present as an adjective, “generalizable” (e.g. “*Los datos... no deberán ser generalizables a toda la población; más bien dadas las limitaciones del...*”), as a participle (e.g. “*No obstante, hay limitaciones que impiden su uso generalizado*”), or as a participle in a passive sentence (e.g. “*Several methodological limitations... For all these reasons, our sample may not be generalized to...*”).

The most typical limitation expressed in medical research articles is the small sample size. This limitation is conveyed by means of the noun “generalization” or “generalización” to express the impossibility of extrapolating results, as shown above, but it is also expressed referring to the small sample size. Collocations take place therefore with the word “limitation” and the adjective “small” in combination with “sample” and “number”. Thus, “limitation” appears together with the structures “small sample size”, “small number of

subjects” “small number of patients” (e.g. “*Another limitation of our study is the small number of patients with...*”). In Spanish, “limitation” mainly occurs with the adjective “reducido” (e.g. “*En cuanto a las limitaciones más importantes de nuestro estudio, destaca el reducido tamaño de la muestra*”), but it also collocates with “número” (e.g. “*Los resultados presentados tienen limitaciones dado el número de niños...*”).

The noun “datos” also appears in the same sentence as “generalización” and “limitations” (e.g. “*... datos insuficientes que pueden limitar la generalización y validez de los resultados*”, “*Los datos de generalización del Grupo I son más limitados que los del G2*”). The connection between “datos” and the adjective “limitados” is also present in the sentence: “*Los datos de mantenimiento aportados son un poco limitados por la dificultad de...*”, thus, the collocation “datos limitados” is independent of the noun “generalización”.

The same setting takes place in the English language where “data” co-occurs with “limitation” and “generalization” (e.g. “*The small number of subjects studied and the method applied constitute a limit in the generalization of data*”) but it also appears only with “limitation” (e.g. “*In spite of the above mentioned limitations, our data support that...*”).

c) The preposition “despite”, “in spite of”- “a pesar de” - different

Another typical structure is the combination of the noun “resultados” with the verb “tener” followed by “limitaciones”:

“Los resultados... tienen sus limitaciones”. Sometimes there are not many words between the two parts of the structure (e.g. “*Los resultados presentados tienen sus limitaciones, dado...*”), whereas others, there is a considerable distance but are still linked lexically (e.g. “*Los resultados de esta investigación respecto a la generalización a situaciones y oyentes tienen sus limitaciones puesto que el tratamiento se aplicó...*”).

No similar structure, was found in the English corpus, however, a very similar structure in both languages is the combination of the preposition “despite” and “a pesar de” followed by the nouns “limitations” or “limitaciones” respectively and a verb, many times expressing something positive or introducing something positive (e.g. “*Despite the existence of some methodological limitations, the results of this study suggest that both types of PMT programs are effective...*” and “*Sin embargo, a pesar de estas limitaciones, nuestros resultados confirman la utilidad diagnóstica del NT-proBNP también en una población...*”). It must be pointed out that the structure “A pesar de las limitaciones” is more recurrent in the Spanish corpus, although instances were also found in English, not just with “despite” but with a synonym connector, “in spite of” (e.g. “*In spite of the above limitations, our data support that...*”). However, the presence of the latter connector is not very representative since it only appears four times in the corpus.

Finally, it can be concluded that the noun “resultados” co-occurs more often in Spanish than in English with the noun “limitaciones” and it appears in other structures as well like, for

instance, with the verb “influir” (e.g. ... “*este estudio adolece de algunas limitaciones metodológicas, que pueden influir en los resultados*”).

d) Linkwords with “limitation” and “limitación”

The last aspect to be mentioned is that linkwords many times start a sentence containing the word “limitations” and “limitaciones”. The most typical linkwords registered are those expressing contrast. This is due to the fact that sentences introducing limitations come after a statement implying some kind of achievement or advantage (e.g. “*However, we know that the rate of improvement in OCD is limited*” and “*No obstante, hay limitaciones que impiden...*”). With this function, the most recurrent linkwords in English is “however”, which has two equivalents in Spanish with the same function “no obstante” and “sin embargo”. Other linkwords go next to the word “limitation” to immediately add something positive and reduce the effect of the limitation (e.g. “*In spite of the above limitations, our data support that...*” and “*A pesar de las limitaciones comentadas el tratamiento con IECA y estatinas es la medida farmacológica que más normaliza la DMF,*). The most typical connectors with this function in English and combining with “limitation” are: “despite” and “in spite of” and in Spanish “a pesar de”, the latter being very frequently used. Sometimes, connectors used to introduce a conclusion are used such as “En resumen” in Spanish or “Hence” in English (e.g. “*Hence, even given the methodological limitation of this study, we believe...*” and “*En*

resumen, a pesar de las limitaciones de este trabajo, creemos que...”) This last exemplifying sentence shows the combination of two connectors which was not the only instance in Spanish (e.g. “*Sin embargo, a pesar de las limitaciones, nuestros resultados afirman...*”).

5. Problems /complications /Difficult- Problemas/ complicaciones/ Difícil

a) Problems-Problemas

The nouns “problems”/ “problemas” have an inherent negative connotation; this is why they are so suitable to express limitations. The general percentage of representation of the term “problem” including all the morphological categories is of 0.11% in English and 0.14% in Spanish, but these data refer to their representation in the whole corpus including all the moves. Logically, the move *Limitation* is the one in which they appear the most.

b) The verbs “to remain”- “incrementarse”, “agravarse”

In order to express that something has not been reached or solved but is still a challenge for the future, the verb “remain” is frequently used in combination with “problem” in this move (e.g. “*Infection of CSF shunts remains a difficult problem to treat.*”). A very similar collocation also takes place in the

Spanish corpus with the verb “permanecer” (e.g. “*En lo que se refiere al desequilibrio entre proporción de errores por..., el problema permanece e incluso se agrava ligeramente*”).

An alternative way to express the same idea is with the adverb “no” and the participle of the verb “resolver”: “no está resuelto” (e.g. *El problema de... no está resuelto*). If the problem does not just remain but becomes even worse the verb “agravarse” is used in Spanish as shown in the recent example “*el problema se agrava*”. Additionally, there are several elements apart from the verb that can convey a negative connotation like “desequilibrio”, “errores” and the adverb “incluso” that highlight the verb and make it more powerful (e.g. “*En lo que se refiere al desequilibrio entre proporción de errores por..., el problema permanece e incluso se agrava ligeramente*”). Instead of the verb “agravarse” it can also be said that the problem becomes bigger, and the verb “incrementarse” is used with this aim (e.g. “*...este problema se incrementa con el valor del coeficiente y disminuye con el tamaño muestral.*”).

c) The adjectives “grave”, “clínicas”, “delicado”, “inherente” and “diferente”- “different”, “serious”

The verb “agravarse” belongs to the same lexical field as “grave”. This is the most recurrent adjective that collocates with “problema” in this move (e.g. “*A pesar de esto, la gran prevalencia de la enfermedad coronaria hace que la insuficiencia cardiaca post-IAM continúe siendo un grave*”).

problema médico...”) Sometimes, the adjective “grave” comes after the noun “problema” as a complement rather than affecting the noun “problema” directly, resulting in the structure “problema de graves implicaciones clínicas” (e.g. “... *o se convierte en un problema de graves implicaciones clínicas*”).

The adjective “clínicas” also appears close to the adjective “grave” as it can be observed in the exemplifying sentence in brackets. Sometimes the syntactic structure is different, namely, the noun “gravedad” is used and “del problema” acts as a complement of this noun (e.g. “... *antes de poder progresar de una manera significativa en reducir la elevada mortalidad... será necesario demostrar la gravedad del problema*”). A synonymous structure found in the corpus to “gravedad del problema” is “magnitud del problema”, which refers more to the size but could be used in the same context as well (e.g. “*El uso de drogas inyectables, desde la perspectiva de la mortalidad, enfrenta gravísimas dificultades constituidas por la ausencia de datos analíticos que permitan evaluar con claridad la magnitud del problema*”).

Other adjectives that appear with “problem” are: “delicado”, “inherentes”, “diferentes”, etc, but none of them has an important representation in the corpus. The adjective “diferentes” underlines the limitation in the sense that it implies the existence of more than one problem (e.g. “*Las terapias conductuales..., contemplan diferentes problemas...*”). The equivalent literal adjective in English is also present in the corpus, “different problems”. In the following instance it is combined and preceded by the adjective “clinical”: “*As quoted*

earlier..., and paediatric SSTE and ITE pose different surgical problems”. Another example with exactly the same function but this time by means of an adjective is “muchos” (e.g. “*pueden ser muchos los problemas que se presenten desde un punto de vista pediátrico*”).

Regarding adjectives, the most common term accompanying the noun “complicaciones” is “graves”, where the adjective can be found before or after the noun (e.g. “*Aún así las broncoscopias no dejan de tener complicaciones potencialmente graves, por lo que deben realizarse por un personal cualificado... para resolver una posible complicación.*”). It is worth pointing out the habitual presence of the connector “a pesar de” in front of “complicaciones” (e.g. “*A pesar de las posibles complicaciones, podemos concluir que...*”). Sometimes the connector co-occurs with “complicaciones graves” as well. The connector is mainly used to state a contrast; most of the times trying to express an advantage despite a disadvantage or the other way round (e.g. “*El epoprostenol en perfusión intravenosa continua mejora la supervivencia y la capacidad funcional en los pacientes con HTP primaria y sus formas asociadas, a pesar de las complicaciones graves relacionadas con el sistema de infusión.*”) The presence of this kind of connectors leads to the overlapping of the *Move Advantage and Limitation*. The equivalent adjective of “graves” in English is “serious” (e.g. “*Infection remains the most serious complication of VP shunt placement.*”).

**d) The adjectives “methodological”, “surgical”
metodológico**

As for the English adjectives, those repeated in the corpus are “methodological”, “surgical” (e.g. “*Among the methodological problems of our study, the absence of a normal control group to allow for... is an important one*” and “*As quoted earlier, ependymoma is a surgical disease’ [27], and pediatric STE and ITE pose different surgical problems*”). The adjective “surgical” has a high rate of appearance in the English corpus; in fact, it is among the 10 most used adjectives with a representation of 0.14%. Comparatively, its representation in the Spanish corpus is only 2%. Therefore, the presence of the English adjective “surgical” in combination with “problem” is not surprising. Other nouns that collocate with this adjective are: “intervention”, “treatment”, “excision”, “procedure”, “technique”, “strategy”, “approach” and “revascularization”. Equivalent collocations in Spanish also take place in the corpus like: “intervención quirúrgica” and “tratamiento quirúrgico”, and other collocations such as “tiempo” and “corrección” do not have a literal equivalent in the English language. In the case of “surgical”, in the *Move Limitation* it only appears to collocate in medical texts with “problem” in English, but the collocation was not present in Spanish. As for the adjective “metodológico”, it does not collocate with “problema” in Spanish. This adjective collocates with “limitación” as shown in the section above. (see section 4).

e) The adjectives “serious”, “importante”, “significant”, “main” and “major”

As already shown, to define a problem as serious or important the adjectives “grave” and “delicado” are the ones used in Spanish. The English adjective “serious” also collocates with “problem” in general terms, but it is not used in the *Move Limitation* in the corpus but it is rather used in a context expressing an advantage (e.g. “*La técnica de Mitchell permite corregir en un solo tiempo quirúrgico y con pocas complicaciones, la severa y poco frecuente malformación del pene epispádico asociado o no a extrofia, resolviendo de forma bastante satisfactoria este serio problema.*”). Other adjectives present in English that clearly belong to the *Move Limitation* and can be considered equivalent to “grave” and “delicado” in this move are: “important”, “significant”, “main” and “major” (e.g. “*In depression, one major problem is that satisfaction judgements are clearly influenced by actual mood state.*”).

Obviously, the adjective “serious” is closer semantically speaking, but all those aforementioned, despite the different nuances in meaning could serve as synonyms as well.

As for the adjective “importante”, of course, it has an equivalent literal meaning in Spanish collocating with “problem”, but, as it was the case with “serious”, in the instances of the corpus it is used rather to express something positive than negative (e.g. “*...la Atención Temprana es factible para la mayoría de los niños de riesgo como los prematuros y*

niños de bajo peso y ...tendría aplicaciones de enorme trascendencia en este importante problema de salud pública”).

f) The adjectives “difficult”, “challenging” and “problematic” – “difícil”, “problemática”

Probably, the collocation of “problem” with “difficult” or “challenging” could also be compared to the adjective “delicado” in Spanish (e.g. “*Infection of CSF shunts remains a difficult problem to treat*”).

The Spanish adjective “difícil” also appears in sentences where the noun “problem” is present (e.g. “*La acalasia es un problema poco frecuente en niños, por lo que es difícil encontrar trabajos aleatorizados en los que se analicen...*”). This term with an inherent negative meaning appears as well when it functions as a noun “dificultad”(e.g. *Aunque las combinaciones terapéuticas parecen sugerentes no se deben soslayar los problemas planteados como... la dificultad en el análisis diferencial de los componentes del éxito terapéutico*²⁷) or as a verb (e.g. “*El estudio presenta algunas limitaciones... las propias de un estudio retrospectivo en un problema tan delicado con implicaciones morales, legales y sociales, que dificultan la obtención de una información objetiva y segura*”).

The noun “problem” appears in the context with the form of an adjective in English “problematic” (e.g. “*This is particularly problematic since the elderly have been shown to be at markedly increased risk for...*”). In Spanish, although the

adjective “problemático/a” comes up in the corpus, its use in this move as an adjective is not very common, in fact, most instances of “problemática” are nouns: “la problemática” and “esta problemática” (e.g. “*Se puede concluir que el estudio confirma la amplitud de la problemática del duelo en las nuevas consultas a...*”).

Finally, the noun “ausencia” is also present in combination with the noun “problemas” to make reference to the deficits of the research expressing limitation, for instance, “ausencia de datos clínicos” or “ausencia de sujetos control” (e.g. “*En primer lugar, la ausencia de sujetos control y el reducido número de casos clínicos presentados en cada modalidad terapéutica, resulta un problema de cara a...*”). As the exemplifying sentence indicates, the co-occurrence of various elements typically found in the *Move Limitation* is very common. In this case, “ausencia” comes up in the same sentence as “problema” and additionally combining with “reducido número de casos”. The use of “problema” to refer to the small sample size is also usual as in “*este problema se incrementa con el valor del coeficiente y disminuye con el tamaño muestral*”. Although, the verb “to lack” is also used in English to refer to the small sample size in this move, no instances were found of “small sample size” in combination with “problem”.

g) The nouns “difficulty”- “dificultad”

This term can be said to have a relevant representation in the corpus in general terms. Taking into account all the possible

morphological categories its representation is of 0.08% in Spanish and 0.06% in English.

The noun “dificultad” is frequently followed by the preposition “de” or “a la hora de”. The equivalent English structure is formed by the noun “difficulty” followed by the preposition “in”. For instance, researchers often express the troubles they can face if they aim at carrying out research in the future for different reasons. Comparable sentences in both languages expressing this idea are: “*La dificultad de realizar estudios de estas características es obvia, ya que requiere la disponibilidad de muestras muy amplias*” and “*However, considering the difficulty in carrying out studies in similar rural settings and using more sophisticated methods, the data show...*” Sometimes a different structure is used to express the same idea in Spanish, for instance, they express their awareness regarding the difficulty in performing further research by means of “ser consciente de las dificultades que supone en futuras investigaciones”: “*Somos conscientes de la dificultad que supone en futuras investigaciones el tomar en consideración simultáneamente todos los factores...*”

It is also habitual that researchers directly refer to the difficulties that those patients they have dealt with in their papers present. To transmit this idea equivalent structures take place in both languages, namely “to have difficulties in” and “tener dificultades para”. Two examples from the corpora showing this are: “*However, considering the difficulties that anorectic patients may have in complying with the administration of drugs, previous and concomitant*

psychotherapeutic interventions should be focused on the problems concerning...” and “Estas pacientes presentan a menudo desajustes menstruales tanto por la propia enfermedad psíquica como a causa de los tratamientos psicofarmacológicos, tienen dificultades para programar ...”

It is not very common to find the nouns “difficulty” and “dificultad” preceded by an adjective. Only in the Spanish corpus instances were registered of the adjectives “gravísimas” and “serias” that help emphasize the noun “dificultad” (e.g. “*El uso de drogas inyectables, desde la perspectiva de la mortalidad, enfrenta gravísimas dificultades constituidas por la ausencia de datos analíticos que permitan evaluar...*”).

Occasionally, despite the rare appearance of the *Move Advantage* in medical research articles, the noun “difficulty” comes up expressing something positive. As it has been observed throughout the *Move Limitation*, there are many words with a positive connotation that are turned into negative by means of other inherent negative words. The following example containing the word “difficulties” shows the opposite process, namely, a negative word being turned into a positive by the use of the verb “resolve” combining with “difficulty”: “*We have thus identified a few factors ...and have contributed to resolving the difficulty of articulating direct subjective QOL interventions*”.

h) The adjectives “difficult”- “difícil”

The adjective “difficult” collocates with the verb “differentiate” because researchers want to express many times that differentiating between two things is not always easy (e.g. “*It is difficult to differentiate NE cysts form other cystic lesions...*”). Additionally, authors resort to the verb to “treat” in combination with the adjective “difficult” to refer to the difficulties in coping with certain kinds of problems or dealing with certain types of patients (e.g. “*SVS patients who have failed ... are a difficult patient population to treat*”). In Spanish, no instances of the adjective “difficult” were present combining with the verb “diferenciar”. As for the word “tratar” another term belonging to the same word field, namely the noun “tratamiento”, was registered in combination with the adjective “difícil” giving rise to “difícil tratamiento” (e.g. “*... en el caso de los hospitales de tercer nivel a los que, cada vez más, se les exige que asuman enfermos de diagnóstico y/o tratamiento difícil y complejo, se les pondría en una situación plagada de contradicciones*”).

Alternatively, instead of the adjective “dificultar” a word belonging to the same word family appears, namely, the verb “dificultar”. This time, the structure becomes the opposite, i.e. the verb “dificultar” collocates with the noun “tratamiento” rather than with the verb “tratar” resulting in “dificulta el tratamiento” (e.g. “*La clasificación de Virchow de 1863 debe ser desechada por no ajustarse al comportamiento clínico y biológico de las anomalías vasculares y provocar una confusión*”).

*en la nomenclatura que dificulta considerablemente su estudio y tratamiento.”). However, the collocation “tratar”+”difícil” was present but it expresses an Advantage rather than a disadvantage as the following expressions indicate: “pueden ayudar a tratar”, “puede tratar de forma efectiva”, “eficaces para tratar” (e.g. “*La dobutamina y el levosimendán son fármacos eficaces para tratar el síndrome de bajo gasto después de la cirugía cardiaca*”).*

Nevertheless, there was also an example in which “tratar” is used to express a limitation but not co-occurring with “difícil” but with the noun “problema” (e.g. “*Aunque...se sabe poco al respecto de su fuerza de asociación, motivo por el que al haber efectuado el estudio con un amplio seguimiento familiar, sí que nos sugiere sobre la importancia de tratar este problema*”).

i) The verbs “comparar”, “ponderar” and “cuantificar” and the verbs “dificultar” and “to make difficult”

The most recurrent verbs in Spanish combining with “difícil” are those used to express that sometimes certain things cannot be evaluated accurately. To convey this meaning “difícil” comes with the verb “comparar” as in: “*Aunque resulta difícil comparar los resultados de este estudio,...*”. Sometimes, the collocation takes place with the noun “comparación” rather than with the verb: “*La variabilidad entre las diferentes intervenciones publicadas hace difícil su comparación y*

definición del procedimiento ideal". The sentence expresses the difficulty on the part of the researcher to describe things when there is no "regularity". In fact, the word "variabilidad" is used for this purpose. The same observation of the researcher can be recognized by means of the noun "regularidad" together with "dificultad" in the following sentence: "*En tercer lugar se detectaron unas regularidades de difícil interpretación*".

With a similar intention, "difícil" collocates with verbs expressing an evaluation such as "ponderar", "cuantificar" (e.g. "*Se trata, en gran medida, de percepciones subjetivas, difíciles de cuantificar y objetivar.*"). On some occasions, the researcher points out that it is the lack of certain data that makes the evaluation difficult" (e.g. "*El diagnóstico correcto de las anomalías... es difícil de llevar a cabo por..., la falta de marcadores específicos...*").

A similar English structure is the combination of the adjective "difficult" with the verb "to make". The following example shows the combination "making it difficult to". It is interesting to observe the overlap between the *Move Limitation* and *Background* due to the existence of the words "in earlier studies": "*In earlier studies, the methods used to quantify outcome measures were not consistently defined, making it difficult to establish the true effects of dorsal rhizotomy*".

The structure "making it difficult to" is equivalent to the Spanish verb "dificultar", which does not exist as a unique term in English. In Spanish it is also used to express the difficulty of attaining accurate results (e.g. "*las propias en un problema tan delicado con implicaciones..., que dificultan la obtención de*

una información objetiva y segura”). As the last exemplifying sentence illustrates, it is also common to combine various words with a negative connotation to express a limitation such as “problema” and “dificultan” in the same sentence.

In general, the adjective “difícil” is often used to express a complaint on the part of the researchers. Alternatively, it is frequently used by researchers to justify themselves for the inaccurate results they have provided the reader with. This last option is very typical in Spanish (e.g. “*De una serie tan reducida es difícil obtener conclusiones sólidas*”).

Finally, it is relevant that the adjective “difficult”, with a negative connotation, occasionally expresses something positive in a sentence. This happens when it is combined with words inherently positive, as it happened with “difficulty” (e.g. “*It is especially suitable for difficult cases with multiple shunt revisions or previous laparotomies performed*”). No equivalent structures were registered in Spanish in this sense.

j) Complication- Complicación

The general percentage of representation in the corpus of the word families of these two terms is 0.04% in Spanish and 0.06 % in English.

Equivalent structures were found between the corpora, and additionally, as it happened with the other terms in this section, advantages were also expressed in both languages by means of these terms. As already shown above, when these words are accompanied by certain terms, they turn to express the

opposite and lose their inherent negative connotation. An example would be the combination of the noun “complications” and the verb “to reduce” (e.g. “*Newer techniques, including stereotactic radiosurgery, conformal radiation and intensity modulation, have been developed to reduce the complications*”). The equivalent occurrence in the Spanish language is a collocation consisting of the noun “complicaciones” + the adjective “postoperatorias” accompanied by the adjective “mínimas” (e.g. “*Actualmente la cirugía laríngea en un tiempo puede realizarse de forma segura con mínimas complicaciones postoperatorias*”). The adjective “mínimas” turning the sentence to express something positive also collocates with “complicaciones”, obviously not in the move Limitation, because it always expresses some kind of advantage (e.g. “*... técnica que presenta mínimas complicaciones y permite una rápida recuperación del paciente y un excelente resultado funcional*”). Interestingly, the combination “complicaciones postoperatorias” appears in the corpus only to express advantage (e.g. “*Las complicaciones postoperatorias fueron pocas y de fácil resolución*”).

k) The nouns “índice” and “rate”

An equivalent structure which constitutes a collocation in both languages is the use of words to quantify “complications”. In Spanish “complicaciones” are evaluated by means of the words “índice” or “porcentaje” thus resulting in the structures “índice de complicaciones” and “porcentaje de complicaciones”

(e.g. “*Los resultados han sido muy buenos a pesar del elevado índice de complicaciones*”). Additionally, it can be stated that the collocation “índice de complicaciones” is frequently preceded by an adjective quantifying the noun “índice”, namely “elevado” or a synonym of this word, “alto” and thus turning into “elevado índice de complicaciones”. Moreover, it is also common for the collocation to be present in a sentence where results are being evaluated, and therefore the noun “resultados” can appear in the sentence as well: “*Los resultados anatómicos, estéticos y funcionales han sido muy buenos, a pesar del alto índice de complicaciones*”. Alternatively, the combination “alta incidencia de complicaciones” also comes up in the corpus, but it is not used to express a limitation.

The literal equivalent to the noun “índice”, “rate” was found in the English corpus with the same use in the *Move Limitations* (e.g. “*The TT group of ACS patients was higher rates of in-hospital complications, including a 4-fold increase in in-hospital major bleeding, as compared to ...*”). As the example shows, also the adjective “high”, comparable to “elevado” and “alto” in Spanish appears together with the noun “complicaciones” in Spanish, this time with the comparative adjective “higher”: “higher rates”. As for “complicaciones postoperatorias” no equivalent was found in English. Thus, this would be an example of “zero correspondence” according to Stig (2007: 25). Alternatively, the noun “incidence” is used in English to express the same idea, this time using the antonym of “high” and giving rise to “lower incidence”: “*Actually, more refined surgical techniques and more effective diagnostic*

examinations have resulted in a lower incidence and more appropriate treatment of complications as compared to reports published in the seventies”.

Another noun that appears with “complicaciones” in this move is “procedimiento”. It is used to explain that using a certain method or procedure implies coping with certain complications. It comes up accompanied by the adverb “no”, the adjective “exento” and the preposition “de” resulting in “procedimiento no exento de complicaciones” (e.g. “*Por supuesto, se trata de procedimientos complejos que no están exentos de complicaciones*”).

1) The nouns “mortality” and “mortalidad”

As stated previously, sometimes researchers combine words with an intrinsic negative meaning (e.g. “complications”) with other terms [e.g. unusual] that turn them to express something positive, namely, an advantage. In the sentence: “*In conclusion thrombosis associated with transvenous pacing leads is an unusual clinical complication occurring in 0.6–3.5% of all implants. However, when it occurs it carries with it significant morbidity and mortality.*” by means of the adjective “unusual” the noun “complication” seems to turn the sentence into a positive statement. However, the appearance in the sentence of the noun “mortality” converts the final information into a limitation. This process of turning words with a positive meaning into negative occurs very often in medical texts. For this reason it is difficult to assess to what extent a corpus

expresses more advantages or disadvantages based on words with inherently positive or negative meaning.

In Spanish, the noun “mortalidad” appears in equivalent structures in combination with “complicaciones graves” (e.g. “*La mortalidad precoz y al año ha disminuido respecto a la de 1995, aunque la incidencia de complicaciones graves sigue siendo similar*”).

It is also relevant that the noun “mortalidad” often combines with the adjective “elevada” to express a limitation and thus co-occurs as well with the noun “complicaciones” (e.g. “*En definitiva, el REIAM es una complicación poco frecuente en la UCIC pero se acompaña de una mortalidad elevada*”). The adjective, can also be placed before the noun “mortalidad” (e.g. “*La IC es una complicación frecuente en el marco del infarto, con una elevada mortalidad hospitalaria...*”).

Finally, the adjective “complicated” is frequently used in this move (e.g. “*Retractor-based flexible arm systems while successful, are more complicated to use and provide a less rigid navigation platform*”). After this adjective, it is usual to find an immediate change of move to *Further Research* to explain that, because of the difficulty of something, more research is needed (e.g. “*It seems to have, independent of its association with bipolar disorders, a more complicated evolution... Further studies are needed to...*”). The adjective “complicated” is more frequently used in English, where 7 instances were recorded (0.01%) than in Spanish. In fact, only two instances of the adjective “complicado” were registered in the corpus (e.g. “*El tema es complicado pero ineludible...*”).

m) The adjectives “secundaria”- “secondary”

Another adjective that collocates with “complicaciones” is “secundarias”, used by doctors to explain that something has been originated from or is due to something else (e.g. “*El 13.9% de los niños con defectos de pared abdominal tratados en nuestro hospital presentaron complicaciones secundarias a la MI no tratada*”).) A literal equivalent structure present in the English corpus exists in the form of “secondary complications”, but it does not appear in the move *Limitation* (e.g. “*Perhaps this is because of the view that procedures such as shunting are more acceptable and avoid possible criticism from a potential complication secondary to cranial remodelling, a more extensive procedure*”).

Finally, other adjectives that take place in the corpus in combination with the noun “complications” are “cutáneas”, “cardíacas” and “obstétricas”, but they are specific of dermatology, cardiology and paediatrics, they do not belong to general medical English.

6. Cannot/Not possible/Not known- No podemos/No es posible/No sabemos

a) The noun “conclusions”, the verbs “can” and “generalize”

It is vital at the end of a paper to inform the reader on the conclusion of the investigated topic and on the possible extension of its application. Many times, researchers have not affirmed that their findings can be applied to a great population or the one they intended on the grounds of their small sample or due to other potential problems they faced during research. To do so, researchers tend to use the noun “conclusions” and combine it with the modal “can”, which normally expresses the capacity of doing something, but in the negative form to express incapacity: “cannot” (e.g. “*However, this conclusion cannot be applied to acute dioxin intoxication as a result of...*”).

The equivalent literal construction takes place in Spanish by means of the noun “conclusiones” and the verb “poder” that is placed after the adverb “no” to express impossibility: “no podemos” (e.g. “*Dado lo pequeño de la muestra no es posible extraer conclusiones en...*”). Additionally, in English, to express the idea that the study cannot be generalized to other people or areas of study the use of verbs such as “generalize”, “extrapolate” often co-occur with the modal verb in the negative form “cannot” in front of the noun “conclusions” in a passive structure (e.g. “*Thus, firm conclusions cannot be extrapolated from such a small series*”).

Although the most habitual is to talk about “conclusions”, also “results” and “findings” are equally used for the same purpose (e.g. “...and that results from different studies in this respect cannot be generalized...”). The equivalent literal noun to “results” in Spanish, “resultados”, also comes up in this kind of structures, but mainly with the verb “saber” originating: “no sabemos...+ resultados” (e.g. “No sabemos hasta qué punto podemos extrapolar los resultados”).

Actually, the word family of the term “conclusion” is also employed to express the impossibility of stating definite results. E.g. it is used as a verb “to conclude” in a passive structure (e.g. “However, it cannot be concluded that hypnosis does not constitute an effective form of treatment”). As the example shows, the presence of the connector “however” is also typical; since the impossibility of concluding results is normally stated after researchers have written what their achievements have been (e.g. “However, this conclusion cannot be applied to acute dioxin intoxication as a result of...”).

In Spanish, the adjective “concluyentes” is also present in the corpus to transmit the impossibility of transmitting concluding remarks because of the small sample size (e.g. “En nuestra muestra no es posible determinar diferencias concluyentes respecto a los ISRS en monoterapia, principalmente por el tamaño de este grupo”).

b) The active voice in Spanish vs. the passive in English: The verbs “poder”, “afirmar”-“state” and “conocer”- “know”

A main difference between the two languages regarding the transmission of the idea of impossibility is the fact that in English most cases are in a passive structure and the subject is never stated. The only exception registered in the corpus is: “we are unable” that was present on one occasion in which the subject is clearly stated (e.g. “*We are unable to give diagnoses of clinical depression to the cases evaluated. This may be regarded as a serious limitation...*”). Apart from the structure “something” cannot be + participle”, in which the agent of the passive is not mentioned, the active is also used in the structure “It is/was not possible to” (e.g. “*It was not possible to directly address this issue in the present study because of the lack of respective instruments designed for children*”). By contrast, in Spanish the active is mainly used. Sometimes the subject is omitted but clearly recognized. Thus, although the impersonal structure “no se puede” is also present (e.g. “*Asimismo... no se puede atribuir el cambio en... se deba de forma exclusiva a la evaluación-intervención realizada en cada caso*”), in most examples the form “no podemos” comes up.

The aforementioned form collocates with the verb “afirmar” (e.g. “*No podemos afirmar con seguridad que nuestra intervención fuera eficaz en la reducción o aumento de...*”). The equivalent literal verb “to state” is also present in the same structure (e.g. “*It is, however, not possible at this time to state*”).

the exact numbers of...). As far as the verb “saber” is concerned, the same differences regarding active and passive occur. Although the impersonal form “no se sabe” appears in the Spanish corpus (e.g. “*No se sabe si esto es debido a una diferencia en la detección... o bien a alguna dificultad...*”), it is more common to use “no sabemos” (e.g. “*No sabemos hasta qué punto podemos extrapolar los resultados.*”). Also the translation of “know” for “conocer” was registered on one ocasión with an active personal structure: “*No conocemos nada acerca de muchas de las variables...*”. On the other hand, in English, normally it is the passive that comes up “something+is/are not known” (e.g. “*The long-term sequelae of these complications are not known*”).

The conditional conjunction “si” is also frequent in Spanish to convey lack of confident knowledge (“*No es posible, en la mayoría de casos, discriminar si se trata de trastornos inducidos por las sustancias, o...*”). To underline this idea of the impossibility of definite results other words such as: “con seguridad”, “de forma exclusiva”, and “irregularidades” are used (vid. examples above). Similarly, other terms with the same function in the English corpus are: “exact”, “clear evaluation”, “rule out”, etc. (e.g. “*Our study design cannot rule out whether amalgam patients only...*”).

7. Incomplete-Incompleto

a) The adjectives “complete”, “incomplete” and “insuficiente”

The idea of incompleteness has been shown by means of different elements up to now in both languages. This idea expresses a Limitation in the sense that research is not complete and as a consequence, it links and overlaps many times with the Move *Further Research*, since more investigation is needed. To express the idea of incompleteness, several points have been crucial: the use of time adverbs as in “... *is still in its infancy*” (vid num 1), but also other relevant terms present in the move *Further Research* such as: the verb “to remain”, the verb “to need” and expressions like ...”*is very far from being defined*”. Similarly, in Spanish the idea of incompleteness is transmitted by means of time adverbs (vid num 1), the verb “quedar”, the verbs “necesitar”, “precisar” and “requerir” and expressions such as “*estamos lejos de*” .

Another strategic verb to express that research is incomplete which actually belongs to the move *Further Research* is the verb “to lack”. Analyzing the verb with more detail, it is observed that before the words like “resources”, “data”, “studies” or “a definitive tool” it underlines the idea of incompleteness or impossibility of carrying out research (e.g. “*This lack of psychopathological resources to explore such an elusive subjective experience is an intrinsic limit of the present experimental protocol,...*”). As it was demonstrated previously,

this structure would be equivalent of the Spanish “no tenemos” or “falta” (e.g. “*Este se debe a que aún no existe una definición uniforme del., por lo que falta claridad y con frecuencia provoca confusión*”).

Other additional elements to express incompleteness are the adjectives “complete” and “incomplete”. The former when preceded by a negation (e.g. “*This discussion is by no means complete, given...*”) and the latter with the inherent lexical value of incompleteness (e.g. “*Since incomplete surgical LAA closure may promote rather than reduce risk of stroke, there is a need for improving...*”). Unlike in English, the Spanish adjective “completo” is mainly used to compare the actual research with previous research and to refer to the degree of agreement or disagreement with other colleagues. For this reason, the adjective “completo” collocates with the verb “estar de acuerdo” and occurs mainly in the *Move Background* in structures like “estar en completo acuerdo”.

As for its antonym “incompleto”, it is hardly employed and additionally it is not used to express limitation. The actual Spanish equivalent adjective is “suficiente” and its antonym “insuficiente” (e.g. “*Por todo ello, podemos considerar que la evidencia farmacoeconómica es aún insuficiente para extraer conclusiones relevantes que ayuden...*”). Nevertheless, this adjective is not always literally equivalent to “incomplete” since many times its translation in English would be “not enough”, but “enough” is rarely used in the corpus (e.g. “*En un 30% no hay información suficiente, muchos de los vínculos que se califican como buenos, la mitad del total, cabe matizarlos con una mayor*”).

profundización;...”). Additionally, an instance in the English corpus of the adjective “sufficient” was present: “However, the clinical data in the literature are not sufficient for conclusion”.

Researchers tend to talk about the results they have achieved at the end of a paper. In fact there is a section in Research Articles called “Results”. Often a brief summary appears in the conclusion. To refer to these results, they use the same noun, namely “results”. In the present move, *Limitation*, to indicate that results have been disadvantageous it is common to find the structure “Los resultados no” in combinations with “son clarificadores”, “son consistentes”, “mostraron ventaja”, etc. (e.g. “*Los resultados no mostraron ventaja global ni local...*”).

Finally, references to time took place in the Spanish corpus to justify the incomplete work or just as a means of expressing a limitation: “*La escasez de tiempo ha provocado que no se pudiera profundizar en algunos aspectos*” or “*... es una terapia efectiva, pero consume mucho tiempo y esfuerzo*”.

3.3.2.4 *Analysis of the move Further Research*

This move aims at informing the reader about a possible extension or about other applications of the investigation. When comparing the two languages 11 common word fields were found as it is shown in Table 3.57 below:

1	Study-Estudio
2	Research/Investigation- Investigación
3	Need- Necesitar
4	Remain- Falta/Queda
5	Improve-Mejorar
6	Hope- Esperar
7	Understanding-Comprender
8	For/in the future- En el/un futuro
9	Interesting- Interesante
10	Still- Todavía
11	Sample- Muestra

Table 3.57 Common semantic fields in the move Further Research in English and Spanish

1. Study- Estudio

As already mentioned in the previous move, “study” has a very high rate of appearance in both languages, so it is not surprising that it is also present to suggest possible studies for the future.

The most recurrent structure for this move is the combination of the comparative “further” with the noun “study” combining with the verb “to need” either in a passive structure (e. g. *Further studies are needed to support this hypothesis*) or simply in an active structure with the verb “to be” (e.g. *“Further follow-up studies and ... are necessary ...”*). In fact, this comparative adjective preceding the noun study is present 0.03% and even the structure “studies are needed” has a representation of 0.02% in the corpus.

Although this is the most typical adjective for researchers to indicate that more studies are needed, other

adjectives and adverbs also collocate with this noun in English such as “prospective”, “future”, “more” and “larger”. Only one example was present in the corpus with the adverb “more” (e.g. “*More studies with SPECT using... are required to predict...*”), however, the two former adjectives have a 0.02% of representation in the corpus. A difference among the 3 is that “prospective” is occasionally found in singular structures. Ex: “*A prospective study is necessary to causally link...*”

In Spanish, the noun “estudios” occurs in the plural although an example of the singular was present: “... *supone una garantía para la viabilidad de un futuro estudio*”.

Regarding adjectives that collocate with the noun “estudios” in this move, it is observed that in Spanish, there are not so many adjectives used to express the idea that further studies are needed. Only the adjective “nuevos” is used to convey the same idea. Ex: “*Nuevos estudios clínicos podrán aportar conocimientos significativos...*” In Spanish it is more frequent, though, to find the adverb “más” as an equivalent function to the English adjectives mentioned above (e.g. further, prospective, etc) but its representation in terms of frequency is relatively low. In fact, despite the adverb “más” having a representation of 0.61% in the corpus, it only appears 8 times preceding the word “estudios” (e.g. “*A pesar de ello, es preciso realizar más estudios para demostrar que la mejoría de la disfunción endotelial se asocia, además, con una mejoría del pronóstico en la MCDI.*”)

A way to indicate the need for more studies is by means of the verb “necesitar” mainly combined with the adverb “más”,

for instance: “Son necesarios más estudios”, although other verbs also collocate with this adverb such as: “Se requieren más estudios”, “harían falta más estudios”. But the interesting contrast is that many other verbs are used instead of “necesitar” to convey this meaning, especially: “se ve oportuno realizar estudios”, the modal “deber”: “se deben realizar estudios”, “deberían estudiarse” and others like “sería deseable realizar un estudio”. However the most common structure is that with the verb “ser”, where there is evidence that the verb “realizar” collocates with “estudios” (e.g. “*es necesario realizar más estudios*”). This sentence is often introduced after a linkword starting the sentence and establishing a contrast such as: “no obstante” and “sin embargo”. Sometimes other words are added with an equivalent but different meaning “podrían estudiarse en mayor profundidad”, which is very similar in meaning but it underlines the idea of studying more deeply. Another example is “de ahí la importancia de estudios en esta dirección”, which is also similar in meaning but it adds the idea of considering it important.

Back to the English adjectives mentioned above, “larger”, is usually combined with other words preceding “studies” such as: “prospective”, “multicenter”, “scale”, etc. Interestingly, all the 3 adjectives are used in the same type of sentence with the noun “study” and the verb “to need” (e.g. “*Prospective studies are needed*”; “*Further studies are needed to support this hypothesis*”). In the case of “larger” sometimes “further” came up in the same sentence, but in this case the word “further” acts as an adverb and not as an adjective, affecting the

verb rather than the noun. Ex: “*Larger prospective studies are needed to further assess this therapeutic approach*”. This function of “further” is not isolated, it also appears in sentences with “further studies”, where the verb “to need” is not present: “*Further studies should be designed and performed to further support...*” This structure with “to further” also combines with the verb “to warrant”, e.g.: “*Larger scale studies are warranted to further characterize...*” But in general it can be stated that “further” is very often used with this function in this move (e.g. “*Studies in ... will help us to further understand the effects...*”). All these instances without the verb “to need” could be equivalent to the Spanish ones stated above, in which different verbs were used instead of need. Finally, it can be concluded that in general “further studies” collocates with the verb “to warrant” as the following example indicates: “*The unexpected finding of ... warrants further study*”.

As for the collocation “future studies”, it is worth mentioning that very often either a future tense or a modal verb is used afterwards (e. g. “*Future studies might be performed with*”, “*Future studies will determine..., etc.*”). Verbs coming immediately after “are necessary” or “are needed to” are: “clarify”, “define”, “elucidate”, “identify”, “support”, “prove”, “show”, “confirm”, “investigate”, “assess”, “characterize”, “focus”, “establish” and “determine”. It can be observed that the adjective “future” and “futuro” is more frequently used in English than in Spanish, where only 7 instances could be found (e.g. “*Se deberían considerar estudios futuros para comparar la eficacia de intervenciones psicosociales, farmacoterapia y*

tratamientos combinados".) It should be underlined that "futuros estudios" appeared 4 times and "estudios futuros" 3. These 3 examples could be a consequence of the English influence on the Spanish language in the scientific world, since normally the order in Spanish is noun plus adjective. It is also interesting that this combination of two words also collocates with the verb "necesitar" or "precisar", as in English ("need" and "require"), but they also collocate with "línea de investigación". Ex: "*Finalmente, precisar la necesidad, dentro de esta línea de investigación, de proseguir con futuros estudios que traten de...*".

The modal "deber" frequently occurs near "futuros estudios" (e.g. "*Futuros estudios deberían poner a prueba estos resultados con...*") with an equivalent value to that of "further studies should".

Back to the most frequently used collocation "further studies", it should be mentioned that sometimes a passive structure comes first followed by "further studies" with a function of complement and preceded by the modal "should": e.g. "*... should be investigated in further studies*". Also the future tense appears frequently with "further studies". e.g. "*Further studies will be necessary*" and even with the verb "to need" "*Further studies will need to identify*". Instead of the verb "to need", the verb "to require" acts as a synonym very often. e.g. "*Further studies are required to determine...*" This synonym has its equivalent in the Spanish language "*Se requieren más estudios para evaluar...*"

It is also common to find the noun “study” in the plural form with no word preceding it and in a sentence stated in the future (e.g. “*Studies will be directed toward the comparison in...*”). Sometimes, the plural noun follows the determinant “such” (e.g. “*...such studies could also help clarify the emerging idea that...*”).

Less common is the presence of a passive with the plural noun, but it might be the case: “*Studies need to be done to address this issue*”.

A very common structure in this move which is also present in the *Move Summarizing* is “the results of this study”. This time the structure comes prior to what researchers consider should be the next step in terms of investigation. Therefore suggestions, recommendations, etc on the part of the researcher are expressed by means of verbs like “show”, “suggest”, etc. Ex: “*The results of this study show that there is a need for*”, “*lead us to make several recommendations for future research on...*”, “*suggest that closer attention to...*”, “*might be fundamental for understanding*”, “*demonstrate the necessity of further investigations on...*”, “*show that there is a need for improvement*”, etc. Many of the examples shown in these lines demonstrate that the structure “the results of this study” also collocates with the verb “to need” and with “further” in this move. In Spanish, though, the corresponding structure “los resultados de este/nuestro estudio” does not appear in this move but rather in the *Move Summarizing*.

2. Research /Investigation- Investigación

The appearance of this noun has a higher percentage in Spanish (0.13%) than in English (0.04%), however, it has to be considered that the synonym “research” in English has a rate of appearance of 0.11 %, thus, taking this into account the rate in both languages is very similar.

The first thing that calls the attention is the fact, that as expected, in English the passive is used in most cases where this noun comes up, whereas in Spanish the so called “pasiva refleja” (e.g. “se” + verb) is used instead (Alarcos 1995: 219). E.g. “se recomienda”, “se precisa”, “se hace necesario...” (e.g. “*Se recomienda la utilidad de seguir investigando con este tipo de trabajos*” vs “*Longitudinal studies are required to investigate whether...*”)

Occasionally, the term “investigación” collocates with the noun “importancia” in Spanish (e.g. *De aquí la importancia para futuras investigaciones...*). The noun “importante” collocates with the verb “to investigate” which corresponds to same word family as “investigación” (e.g. “*Hence, the importance of investigating ... must be emphasized*”).

Unlike in Spanish, “should” in combination with “investigation” and “research” is widely used in English both in active and passive structures (e.g. “*This technique should be the subject of further clinical investigations*” and “*These interrelationships should be investigated in further studies*”). The modal “can” is the second in terms of frequency after

“should” and it appears mainly in the past form “could” (e.g. “*Further research on the PCLS could explore...*”). The use of this modal in this tense is not so frequent in Spanish. However, there is a correspondence in the present tense in both languages (e.g. “*La diversificación de los tratamientos puede ser, el desafío más importante para la investigación*” and “*Further investigations and early treatment can stabilize...*”). It is also worth mentioning the presence of Spanish structures that implicitly include this modal in terms of meaning: (e.g. “*Lo que señalamos como posible línea de investigación futura*” which is equivalent to “*Lo que puede ser una línea de investigación futura*”). Other modals in English in combination with “investigation and research” are “may” and “must”. The former occurs more often than the second, but neither of them as much as “should” or “can”. Nevertheless, equivalence is present in the case of “must” (e.g. “*De aquí la importancia para futuras investigaciones que deben considerar y controlar...*” and “*The obsessive thoughts must be investigated in order to find...*”). As for the modal “may”, there is no direct equivalent in Spanish in the corpus, so it can be concluded that the use of modals is much frequent in English than in Spanish.

As it already occurred with the noun “study”, the verbs “need” and “require” frequently appear in combination with “investigation” and “research”. Again, the passive form is mainly present although some cases of active also occur: (e.g. “*Further studies are needed to investigate the impact of...*”, “*Further biological investigations are necessary to confirm...*”). Regarding Spanish there are also evaluative structures like “es

preciso” and “es necesario” (e.g. “*Es preciso por tanto seguir investigando para comprobar...*” and “*Es necesario, por tanto, a nuestro entender, crear o adaptar...*”)

It is common in Spanish to use the verb “seguir” plus a gerund (e.g. “*Se recomienda la utilidad de seguir investigando con este tipo de trabajos...*”). This verb, although also present in English, it is much less widely used (e.g. “*Therefore, continuing to investigate... is of crucial importance...*”).

As a conclusion it can be said that instead of the passive structure used in English, Spanish uses other ways to convey the idea that further studies should be carried out.

Regarding adjectives, in both languages the nouns “investigation”, “research” and “investigación” collocate with the adjective “future” and “futura”. As for the adjective “further”, it was only registered twice with “investigation” but it appeared 13 times with the noun “research”.

In Spanish all the instances with the adjective “futuro” appear in the plural (e.g. “*Es en esa dirección hacia donde orientamos nuestras futuras investigaciones*”), whereas in English it appears both in singular and plural (e.g. “*Whether patients with such atrial structural characteristics can obtain better long-term outcomes with a rhythm control strategy warrants future investigation...*” and “*Nevertheless, these variables should be controlled in future investigations to get more consistent results*”). The previous example illustrates the usual position of the adjective in English before the noun, whereas the adjective “futuro” varies its position in Spanish since sometimes it precedes the noun, namely as an explicative

adjective (as shown in the example above) and sometimes it follows it, namely when it has an especificative value (Alarcos, 1195: 81-82) (e.g. “*Las líneas de investigación futuras deberían valorar...*”). Additionally, as this latter example shows, in Spanish these terms appear combined with the noun “líneas”. The grammatical function of some items varies though, the noun “investigación” turning into a complement of the noun “líneas”.

Regarding the grammatical form of “investigación” and “investigation” in the two languages, there are also some differences. In Spanish “investigación” is mainly found as a noun or as a gerund following the verb “seguir” (e.g. “*La segunda línea de investigación se encargaría...*” or “*Se recomienda la utilidad de seguir investigando con...*”). In English the world field of “investigation” occurs in the corpus. It mainly appears as a participle in passive structures with modal verbs (e.g. “*The interrelationships should be investigated in further studies*”). It also appears frequently as an infinitive after passive structures (e.g. “*Further studies are needed to investigate...*”) and as a noun (e.g. “*Whether ... warrants future investigations*”).

3. Need- Necesidad

The rate of appearance of “necesidad” is 0.13% taking into account the whole world field (*necesario* 55, *necesarios* 9, *necesidad* 51, *necesariamente* 8, *necesita* 8, *necesitan* 6, *necesitará* 1, *necesitamos* 2), but it mainly appears as an adjective and as a noun, each category representing 0.05%, thus

0.1% of the total corpus together. In English the percentage of appearance in the different grammatical categories is of 0.27% (need 64, *needs* 16, *needed* 39; *is needed* 12 and *are needed* 18, *necessary* 35, *necessarily* 2) being the verb and the noun the most recurrent with 0.20%. These data indicate that this word has a higher rate of appearance in English than in Spanish.

As it was already shown previously, this term combines with the two items above “study” and “investigation” in both languages. In Spanish, it is common to find “líneas de investigación” in the immediate sentence after that containing “necesitar” (e.g. “*Es necesario por tanto, a nuestro entender, crear o adaptar... Queda abierto, pues, un interesante camino para futuras investigaciones*” or “*...se hace necesario seguir profundizando... Es en esa dirección hacia donde orientamos nuestras futuras investigaciones*”).

A very common structure in English is “there is a need for” (e.g. “*There is a need for improving...*”), although in general it can be stated that this noun collocates with the preposition “for” (e.g. “*The need for enhanced diagnostic methods that allow for ... may contribute to...*”). As for the noun “necesidad”, it sometimes appears preceded by the verb “destacar” and followed by the preposition “de” plus an infinitive (e.g. “*Queremos destacar la necesidad de hacer estudios a largo plazo acerca del tratamiento...*”). Sometimes it is preceded by the adjective “urgente” (e.g. “*Esto daría cuenta de la urgente necesidad de crear modelos de trabajo...*”), this also happens in English (e.g. “*Despite our encouraging results... emphasizing the urgent need to develop...*”).

Two clear collocations of the term “necesidad”, not just as a noun, but also in its function as an adjective “necesario” are the verbs “evaluar” and “realizar” (e.g. *Se plantea, la necesidad de evaluar de forma fiable...*” and “*Basándonos en los datos de nuestro estudio nos replanteamos la necesidad de realizar un ...*”). As for the adjective, most of the times it comes in the masculine form (e.g. ... “*sería necesario realizar una revisión en profundidad del...*”), although sometimes the feminine is present (e.g. “*Estimamos necesaria la continuidad en esta línea de investigación...*”).

Finally, the adjective “necesario” often appears together with the verb “resultar” (e.g. “*Resulta necesario fomentar y facilitar investigaciones...*”). In addition, the pair “resulta necesario” often occurs with the verb “desarrollar” (e.g. “*Resulta necesario desarrollar abordajes preventivos que tengan como objetivo la población general*”). Regarding English, the verb “to evaluate” also combines with “need” (e.g. “*... needs further evaluation*”) sometimes, the verb appears more distant from the noun “need” (e.g. “*Our findings are surprising and may indicate a need for a prospective, randomized clinical study to properly evaluate the role of...*”). A similar verb in meaning also present in English is the verb “to assess”, which is frequently followed by “whether” (e.g. “*There is a need for a ... to assess whether...*” or “*... needs to be further examined to assess whether...*”). The verb “realizar” does not have a literal equivalent in our examples in English, but instead the term “examine” both as a verb and as a noun are used (e.g. “*... needs to be further examined*” or “*The*

clarification of such effects in turn justify the need for examination of actual instances of such mother-child interaction to determine...”).

Additionally, the verb “to explore” could also be considered a collocation in the same structure (e.g. “*These findings need to be further explored*”). This latter example gives evidence that “further” also occurs frequently with “need”. They mainly collocate when “need” is followed by “to be” and when it is used in the third person singular (e.g. “... *needs to be further evaluated*”). In Spanish the adverb “más” could act as an equivalent to “further” in these examples above (e.g. “*Son necesarios más estudios*”).

The adverb “más” and “mayor” collocate with the adjectives “necesario” and “necesaria”, although its meaning is not equivalent to “further” in English. These collocations are mainly used to refer to the need of carrying out a study with a larger sample. The comparative “larger” is expressed in Spanish with “más” and “mayor” (e.g. “*Son necesarios más estudios controlados y con muestras más grandes para obtener más datos sobre...*” or “*se necesitan series más amplias*”). Examples of “mayor” are: “*Son necesarios estudios con un mayor número de pacientes*” or “*Consideramos necesaria la realización de estudios controlados, con muestras de mayor tamaño*”).

The comparison with the adverb “more” is also present in the English corpus combining with “need”. It is equivalent in meaning to “más” in Spanish. However, “more” collocates with the noun “data”, which collocates with the verb “need” too (e.g. “*More data are needed to determine...*”). Although the noun

“dato” appears close to the verb “necesitar” in the corpus, it did not appear as often and as close or immediately after the verb as it was the case in English (e.g. “*Basándonos en los datos de nuestro estudio nos replanteamos la necesidad de realizar un ...*” or “*Son necesarios más estudios, controlados y con muestras más grandes, para obtener más datos sobre...*”). It should also be mentioned that the adverb “more” also combines with other nouns such as “work” or “measures” especially in passive structures when the verb “to need” is in the sentence (e.g. “*More work is needed to ...*”).

Another term that collocates with “necesitar” is “profundizar” (e.g. “*A la vista de las enormes dificultades... sería necesario realizar una revisión en profundidad del...*” or “*se hace necesario seguir profundizando en una... que permitan*”). As the latter example shows, also “permitir” is frequently present in the same sentence as “necesitar” (e.g. “*Además se necesitan estudios longitudinales que nos permitan comprender los procesos...*”). The verb “allow” also appears with “need” in the English corpus, but it is not as frequent as in the Spanish corpus (e.g. “*The need for enhanced diagnostic methods that allow for... may contribute to...*”). The equivalent literal meaning to “profundizar” was not present in the English corpus. However the verbs “remain” and “improve” collocate with “need”. E.g. “*The need for ... remains.*” Sometimes “need” and “remain” do not collocate but they appear very close in the same sentence: e.g. “*Despite encouraging early results, there thus remain many questions and considerations that need to be addressed...*” No literal equivalent was to be found in our

Spanish corpus, where expressions such as “es necesario” or “se hace necesario” would express an alternative equivalent meaning. As for “improve”, it collocates in both languages with “need”. An example in English: “... *further developments to improve... are necessary*” or “*there is a need for improving...*”

In Spanish it appears with the expressions “es necesario/a”, “parece necesario”, and “resulta necesario”, the verb “mejorar” preceded by the preposition “para” (e.g. “*Es necesaria la colaboración multidisciplinar en... para mejorar...*”, “*Parece necesario que otras características clínicas,... han de incluirse para mejorar la validez del test...*” and “*Resulta necesario continuar realizando esfuerzos para mejorar la eficacia de los recursos existentes...*”). The verb “precisar” acts as a synonym of “necesitar” and appears in similar structures in combination with “mejoría” (e.g. “...*es preciso realizar más estudios para demostrar la mejoría de la disfunción...*”).

As a last characteristic of the use of the verb “to need”, it is seldom used in a personal way, with a clear subject preceding it. In fact there are only very few instances where an explicit subject can be recognized (e.g. “*Nos parece necesario*” or “*Necesitamos que*” and “*We need to progress beyond...*”).

Finally, it is relevant the fact that many times the term “need” implies a contrast between what has been stated previously, normally an advantage, and the fact that despite the advantage further research is required. This is why many times sentences containing the verb “to need” start with connectors introducing contrast such as: “A pesar de”, “No obstante”, “Sin

embargo”, etc. The most frequent in the corpus is “no obstante (e.g. “*No obstante son necesarios trabajos semejantes en esta línea*”).

As stated above, the verb “precisar” acts as a synonym of “to need”. The same happens with the verb “requerir”, which often acts as a synonym of “need”. The most repeated structure with “require” in the corpus is: “*Se requieren más estudios que/ para ...*” “Precisar” has a higher rate of appearance. It mainly comes up in structures where the need to continue a study is stated. For this reason it collocates with verbs such as: “proseguir”, “continuar” and “seguir” (e.g. “*Finalmente, precisar la necesidad, dentro de esta línea de investigación, de proseguir con futuros estudios que traten...*”, “*En este sentido, es preciso continuar estudiando el proceso con el fin de*” ... and “*Es preciso, por tanto, seguir investigando para comprobar...*”) but also with the adverb “más” and the comparative “mayor”, which can equally express the same idea (e.g. “*A pesar de ello, es preciso realizar más estudios para demostrar que...*” and “*Se precisa mayor investigación clínica en relación al uso de...*”).

The literal equivalent in English “require” is mainly used in passive structures (e.g. “*Further research is required to confirm*”). As the last example indicates, it collocates with “further research” and “further studies” (e.g. “*Further studies are required to determine the predictive value...*”). When this structure is present, the verbs coming afterwards imply a solution to the stated doubt or problem. For this reason, “require” collocates with “conclusions” preceded by adjectives such as “clear” or “definite” (e.g. “*Further research would be*

required to get clearer conclusions”... and “Studies on larger samples are required to reach some definite conclusions”). But other verbs also occupy the position after the passive structure with the verb “to require” e.g. “to confirm”, “to determine”, “to investigate”, “to evaluate” and “to test” (e.g. “...and long-term follow-up is required to evaluate the outcome”).

4. Remain- Falta/ queda

As already stated above, the verb “to remain” collocates with “need”. It has been pointed out that sometimes it does not appear next to “need” but quite close to it in the sentence (e.g. “*Despite encouraging early results, there thus remain many questions and considerations that need to be addressed*”). The idea that the verb “to remain” conveys (i.e. that there are still things to be investigated and doubts or problems to clear or sort out) explains that this verb collocates with the verb “to be” and with verbs implying “implementation” in the participle such as: “done”, “studied”, “demonstrated”, “seen” and “addressed” (e.g. “*A fair comparison using ... remains to be done*” or “*This problematic issue remains to be studied in a larger number of patients*”). Additionally “remain” also collocates in this sense with the lexical term “challenge” and with the adjective “crucial” (e.g. “... *remains a challenge for future research*” and “*Further investigation of..., including ..., remains crucial*”).

Sometimes, the adverb “still” was found in combination with “remain” to underline the idea that something has to be carried out in the near future but has not been done yet (e.g.

“*One important problem still remains to be solved*” or “*The nature of vulnerability still remains to be elucidated*”). This last example illustrates that “remain” often comes up with the noun “problem” (e.g. “... *remains a challenging problem for neurosurgeons*”).

The idea expressed by “remain” in the examples above can be compared with the Spanish expression “*haría falta/harían falta*” (e.g. “*Harían falta nuevos trabajos de cara a explorar factores de confusión en... y poder corroborar...*”). Logically, this Spanish expression often precedes “*estudios*” and “*trabajos*”. In terms of frequency, the verb “to remain” has a representation of 0.07% in the corpus whereas the expression “*haría/n falta*” in the Spanish corpus has none. This fact evidences that the verb “remain” would have other equivalences in terms of translation in Spanish using the verb “to need” and combining it with different other verbs such as for instance: “*sería necesario llevar a cabo*”. Also the verbs “*precisar*” and “*requerir*” could be used as a possible translation for “remain” in any of the forms shown in the section above (3).

The Spanish verb “*quedar*” is also present in the *Move Further Research* with an equivalent meaning to that of “*faltar*”. (e.g. “*quedan pendientes de estudio*”, “*quedan por resolver*”, etc). This verb clearly overlaps with the *Move Limitations* and was therefore already studied in the mentioned *Move*.

5. Improve- Mejorar

First of all, the family word of “improve” has a rate of appearance of 0.19% in the corpus and its rate as a verb is 0.12%, which indicates the importance of this word. As for “mejorar”, the word family’s representation is 0.08%. Consequently, it is more commonly used in the English language. As already stated in the section dealing with the verb “to need” (number 3), the latter collocates with “improve”. The most common structure is a passive one were “the agent is needed to improve something” (e.g. ... *“more work is needed to improve the coronary”*). Sometimes the verb “to improve” precedes the verb “to be” in the passive structure (e.g. *“... and further software developments to improve the... are needed”*). The most common and similar structure in Spanish is the combination of “parece necesario”, “es necesario/a” and “resulta necesario” followed by the preposition “para” and the infinitive “mejorar” (e.g. *“...resulta necesario continuar realizando esfuerzos para mejorar la eficacia de los recursos existents...”*). Also the combination “es preciso” appears in the same construction, but this time the preposition “para” is followed by the noun “mejoría” and some words are found in between (e.g. *“A pesar de ello, es preciso realizar más estudios para demostrar la mejoría de la disfunción...”*). A different structure which is present in Spanish is the verb “mejorar” with the preposition “a” in front of it (e.g. *“... los esfuerzos futuros deben ir dirigidos a mejorar la identificación de ...”*).

The noun “strategies” is also found close to the verb “to improve” (e.g. “*New strategies to improve local control are needed, as a more...*”). This is also the case in Spanish, where literal equivalents are also present in the corpus (e.g. “*Para mejorar la duración de la LM deberían estudiarse estrategias de información adecuadas...*”). This last example also shows that the verb “deber” often comes up in sentences in which the verb “to improve” is present (e. g. “*El estudio ha revelado áreas que deben ser mejoradas*”). Another noun that collocates with the verb “to improve” is “outcome” (e.g. “*Our and other’s observations support the efforts to investigate additional strategies for... to improve the outcome of patients with...*”)

6. Hope- Esperanza

Although these terms come in both corpora, they are not very widely used to express further research in the corpus of medical conclusions. In Spanish, two adjectives that collocate with “esperar/esperanza” are “futuro” and “próximo”. These two adjectives are also a collocation themselves as the examples illustrate: e.g. “*En el futuro próximo se espera obtener datos importantes sobre...*” or “*Así mismo, estimamos necesaria la continuidad en esta línea de investigación en un futuro próximo*”. This last example shows how the most recurrent words end up being combined among them and repeated over and over again. The adjective “próximo” does not necessarily have to be found together with “futuro” when combined with “esperar” (e.g. “*Todo lo cual, esperamos poder darlo a conocer*”)

en próximas publicaciones”). The last sentence shows the presence of the subject, of the researcher using “(nosotros) esperamos”. This presence of the authors also takes place in the English corpus in which the plural pronoun “we” is clearly placed in front of the verb (e.g. “*We hope to conduct a large-scale study*”).

It is relevant to mention that the noun “study”, as shown in the last example, often comes in the same sentence with “hope” (e.g. “*While the current study leaves these questions unanswered, it will hopefully promote interest...*”). In the last exemplifying sentence in brackets, it is observed that the verb also functions as an adverb, “hopefully”, in this move in English. Although the presence of the personal pronoun also comes up in English with the verb “hope” as stated above, the most common structure with this verb is “the future hope is”: “*Beyond that, the future hope is that it will be used as an off-the-shelf orthotopic biventricular cardiac replacement system that will provide myocardial infarction*”.

7- Understanding- Comprehender

To start with, the representation of the verb “to understand” in the English corpus is equal to that of both “comprehender” and “entender” together in the Spanish one, namely 0.03%. This shows that the rate of appearance is not very significant, but they are used in the move *Further Research* in both languages. The most common form belonging to the word family of the English aforementioned term is

“understanding”, used both as a noun and as a verb after the preposition “for” (e.g. “*Future work to identify ... will be important for understanding molecular mechanisms of heart disease*”). When used as a noun, it collocates with the possessive pronoun “our” (e.g. “*We believe that the increased documentation of ... will add to our understanding of...*”). Only one example of this kind in which the possessive “nuestro” preceded the noun “entender” was present in the Spanish corpus: “*Es necesario por tanto, a nuestro entender, crear o adaptar un instrumento de medida...*”

It is also frequent to find the presence of a modal verb when dealing with the word “understand”, especially “might” (e.g. “*Higher P max and P d in higher anxiety levels compared with lower anxiety levels might help us to understand the anxious predilection of adrenergic AF*” or “*The results of this study suggest that closer attention to... might be fundamental for understanding...*”). Nevertheless other modals such as “can”, “may” and “ought” were also present in the corpus in combination with “understand” (e.g. “*However, the careful assessment of ... may aid our understanding of ... and our knowledge about should be placed on a firmer ground*”). This last example evidences another typical appearance in this context, which is the presence of the verb “to help”, which is a synonym in this case of “to aid”, as shown in the last exemplifying sentence (e.g. “*Studies in groups of people more prone to AF such as the elderly and hypertensive population will help us to further understand the effects of anxiety on event outcomes*”). The appearance of the verb “ayudar” also collocates

in Spanish with “comprender” (e.g. “*Es obligado realizar estudio citogenética que nos ayude a comprender mejor estos tumores y a...*”). The adverb “mejor” after the verb “comprender” can also be considered a collocation.

Regarding modals in Spanish, only an equivalent to “can” was used in the Spanish corpus, namely “poder” (e.g. ... “*el estudio de las diferencias humanas en los procesos de composición escrita puede aportar luces para comprender y mejorar la calidad de los textos de los alumnos...*”).

As it was expected, the next collocation to be mentioned is that of the noun “study”, as illustrated in the example above but also the noun “research” and the verb “to need” come up in combination with “understand” (e.g. “*Longitudinal research is needed to understand clearly the interactions between normal developmental variations in children and the risk factors of mental health problems...*”).

As for the Spanish corpus, no equivalent to “research” was found in combination with “comprender”, but the noun “estudio” also comes up in the same sentence as “comprender” (e.g. “*Además se necesitan estudios longitudinales ... que nos permitan comprender los procesos...*”). Finally, also the adverb “clearly” collocates with the verb “understand” in the texts, as shown in the example above.

8. For/in the future - En el/un futuro

Only one instance of the structure “for the future” was found in the English corpus (e.g. “*Perhaps, prescient, the*”).

description of ... provides the likely clinical scenario for the future”), whereas a few were found for “in the future”. This structure collocates with the word family of the lexical item “promise”; as an adjective (e.g. “... *but results of cytogenetic studies are promising, and we continue to hope that in the future, tailored postoperative treatments will be available on a molecular basis*”) and as a noun (e.g. “*The promise in the future of smaller, more durable pumps with no percutaneous connections will further expand the indications for these devices for a broad spectrum of HF patients*”). As already stated above, “en el futuro” and “en el futuro próximo” comes up with the verb “esperar” (vid. Section number 6) (e.g. “*En el futuro próximo se espera obtener datos importantes sobre...*”). In the same section, it was observed that in English “future” collocates with “hope” coming either next to it (e.g. “*The future hope is that...*”) or a bit more distant (e.g. “...*and we continue to hope that in the future...*”).

A similar structure using the noun “future” also present in this move is that in which the adverb “near” is sometimes found collocating with “future” forming “in the near future” (e.g. “*STE will most likely become a choice indication for... in the near future*”).

Logically, most examples in English with this expression use the auxiliary verb “will” to make future tenses (e.g. “*The classification of SCMs will continue to be a matter of debate in the future*”). However, on one occasion in the corpus a different tense is used (e.g. “*Though technical refinements and ... are needed ... has the potential to be an invaluable diagnostic*”).

modality in the near future”). As for the tenses in Spanish with “en el/un futuro” much variation takes place. The future tense is present (e.g. “*Una revisión desde estos mismos planteamientos en el futuro podrá confirmar o no las tendencias...*”), but also the present and conditional (e.g. “*En el futuro próximo se espera obtener...*” or “*Todas estas consideraciones se centran en la...pero también debería valorarse en un futuro la repercusión...*”)

9. Interesting- Interesante

These two words are not very often used in the corpora, but they should be mentioned since they appeared in the *Move Further Research* in both languages representing 0.01% in Spanish and 0.02% in English. The most recurrent structure in Spanish is the one formed with the conditional form of the verb “to be”, namely “sería”, plus the adjective and then followed by an infinitive (e.g. “*Por otra parte, sería interesante comprobar la relación entre...*”). The verbs that come up after “sería interesante” are associated with the performance of another paper, like, “comprobar”, “estudiar” “plantearse”, “realizar” and especially “ampliar” and “replicar”, which appear to collocate with the adjective “interesante” (e.g. “*Sería interesante replicar el estudio ampliando...*”).

A variation of “sería interesante” also found in the corpus is “se considera interesante” (e.g. “*En cualquier caso, se considera interesante ampliar la duración de las sesiones para futuros grupos*”). An example with the conditional was also

found in the English corpus (e.g. “..., *it would be interesting for future studies to combine a within-subject design using different behavioural tasks with electrophysiology and functional imaging methods.*”) As this instance shows, the noun “study” collocates with “interesting”. Sometimes, it does not come immediately afterwards, but it is a bit more distant in the sentence: “*In this regard, an interesting suggestion comes from the obtained results: it is possible that further studies will demonstrate that...*” In this structure, the noun “study” is accompanied by the typical adjectives collocating with “study”, namely, “further” and “future”. The close relationship between these two terms is equivalent in the Spanish language, where either the infinitive form “estudiar” or the noun “estudio” combine with the mentioned structure “sería interesante” (e.g. “*Sería interesante replicar el estudio ampliando...*”). Other nouns also present in sentences in which the adjective “interesante” comes up are “trabajos”, “muestra” and “investigaciones” (e.g. “*En nuestra opinion, en futuros trabajos sería interesante tener en cuenta...*”).

A similar structure in both languages using the combination of “investigations” and the adjective “interesting” is one where the nouns “route” and “camino” are used in a very similar way in this move (e.g. “*Another interesting route to investigate which ... may modify ...*” and “*Queda abierto, pues, un interesante camino para futuras investigaciones*”). Both in English and Spanish the adjectives “interesting”/ “interesante” affect the nouns “route” and “camino”, but also in both cases the verb and noun “investigate” and “investigation” come

immediately afterwards respectively. Finally, it should be mentioned that the adjective “interesante” is sometimes introduced by “lo” in structures like: “Lo que sería interesante” and, “lo más interesante sería ...”

10. Still- Todavía

The first difference between these two adverbs is their rate of appearance. “Still” has a higher rate than “todavía”: 0.02% and 0.05%, respectively. The adverb “todavía” is commonly found to emphasize the idea that investigations are at the very beginning. This is possible thanks to the adjective “temprano” or expressions such as “todavía nos encontramos en los albores” (e.g. “*Todavía es temprano para poder establecer..., debido a muchas incógnitas que quedan por resolver y a los interrogantes suscitados en temas...*” or “*Todavía nos encontramos en los albores de este nuevo concepto... pero el conocimiento...ha abierto nuevas expectativas y nuevas vías de investigación*”).

The verb “quedar” in the first example is comparable to the verb “to remain”, which collocates with the adverb “still”. In fact, it can be stated that “still” and “remain” collocate, although their position in the sentence is not close. Normally when they appear together the verb “to remain” is found in a passive structure: e.g. “*However, CPC still has an extremely poor prognosis and the efficacy of ... remains to be established*”. Regarding the second example (e.g. “*Todavía nos encontramos en los albores de este nuevo concepto...*”), it can also be

considered equivalent in terms of meaning to the English version “... *is still in its infancy and its role in paediatrics is far from being defined*”.

This last example brings up another collocation, similar in meaning: the combination of the verb “to be” with the time expression “to be far from” (e.g. “*We are still very far from accomplishing this*”). A last characteristic to be mentioned is the fact that most terms coming after the adverb “still” have a negative meaning: “critical”, “limited”, “poor”, “hindered”, etc. (e.g. “*However, CPC still has an extremely poor prognosis and the efficacy of ... remains to be established*”). Sometimes items with a negative meaning are found in front of the adverb like “confusion” (e.g. “... *much confusion still exists with respect to...*”). The most typical is “problem”, which might come directly before or slightly more distant (e.g. “...*, but the problem of genetic... is still far from being disentangled*” and “*One important problem still remains to be solved*”). This is the first evidence that there is a clear overlap and relation between the *Move Further Research* and *Limitation*. In fact, the first common linguistic field in the *Move Limitation* has to do with time adverbs and “still” is one of the most recurrent ones (see p. 326). Finally, as expected, the verb “to need” is also present many times in sentences containing the adverb “still” (e.g. “...*, there are still a number of critical areas in need of exploration*”). Although “necesitar” is also frequently present in many structures in Spanish, no instances were found in the corpus in combination with “todavía” probably due to the low rate of appearance of this Spanish adverb.

11. Sample- Muestra

The first difference is that the noun “muestra” appears twice as frequently as “sample” in the corpus, namely 0.16 % and 0.07%, respectively. Additionally, this noun is a clear evidence of the aforementioned overlap between *Further Research* and *Limitations* mentioned in the section above. When researchers make use of these words in the aforementioned moves, it is mainly because they suggest the results of the study are not those expected or they could be improved by means of a greater sample. As saying that implies an intrinsic limitation, in their suggestion the overlap takes place.

A similar structure present in both corpora representing this would be: “*This study is based on a small sample; therefore the results need to be replicated with a larger group of patients*” and “*Cuando se disponga de una muestra mayor sera conveniente replicar el análisis factorial del estudio*”. As the examples show, “replicar” and “replicate” are commonly used with the noun “sample”. The main two verbs that collocate with “sample” in Spanish in this move are “aumentar” and “ampliar” (e.g. “*Las acciones futuras se dirigirán a ampliar la muestra*”).

Unlike in Spanish, in English it is not the verbs that collocate with “sample” in *Further Research*, but the comparative adjective “larger” (e.g. “*Further research should be directed toward the evaluation of... and should confirm these results in a larger sample group...*”). Sometimes the adjective is

not in the comparative form: “... *ought to be considered and cross-culturally studied in large samples in an effort to understand...*” In Spanish “muestra” collocates with a comparative adjective as well, as shown above (e.g. “muestra mayor”). However, there are other adjectives that collocate with “muestra” that come up after the adverb “más” such as: “representativa” (e.g. “...*sería adecuado disponer de una muestra más representativa de la población general...*”).

The adverb “más” also appears with the adjective “amplias” (e.g. “... *aunque deberían realizarse estudios controlados con muestras más amplias*”). Also the noun “tamaño” collocates with “muestra” in this move (e.g. “*Aunque estos Buenos resultados deben ser tomados con precaución por el tamaño de la muestra, nos anima a estimular a los profesionales de AP a la realización de...*”). It is frequently specified that the “sample” represents “patients”, this redundancy turns into a collocation as well (e.g. “*Estos hechos son aspectos a mejorar en futuros estudios que a su vez deberían aumentar la muestra de pacientes incluidos*”). Finally, it is also noteworthy the presence of the noun “investigation” and the adjective “future” in the same sentence as “muestra” or very close to the noun, either before or after it (e.g. “*Varios aspectos nos parecen importantes para una futura investigación: 1. Ampliar la muestra de los subgrupos de pacientes...*” and “... *concluimos que es importante medir..., lo que señalamos como posible línea de investigación futura*”). In these two sentences, the adjective “importante” is also used in both of them. At this stage it can be affirmed that the same

reduced circle of words is combined over and over again in the move.

12. Elucidate- Dilucidar

This verb is more frequent in English than in Spanish, according to the instances found in the corpora. The main use of this verb in Spanish is “dilucidar” in combination with the noun “cuestión” (e.g. “*Qué intervenciones son eficaces... y cuáles deben ser sus componentes básicos es una cuestión a dilucidar en estudios controlados*”). This collocation is also present in the English corpus (e.g. “*Thus, subsequent investigations of women athletes should include extensive questions about all of these areas, to better elucidate what appears to be a largely unrecognized public health problem*”). Sometimes, the noun “question” is placed in the distance but the link between “elucidate” and “question” can be easily identified. The next example shows the aforementioned words but with another item belonging to the same word family as “elucidate”, i.e. with the noun “elucidation”: “*Despite encouraging early results, there thus remain many questions and considerations that need to be addressed before therapeutic angiogenesis becomes a new option...These issues include..., elucidation of the stability of newly formed vessels...*” With this example, the link between “to elucidate” and the verb “to remain” is also evident. Another example proving this relation is “*The nature of vulnerability still remains to be elucidated.*”

Other closely related words to the term “elucidate” that should be mentioned are “study” and “need” (e.g. “*Thus, more experimental embryological studies are needed to elucidate the unified theory.*”). Finally, one more example is going to be included that evidences that all the items contained in the list of equivalences between the two languages pointed out at the beginning of this chapter are combined and that, actually, in this move the amount of vocabulary to express that more research should be carried out is limited to the list (e.g. “*As the number of patients in this series is small, further randomized studies are necessary to elucidate the benefits of primary excision...*”).

As far as modal verbs are concerned, it is relevant to mention that the most outstanding in this move in Spanish is “deber”. Regarding English, the most recurrent is “should”, followed by “may”, “can” and “could” and “must” in this order.

The last term of the list containing 12 items equivalent in the two languages shows that sometimes words are much widely used in one language than in another (e.g. “to elucidate”). Many times there are words which are non-existent in one language, but are used in the other. For this reason, recurrent terms only present in one of the two languages in the *Move Further Research* are shown in the list below:

a) Habrá que- No ha hecho más que empezar

The expression “no ha hecho más que empezar” in combination with “hay que” or the future form “habrá que” is very often used (e.g. “*La terapia... no ha hecho más que*”).

empezar. Habrá que demostrar su eficacia...). These two expressions are also present separately. The second, “habrá que” can be considered to collocate with the verb “demostrar”, as already shown in the example above. An alternative for “habrá que” in the corpus is “habrá de”, which appears in sentences containing the verb “permitir” (e.g. “*Por otra parte, habrían de realizarse estudios epidemiológicos, ... que nos permitieran...*”). After both expressions: “habría que”/ *habrían de*” the idea that something should be done more thoroughly is conveyed by means of structures such as: “definir con mayor claridad”, “estudiar con más detalle”, “realizar otros estudios más exhaustivos”, etc. (e.g. “... *para ellos habrían de realizar otros estudios más exhaustivos para valorar cómo influyen estos factores...*”). As it is observed in this last example, another word which easily collocates with “habrá que” is “otros” to imply that other studies, variables, etc, should be further studied.

b) Ser conveniente

An adjective that frequently occurs in the Spanish text is “conveniente” preceded by the verb “ser”. It co-occurs with the items analyzed in the lists of equivalences between the two languages. For instance, it appears with a verb implying the need to continue investigating: “*Sin embargo, dada la tendencia de los resultados se considera conveniente seguir trabajando en la misma línea*”, or with the idea that research should be “replicated” with “other samples” : “*Convendría, cuanto antes, replicarlos con otras muestras*” and with the collocation

“futuros trabajos”: “*Otro aspecto que se debe tener en cuenta es el de la conveniencia de incluir en futuros trabajos un grupo control para comparar...*”. This combination did not have a literal equivalent in the English corpus.

Another structure present in the Spanish corpus is “ser susceptible de” followed by participles. Again, other items present in the analysis above such as “nuevos trabajos” or “para mejorar” take place in the same sentence. (e.g. “*El test de la homofobia, que aunque es susceptible de ser estudiado y testado en nuevos trabajos, incluso ser modificado para mejorar su utilidad clínica...*”).

c) Sería deseable

The presence of the verb “to be” in combination with the adjective “deseable”, or with the noun “menester” is also relevant in this move, especially in the conditional tense in structures like “sería deseable” or “sería menester” (e.g. “*Sería deseable que para próximas ediciones fuera considerado, al menos como patología a estudiar y profundizar*” and “*Por último, sería menester discriminar aquellas técnicas y elementos factibles de ser...*”).

d) Modals: “deber”

Finally, it is important to mention that “deber” would constitute the first modal verb in terms of recurrence in the Spanish corpus in the *Move* Furhter Research.

e) Prospective

The adjective “prospective” is representative of this move with a percentage of appearance of 0.04%. As mentioned in section number one of this move this adjective and other similar ones like “further” and “future” collocate with “study”. In fact, its appearance is mainly in passive sentences combining with the verb “to need” (e.g. “*Further prospective studies are needed in this understudied field...*”). As it is shown in the example, “prospective” also collocates with the adjective “larger” and the adverb “further” on many occasions, the latter adverb presented with a split infinitive (e.g. “*Larger prospective studies are needed to further assess this therapeutic approach.*”)

The collocation “prospective study” is the most recurrent regarding the adjective, however, other nouns that also collocate with the adjective “prospective” are “trial” and “effort” (e.g. “*Further prospective trials may be useful to standardize treatment*”). This collocation is sometimes reinforced by an additional participle constituting a 3-element collocation, namely, “prospective randomized trial” (e.g. “*There is a need for a prospective randomized trial to assess...*”).

The collocation “prospective effort” also appears frequently as a part of a 3-element collocation with the presence of “further” giving rise to “further prospective efforts” (e.g. “*Further prospective efforts remain needed...*”).

Finally, an example with “data” was also found in the corpus (e.g. “*Prospective data are needed to clearly...*”).

f) To warrant

This verb is widely used in this move in English. It was already mentioned in the section “study”, as it collocates with it to imply that more research will be done. This same idea can be conveyed when it collocates with other nouns like “investigation”, “trial” and “finding” (e.g. “*Whether patients can obtain better outcomes... warrants future investigation*”, “*A prospective, ideally multicenter randomized trial is warranted to determine whether the finding of the present study...*”). The noun “finding” present in the last exemplifying sentence, is not an isolated case. These two terms co-occur in other instances (e.g. “*The unexpected finding of a consistent increase in... warrants further study*”). In general, it can be affirmed that the verb “to warrant” is more commonly used in a passive structure (e.g. “*Larger scale studies are warranted to further characterize...*”).

g) Evaluation

The noun “evaluación” has a high rate of appearance in Spanish; in fact it is higher than in English. In Spanish it occurs 0.06% in the corpus, whereas in English the representation of “evaluation” in the corpus is 0.05%. In Spanish it is mainly used forming the compound “evaluación-intervención”, and it also

collocates with the adjective “diagnóstica” and “funcional”. However, “evaluación” is not present in the *Move Further Research*.

By contrast, it is present in English in combination with the verb “to need” and other synonyms such as “to require” or “to await” (e.g. “*This strategy awaits further evaluation*”). Sometimes, a different grammatical category is used instead of the verb “to need”: “*Such questions imply the necessity of process evaluation to better delineate...*”)

Occasionally, other categories belonging to the same word family as “evaluation” are used in this move. The next sentence underlines by means of the adverbs “more accurately” placed after the verb “to evaluate” the idea that new techniques are necessary: “*Therefore, neuropsychological findings should be supported by neuroimaging techniques to evaluate more accurately cognitive dysfunctions in these patients*”.

h) Findings

Although the noun “hallazgos” also has a representation of 0.02% in the corpus, it is not used in the move *Further Research*. Its appearance is to be found for instance in the move *Background*, since it is very commonly used to compare the results of the paper with that of other authors (e.g. “*Estos hallazgos son concordantes con...*”). In English, its rate of appearance is very high, (0.18%), thus, it is not surprising that it comes up in the move *Further Research*. In fact, its appearance with the words “evaluation” and “to warrant” has already been

mentioned above. This word appears both in singular (vid. examples for “to warrant” and “evaluation” above) and plural and many times preceded by the demonstrative “these” (e.g. “*These findings need to be further explored in different settings*”). It is frequently embedded in a passive structure as shown in the last exemplifying sentence and in the following sentence: “*Therefore, neuropsychological findings should be supported by neuroimaging techniques to evaluate more accurately...*”

i) Modals

As far as modal verbs are concerned, their use in this move is higher in English than in Spanish. The most widely used in a descendent order are: “should”, “may”, “can” and “could” and “must”. In this move, the most recurrent modal verb coincides in both languages being “should” and “deber”.

3.3.3 Neologisms in the Spanish medical corpus analyzed

The impoverishment denounced by several authors (Iscla; Aleixandre, 2003) of the Spanish medical language caused by different factors is also reflected in the Spanish corpus of this research work. From the many areas covered by different linguistic experts in relation with the influence of neologisms on the Spanish medical language and regarding the influence of English on Spanish and the consequent translation problems in the last years, most aspects have been registered in the present

study. Although most areas were analyzed at the semantic level, special attention has been devoted to “derivation” and “loanwords”, explained according to Haugen’s nomenclature in 1950. Additionally, other aspects not mentioned in the literature above but present in the study have been analyzed e.g. “compounds” and “hyphenated words”. At the syntactic level, all the aspects were registered in the corpus: “copying syntactic structures from other languages”, “abusive use of the passive and gerund” and “ellipsis”.

- In the field of semantics:

a) Words from other languages adapted to the Spanish spelling

The appearances of adapted terms from other languages (mainly from English) that have been accepted by the Real Academia are registered in the corpus. Some instances are the terms “estrés” “estándar”, “láser”, “déficit”, “ítems” and “hobbies”.

b) English words in the Spanish corpus (Calques or loans)

The incorporation of words from other languages without morphological adaptation denounced by several Spanish linguists (Martínez, 2006; Aleixandre, 2003), especially when a Spanish equivalent exists, takes place in the Spanish corpus. For instance, Navarro (1992:577) states that “feed-back” is not

admitted, it is what Haugen (1950) calls a “substitution”. In fact, Navarro adds that it can be replaced by “retroalimentación” in some cases as in the example of the corpus “*Lo más destacable es el hecho de que esta evolución positiva que se da en la capacidad de interacción de los pacientes se relacione estrechamente con el avance que observamos en los terapeutas, que hicieron servir como feed-back el análisis de las sesiones anteriores.*” Rodríguez (1997:207) adds other possible translations for the word like “retroacción” and “retroalimentación”. Some other examples of English words in the corpus are: “Problem solving”, “Life events”, “Coping” used in “estrategias de coping”, “test-retest”, etc. (e.g. “*Los especialistas reafirman la importancia de estos life events que ocupan uno de los primeros lugares en el rango de impacto de riesgo*” or “*Es necesario que en futuros trabajos se estudien la factibilidad, la fiabilidad en términos de consistencia interna y estabilidad test-retest, y la validez*”).

The noun “screening” is present in the corpus especially in combination with the noun “instrumento”, when the equivalent “cribado” should be used instead (e.g. “*los síntomas que en concreto se seleccionen para confeccionar un instrumento de screening del TS no parecen ser relevantes para que el test resulte válido,...*”). Other suggestions to translate “screening” into Spanish according to Navarro (1994:147) are “selección sistemática” or “detección sistemática”. Sometimes, the neologism is a compound made of an existing term and an English one, for instance: “inmortales like” (e. g. “*...con el consecuente bloqueo de la apoptosis, dejando en*

funcionamiento clones celulares inmortales- like [en queratinocitos], proceso que...”).

Considering that one of the disciplines in the corpus is cardiology, it is not surprising that the anglicism “stent” appears repeatedly in the corpus (e.g. “*Con el tipo de stent utilizado ha sido posible intentar el implante directo en todos los casos...*”). According to Navarro (1994:513), the most suitable translation for this substitution is “endoprótesis vascular” and prefers other possibilities such as “resorte intravascular”, “tutor intravascular” or “tubo expansible” rather than the anglicisms “stent” and “estén”.

c) **Derivation**

While reading the Spanish corpora a list was created containing words that are not registered in the Spanish dictionaries at the present time. The dictionaries consulted were the D.R.A.E. (2001), Diccionario de la Real Academia Española, and Diccionario abreviado del español actual (2000). The authors of the journals created a significant number of previously unattested words using different strategies. The most common method to create new words used in the Spanish medical corpus analyzed is the strategy of derivation by means of the addition of different suffixes and prefixes to existing words.

a) Use of existing prefixes to form neologisms. According to Gómez (2002:19), prefixes do not change the grammatical

category of words but their function is precisely to give a new meaning to the word they are added to. In the corpus the following prefixes were identified:

- Creation of new adjectives:

- The prefix “in-”, usually conveys a negative meaning to the adjective with which it is combined and turns it to mean the opposite. According to Gómez (2002:20) the prefix “in” means “no”. In the DRAE (2001) the same meaning for the suffix is given (e.g. “negation” or “deprivation”). An instance is the adjective “inalcanzable” that means something cannot be reached. However, in the DRAE an additional meaning for the suffix “-in” is “inside” (e.g. “incluir”). In the corpus, this prefix is applied with the first meaning indicated, “negation”, to adjectives that do not admit it such as “inespecífico” or “inefectiva”. It is difficult to know whether “inespecífico” is simply a new word coined by the author or whether its origin is related to the English “unspecified”.

- Another prefix used to form adjectives in the corpus is “multi-”. It is used in Spanish to mean “many” (DRAE, 2001) and it is applied to words like “factorial” giving rise to the up to now unattested adjective “multifactorial”.

- The prefix “hetero-” used to denote “different”, “unequal”, “another” (e.g. DRAE, 2001) like in “heterosexual” is applied to the participle “aplicada” giving as a result “heteroaplicada”.

- The prefix “uni-” meaning “one” (e.g. “unidireccional”) generates terms like “unimodal”.

- Creation of new verbs and nouns:

- Many prefixes can be used to create new verbs and nouns. For instance, the prefix “sobre-”, which in Spanish conveys intensification of the noun it precedes (e.g. sobrealimentación). Other meanings registered in the DRAE (2001) are “repetition” (e.g. “sobrecenar”), “sudden action” (e.g. “sobrecoger”) and “addition” or “superposition” (e.g. “sobrecalzada”). For example, the prefix can be applied to the adjective “valorado” but also to the verb “valorar” resulting in the adjective “sobrevalorado” and the verb “sobrevalorar”, respectively. The application of this prefix is with the meaning of “intensification” is applied to other nouns creating neologisms such as: “sobrepuntuación” and “sobrerepresentación”.

- Similarly, the prefix “infra-” can be considered an antonym of “sobre-” in the sense that according to the DRAE (2001) the meaning it conveys to the words it is added to is “inferior”, “under” (e.g. “infrahumano”). For instance in “infravalorar” which means “undervalue” or “underestimate”. In the corpus, it is applied to the noun “tratamiento” creating the new term “infratratamiento”.

- The prefix “sub-”, which is similar in meaning to “infra-” also conveys the idea of “under” (DRAE, 2001) as in

“subsuelo” or “subestimar”. The DRAE (2001) registers other occasional meanings for this prefix such as “inferiority”, “diminish” and “decrease”. In the corpus it generated previously unrecorded terms such as “subtest” and “subescala” where the suffix’ meaning is “under”.

- The prefix “co-” is registered in the DRAE (2001) with the meaning of “cooperation” (e.g. “confluir”). It is applied to the neologism “responsabilización”, giving rise to the previously unattested noun “corresponsabilización”.

- According to the DRAE (2001), the prefix “inter-” conveys the meanings: “between” or “among” and “in the middle” (e.g. “intercostal”). In the corpus it generates nouns like “intersujetos”.

- The prefix “hiper-” gives an intensifying meaning to the word it is added to (Gómez, 2002:20); thus, it provides the word with a meaning of “superiority” or “excess” as in “hipertensión” (DRAE, 2001). This prefix gives rise to nouns like “hipervigilante”.

- The prefix “intra-”, meaning “inside-” (DRAE, 2001) as in “intravenoso”, is used together with the noun “grupo” and results in the noun “intragrupo”.

- Sometimes the same prefix is added to different grammatical categories like “-auto” and “-re”.

- The prefix “auto-” meaning “own” (DRAE, 2001), e.g. “autobiografía”, is added to words that have different morphological categories such as the noun “autoinforme”, the verb “autodirigir” or the adjective “autoseleccionados”, forming words that are not registered in the dictionary.

- The prefix “re-” can provide the word to which it is added with different meanings according to Gómez (2002:19) and to the DRAE (2001). It can be used to intensify (e.g. “relimpio”) but it can also mean “backwards” (e.g. “reular”, “refluir”) and it is frequently used to indicate repetition (e.g. “releer”). This latter meaning is the one registered in the medical corpus and it is added mainly to nouns but also to adjectives, creating the neologisms “reintroducción”, “rediagnosticados”, “redefinición” or “reactualización”.

Although normally prefixes do not change the grammatical category of the word, Gómez (2002:19) indicates that some modern prefixes such as “post-” but also “multi-“ and “anti-“ do turn nouns into adjectives.

- The prefixes “pre-“and “post-”, indicating normally “before/earlier” and “after/later”, respectively (DRAE, 2001). They are added to words with or without a hyphen creating the neologisms “pre-tratamiento”, “pre-adulto”, “presuicida”, “post-tratamiento”, “postinfarto”, “postalta”. The original nouns “tratamiento”, “adulto”, “suicida”, “infarto” and “alta” in the examples stated above turn into adjectives in some instances of

the corpus (e.g. “angina postinfarto” or “revascularización postalta”).

b) Regarding the use of suffixes to create new words that are not registered in the dictionary, Gómez (2002:20) distinguishes between two kinds of suffixes: on the one hand, those comparable to prefixes because they give a word a new meaning and on the other hand those that only have a grammatical meaning. An example for the latter group is the suffix “-ndo” to convert the infinitive “amar” into the gerund form “amando”. However, it is the first group, namely suffixes that convey a new meaning to the word, that are interesting for the analysis of neologisms. These suffixes provide the previous word with a concept different from the original one (Gómez, 2002:21). There are two kinds of suffixes with non-grammatical meaning:

- Those that provide a nuance to the word in terms of affectivity from the point of view of the speaker. Gómez (2002:21-23) differentiates 5 kinds of these suffixes: diminutives, augmentatives, despectives, superlatives, and familiar. To illustrate this with an example, diminutive suffixes can express affect on the part of the speaker (e.g. “abuelita” for “abuela”, “grandmother”, corresponding to English “granny”). Gómez (2002:23) calls these kind of suffixes “potestativos” or “apreciativos”.

- The second type of suffixes is referred to as “obligatorios” (Gómez, 2002:23). In this group of suffixes a change of the grammatical category of the word often takes

place. The second relevant characteristic of these suffixes is the many different meanings they can give to the words. Gómez (2002:23) makes a list to collect them according to the meaning they have under different headings: suffixes indicating origin (e.g. “-ano” like in “asturiano”), place (e.g. “-era” like in “gasolinera”), action or effect (e.g. “-dura” like in “mordedura”), profession (e.g. “-ista” like in “modista”), agent or actor (e.g. “-in” like in “bailarín”), quality (e.g. “-tud” like in “amplitud”), modality (e.g. “-mente” like in “fácilmente”), relation or ownership (e.g. “-in” like in “infantil”), etc.

Neologisms registered in the Spanish corpus are made of the so called “obligatorios” suffixes. Words change their grammatical category by the addition of the mentioned suffixes and give rise to previously unattested terms:

- Creation of new adjectives:

- The use of the suffix “-able” is normally used in Spanish to create a new verbal adjective that indicates that something is possible, namely a passive possibility (DRAE, 2001). For instance, the adjective “realizable” that means that something can be done. This suffix is applied to other words, creating non-registered adjectives like “diferenciable” or “objetivable”.

- Another suffix used to create non recorded adjectives is “-al”, which is normally added to mean “group” as in “instrumental”, meaning “group of instruments” (Gómez, 2002:23). This suffix can also indicate “relationship” or

“ownership” according to the DRAE, 2001. In the corpus the second meaning of the suffix is the one recognized in the adjective “atencional” (e.g. “*Asimismo, se investigan la integración perceptiva en el tiempo debida a la atención, la evolución diacrónica del sesgo atencional o la naturaleza de la orientación atencional en el tiempo...*”). Similarly, in the previously unattested adjective “comportamental”, the meaning of the suffix implying “group” is not recognized (*Entonces y tal como señalábamos en la introducción las personas liberales son más erotofílicas y las conservadoras más erotofóbicas, aunque esto no tenga porque ser siempre así, dado que los componentes de las actitudes (cognoscitivo, afectivo y comportamental)...*). Thus, taking the example into account the adjective “comportamental” seems to be a loan translation of the English term “behavioral” rather than to only a case of derivation.

- The suffix “-ador” normally refers to someone who carries out an action, i.e. to the “agent” according to the DRAE (2001). For instance, “cazador”, from the verb “cazar”, corresponding to “hunter” and to the verb “to hunt” in English. It is applied to the verb “optimizar” in the corpus, originating the adjective “optimizador”. Other meanings of the suffix registered in the DRAE (2001) are “device” (e.g. “acelerador”), place (e.g. “comedor”), but none of the above meanings of the suffix can be recognized in “optimizador”.

- The suffix “-ado” belongs to the category expressing “collectivity or group” according to Gomez’s classification (2002:23). For instance, the adjective “profesorado” means “a

group of teachers”. Apart from this meaning, other are attributed to this suffix in the DRAE: it can express “similarity” like in “aterciopelado”, it can also form nouns that indicate “action” and “effect” like “afeitado” or alternatively the suffix indicates a position like in “obispado”. In the corpus the adjective “mistificado” seems more a case of what Gómez (2002:21) calls a “postativo or apreciativo” suffix with no grammatical change of the word, however, no nuance different to the adjective “místico” can be appreciated in “mistificado” (e.g. “*En tal priorización de necesidades, no se ha de obviar la raigambre sociocultural del mistificado fenómeno de las drogodependencias, ni los efectos derivados de ...*”); thus, it might be a loan translation of the English participle “mystified”. As for the adjective “aleatorizado”, the same unjustified use of the suffix takes place; the appearance of this adjective in the medical corpus can only be associated with a loan translation of the English word “randomized” (e.g. “*No se trata de un estudio verdaderamente aleatorizado.*”).

- The suffix “-oide”, sometimes also “-oides” stems from the Greek and it conveys the meaning of “en forma de” and “parecido a” meaning “similar to” (DRAE, 2001), for instance the adjective “androide” means “similar to a man”. This suffix is added to the noun “cannabis” in the corpus originating “cannabinoides”, that should be understood as “similar to cannabis”.

- The suffix “-ista” is conferred the meaning of “partidario de” in the DRAE (2001) meaning “supporter” or “fond of” but as pointed by Gómez (2002:24) it is also used to

indicate profession. For instance “modista” is added to “moda”, “fashion” and is used to refer to a “dressmaker”. The neologism registered in the corpus is “salubrista”, in which the first meaning “supporter” can be recognized. Thus, “salubrista” is the adjective referring to someone who is a “supporter of health”, in Spanish “partidario de la salud” (e.g. “... *en aras de promover actitudes y conductas optimizadoras vinculadas a una política salubrista, reivindicación necesaria a todos los efectos*”).

- Creation of new adverbs:

- The suffix “-mente” normally turns adjectives into adverbs (DRAE, 2001) (e.g. “rápido” > “rápidamente”). Gómez (2002:25) states that the meaning this suffix conveys has to do with “modality”. In the corpus this meaning of the suffix is recognized when it is added to adjectives that do not admit it such as “admitidamente”, “espuriamente” “comprehensivamente” or “artefactualmente”, “transculturalmente”. The following example illustrates the use of the existing adverb “admittedly” in English: “*Additionally, the number of patients studied was limited and, admittedly, the specific correlations that we have found are only indications*” versus the neologism “admitidamente” in Spanish: “*Admitidamente, nuestra muestra es limitada y sesgada, se trata de parejas que fueron derivadas a un centro...*”

- Creation of new nouns:

- Creation of abstract nouns by means of the suffix “-dad” (and its forms “-idad”, “-edad”). Sometimes a suffix can convey more than one meaning depending on the word it accompanies. The suffix “-idad” belongs to the group expressing “action or effect” (e.g. “caducidad”) in Gómez’s classification (2002:23-25) but also to the group of suffixes expressing “a quality or trait” (DRAE, 2001) (e.g. “tenacidad”). The meaning of the suffix can be recognized in the instances of the corpus neologisms “comparabilidad”, “sugestibilidad” and “evitabilidad”.

- The suffix “-miento” is present in Gómez’s classification and it is included in the group of suffixes indicating “action or effect”. As an example the noun “pensamiento” is given, since it is an effect or consequence of the noun “pensar”, “to think”. This suffix is used in the corpus to form previously unattested nouns like “modelamiento”, formed on the basis of “modelar” which seems to be a loan translation of “modelling” (e.g. “*Ellos también concluyen que la influencia social está presente por medio del modelamiento, y de una persuasión directa a través de la presión para satisfacer las demandas de otras personas o del entorno social.*”)

- Another suffix that is included in the group of suffixes expressing “action” or effect in Gómez’s classification (2002:23-25) is “-ción”. In the corpus the noun “conceptualización” appears, in which the suffix meaning can be

identified. Thus the noun “conceptualización” would be the action or effect of the verb “to conceptualize”.

- Creation of new verbs

- Although less frequent in the corpus, unattested verbs formed by means of verbal suffixes were identified (such as “-izar”, e.g. “hipotetizar”). It is very likely that the verb “hipotetizar” is modelled on the English verb “to hypothesize” as a consequence of the inexistence of a Spanish verb. In Spanish it would be necessary to use the expression “hacer una hipótesis” to express the same meaning. This is another example of the difficulty that the Spanish language has to create short terms equivalent to the English, as mentioned by Ballester (2003:4).

d) False friends in the corpus

Some examples of neologisms that are a probable consequence of the spread of false friends in the Spanish language when carrying out translations are present in the corpus. For instance “tableta” is used instead of “comprimido” based on the English term “tablet”, which is the abbreviation of “compressed tablet” (Navarro, 2000:985), (e.g. “*Por ejemplo, prescribir la mitad de una tableta de sertralina de 100 mg al día es la mitad de caro que prescribir una tableta de sertralina completa de 50 mg al día.*”). The noun “evidencia” is very commonly used with the same meaning as “evidence” in

English, rather than with the Spanish meaning of “*certeza clara ... que nadie puede dudar de ella*” (see Navarro, 2005 above).

According to the Collins Cobuild English Language Dictionary (1987) “evidence” is “anything that you see, experience, read, or are told that causes you to believe that something is true or has really happened” (e.g. “*Este resultado es consonante con la evidencia obtenida en el ámbito de estudio del estrés familiar que subraya la estrecha relación e interdependencia existente entre el funcionamiento familiar y la presencia de estresores normativos y no normativos...*” or “*hay que tener en cuenta que los artículos revisados generan poca evidencia científica ya que se tratan...*”). It is important to mention that the translation for the noun “evidence” in the Collins Spanish Dictionary is “evidencia”, which makes the misuse of the word more understanding.

Navarro (2002:297) provides us with multiple options for the translation of the term “dramatic” into Spanish depending on the context: “impresionante”, “espectacular”, “considerable”, “sensacional”, “gravísimo”, “claro”, “intenso”, “tajante”, “sorprendente”, “imponente”, “súbito”, “descomunal”, etc. The adjective “dramática” is present in the corpus and has been probably used instead of “considerable”, “impresionante” or “espectacular” possibly as a semantic loan based on the English adjective “dramatic” (e.g. “*La escasez de donantes pequeños limita de forma dramática la posibilidad de conseguir injertos para candidatos de bajo peso lo cual hace necesario la búsqueda de alternativas quirúrgicas*”).

The presence of the adjective “severo” in the corpus seems to be a misuse as a consequence of influence of the English “severe” often translated as “severo” rather than as “serious”, in Spanish “serio” o “grave” (e.g. “*A pesar de que los TP inicialmente presentan un estado psicopatológico más severo el grado de cambio clínico logrado no diferencia a los dos grupos,...*”). Navarro (2000:912) explains that the adjective “severo” in Spanish means “serio” or “riguroso” and that it can only be used to define a person’s character. Therefore he distinguishes several possible equivalences to the English adjective “severe” depending on the context (e.g. “grave” as in “*severe situation*”, “intenso” or “fuerte” as in “*he had severe loss of blood*”, translated for “una fuerte hemorragia”, “extenso” as in “*severe acne*”, etc.). Finally he mentions the wide use of the inappropriate adjective “severo” in Spanish and underlines that linguists tend to stop correcting the inappropriate use of the term.

The adverb “comprehensivamente” is recorded in the corpus. It is modelled on the English adjective “comprehensive”. Navarro (2000:210) suggest different options for the translation of the adjective “comprehensive” according to the context: “amplio”, “completo”, “extenso”, “general”, “exhaustivo”, etc. In the corpus, the adverbs “ampliamente” or “exhaustivamente” should have replaced “comprehensivamente” (e.g. “*...un fenómeno multidimensional que se ha de abordar comprehensivamente*”).

e) Words with a meaning not registered in the dictionary

A clear example in the corpus is “adolecer” which is defined in the Spanish dictionary as “something having a defect or vice” but in the corpus it is used with the meaning of “to lack” (e.g. “*De manera similar a trabajos previos, este estudio adolece de algunas limitaciones metodológicas, que pueden influir en los resultados, como por ejemplo, control de la sintomatología ansiosodepresiva...*”).

- Compounds

In the semantic field the appearance of not previously attested compounds was registered in the corpus. Sometimes two simple words are joined forming a new compound. This compound appears either with or without a hyphen between the two words (e.g. “socioeducativo”, “ansiodepresiva”, “cortomedio” or “coste-beneficio”, “coste-eficacia”, etc.).

Regarding the formation of compounds by joining simple words the noun “coste” is frequently used forming different morphological compounds: “coste-eficacia”, “coste-efectivo” and “coste-efectividad” both with hyphen and without “coste efectividad” (e.g. “*Tampoco disponemos de datos económicos ni de coste-eficacia acerca de...*” or “*prescribir tabletas más grandes tiene mayor coste-efectividad*”).

The anglicism “cost-effectiveness” is criticised by Navarro (1994:507), as he considers “rentabilidad” is concise and even

shorter than the often occurring expression in medical texts “relación costo-efectividad”. Additionally, he points out that in this example shortness has not been a reason for adapting the anglicism.

Another combination present in the corpus with the noun “coste” is “coste-beneficio”. Other compounds are formed by the combination of two adjectives (e.g. “obsesivocompulsivo”, “farmacoeconómica” “ansiosodepresiva” and “cortomedio”). The latter compound adjective collocates with “plazo” (“term” in English) and is present in many different ways in the corpus. Sometimes researchers write “a corto y largo plazo”, others “a medio y/o largo plazo”, both of which are more logical for the Spanish language than “cortomedio”. Another term frequently used to form non-attested compounds is “evaluación”, forming hyphenated and non-hyphenated compounds like “evaluacióntratamiento” and “evaluación-intervención”.

- In the field of Syntax:

Elipsis

The use of the adverb “no” preceding nouns and adjectives to form antonyms, criticised by Martínez (2006:87), takes place in the corpus: e.g. “no causantes”, “no sexual”, “cirugía no cardiaca”.

Other instances of ellipsis registered in the text imply the omission of nouns rather than of verbs due to the fact that doctors give it for granted that the omission does not affect the

readers' understanding. They refer for instance to “sala de psiquiatría de agudos” and “hospitalización de agudos” where the noun “pacientes”, “patients” is omitted (e.g. “*En nuestra muestra, la venlafaxina retard en monoterapia se muestra retrospectivamente eficaz y rápida en monoterapia en el tratamiento de pacientes con depresión mayor en régimen de hospitalización de agudos,...*”)

Abusive use of the gerund and passive

The misuse of the gerund mentioned by Ballesteros (2003) is a common trait of the Spanish medical corpus. The use of the gerund is clearly against conciseness, since it results in long complex sentences with commas (e.g. “*En la actualidad no está justificada la utilización rutinaria de ATP típicos en el tratamiento de los síntomas no cognitivos de la demencia o SPCD, debiendo reservarse únicamente para situaciones puntuales de emergencia, aprovechando su presentación intramuscular.*”)

Regarding the use of the passive, the criticisms by Ballesteros (2003) and Martínez (2006) seem relevant since in fact they are right when they state that the influence of the English passive in Spanish results in less fluent texts and sometimes more difficult texts regarding understanding. An example of the cumbersome passive in the medical corpus is “*Aún más, el nerviosismo (PS-7) y el desgaste personal (PS-8, PS-24), las disfunciones en el sueño (PS-19) y el dolor de origen cardiovascular (PS-29) pueden ser disminuidas por la S ATs p*”.

The option of the so called “pasiva refleja” suggested by Martínez (2006:89) would turn the last example in a more comfortable text to read, namely “*se pueden disminuir con ATsp*”.

Copying the syntax of other languages

Several examples were mentioned by Martínez (2006:86) as syntactic structures copied from other languages are registered. In the corpus, for instance, the expression “jugar un papel” as a copy of the English or French verb “to play” or “jouer” (e.g. “*Las malformaciones venosas puras deben ser tratadas por un equipo multidisciplinar en el cual juega un papel fundamental el cirujano pediátrico,...*”).

Although, Navarro (2000:778) refers to the possible acceptance by the DRAE of the term “jugar un papel” in the near future due to its wide use, it has not been accepted yet. He suggests different possible equivalents in Spanish depending on the context: “intervenir”, “desempeñar una función”, “representar un papel”, “tener importancia”, “participar en”, “formar parte de”, “estar implicado”, “ser un factor importante”, etc.

Another example is the expression “en base a” copied from the English “on the basis of” instead of the Spanish corresponding structures like “basándonos en” or “teniendo en cuenta” (e.g. “*Según nuestros datos parece que el Ibuprofeno presenta menor eficacia que Indometacina en base al resultado de porcentaje de reapertura del ductus...*”). However, the most

recurrent instance is the incorrect use of “de cara a” instead of “con el fin de” or the preposition “para” taken from Frech (e.g. “*De cara a futuras investigaciones sería oportuno realizar trabajos de carácter longitudinal que evalúen la alexitimia y otras variables que puedan relacionarse con ella...*” or “*la ausencia de sujetos control y..., resulta un problema de cara a aseverar...*”).

The English structure “dependiendo de” modelled on the English “depending on” takes place in the corpus. In the following example the expression can be replaced by “de acuerdo con” or “según”: “*...pero sí es una variable que modula la ejecución dependiendo de las exigencias de la tarea experimental*”.

Another aspect that draws our attention is the use of explicative adjectives preceding the noun. Although in Spanish the adjective can take both positions (Gómez, 2002), either in front or after the noun, it is more common to use the so called “adjetivos especificativos”, i.e. after the noun, because this is the one that specifies and brings relevant information to the noun. Normally the preceding adjective is called an “epithet” where information is not relevant but its function is a rhetorical one. For instance, if the adjective “blaca”, white, precedes the noun “snow”, “nieve” it does not specify or give any information about the snow.

The Spanish medical corpus, however, has much more instances of adjectives preceding the noun than following it, and occasionally an adjective is found in combination with the same noun as an epithet and following the noun (e.g. “futuros

estudios” and “estudios futuros” or “futuras investigaciones” and “investigaciones futuras”), although, the epithet structure is more recurrent. This probable influence of the English language is reflected in the following examples: “específico apoyo”, “adecuada utilización”, “anteriores estudios”, “buena evolución”, “minucioso control”, “reducido porcentaje”, “elevado grado”, “serio problema”, etc. (e.g. “*Estamos a la espera de los resultados de los estudios multicéntricos en los que se concreten los criterios ecocardiográficos definitivos de asincronía que indiquen la adecuada respuesta al tratamiento de resincronización...*”).

Finally, the presence of participles in the Spanish corpus is very high. Some examples of recurrent participles are: “encontrados”, “analizados”, “obtenidos”, “estudiados”, “realizados”, “utilizadas”, etc. (e.g.: “resultados obtenidos”, “muestra estudiada”, “investigaciones realizadas”, etc.)

Neologisms and their inclusion in different dictionaries

The dynamism of the Spanish language as a reflect of advances in science is patent in Martínez’ words (2006:89) when he affirms: “son el uso y las costumbres las que culminan en la norma lingüística”. De este modo, la norma va siempre por detrás del uso. Martinez’ statement is in keeping with the analysis of the Spanish corpus for which a second dictionary was consulted, “Diccionario abreviado del español actual”. This dictionary has already incorporated many neologisms included in the Spanish corpus that were not accepted at the time of the

analysis (2007) by the latest edition of the RAE dated from 2002. However, according to Martínez some of these neologisms might be accepted in the future. In fact, it is common for the Real Academia to reject words and accept them after a while. As pointed by Fernández (2004:15), the DRAE tends to accept neologism later than other dictionaries like “María Moliner”, “Manuel Seco” and “Manuel Alvar”. Some examples of the neologisms in the corpus that have not been accepted by the Real Academia but that are present in the Diccionario Abreviado del español actual are mainly new words resulting from derivation. For instance, adverbs with the suffix “-mente”: “complementariamente”, “desafortunadamente”, “diferencialmente”, “estadísticamente”, “éticamente”, “frontalmente”, “inteligentemente”, “médicamente” and “positivamente” (e.g. “*Por ello se trataría de no rechazar frontalmente los tratamientos...*”). Other words present in the Spanish corpus that are registered in the Diccionario Abreviado del Español actual formed with suffixes are the following:

- The suffix “-ado” as in “mistificado
- The suffix “-ible” forming adjectives as in “prevenible”
- The suffix “-ivo” forming adjectives, for instance “evaluativo”.
- The suffix “-ción” forming nouns like “frecuentación” (e.g. “*En nuestro medio se encuentra una baja frecuentación por cefaleas como principal motivo de consulta.*”)

Other words formed by the addition of prefixes that were included in Diccionario Abreviado del español actual but not in the DRAE were:

- The prefix “co-” (e.g. “correlacionar” or “corresponsabilización”)
- The prefix “des-” (e.g. “desatender”)
- The prefix “in-” to create adjectives like “inefectiva” and “inespecífica”.
- The prefix “multi-” (e.g. “multifactorial”)
- The prefix “inter-” (e.g. “intercurrente”)
- The prefix “post-” (e.g. the adjective “postoperatorias”)

The acceptance of certain terms by Diccionario Abreviado del español actual that are not accepted at the present by the DRAE is a patent proof that we are dealing with a “living language” in medicine which is continuously evolving and adopting neologisms. This difference in the dictionaries acceptance on neologisms suggested that more differences could be found if more dictionaries were consulted. Obviously, the most suitable and likely dictionary to find the inclusion of medical neologisms not accepted in the previously checked dictionaries is a medical dictionary. The selected dictionary to carry out the investigation was: “*Diccionario de medicina Océano Mosby*” dated from 1994. Contrarily to what was expected, the difference regarding the neologisms included was not very significant. Most of the terms not present in the

previously checked dictionaries (DRAE and Diccionario Abreviado del español actual) were not included except for:

- “aleatorizado”, this term was also suggested by Navarro (2000:842) and Martínez (2006:84), when supporting the use of a Spanish terms rather than the use of anglicisms, in this case, “randomizado”.

- “obsesivo-compulsivo”, which appears with a hyphen between the two adjectives.

- “deprivación”, which was closer in terms of meaning to the adjective “depressant” in English.

In this case its meaning is not equivalent to the one in the Spanish corpus in which it refers to “deprivation” (e.g. “...alteraciones del funcionamiento de estructuras neurológicas que se deben a la acción aguda o sostenida de las drogas administradas, o bien son productos de la deprivación total o parcial del consumo...”). Therefore, the use of “deprivación” in the Spanish corpus could be a neologism based on the English noun “deprivation”. Navarro (1992:577) states that the word “deprivación” does not exist in Spanish and suggests “privación”, “deshabitación”, “carencia” or “pérdida” as possible translations in Spanish. Some examples of these translations given by Navarro (2000:267) are “affective deprivation” for “carencia afectiva” and “deprivation of sleep” for “privación del sueño”. From the different meanings proposed by Navarro, in the example sentence of the corpus shown above, the correct version for “deprivación total o parcial del consumo” is “privación total o parcial del consumo”.

Neologisms

As the English language is so spread in the scientific world, Spanish researchers tend to model their neologisms on the English language. Some neologisms have a very close aspect to the original English word. According to Navarro (1992:575) the formation of anglicisms is based on literal translation of English terms (e.g. “computerized”), although in some cases they are incorporated in our language directly and with no modifications (e.g. “shock”, “test”). As already stated at the beginning of the chapter, many times it is difficult to determine what the origin of a word is and it is not an exact science but something that can sometimes only be inferred by observing different possibilities. The neologisms present in the corpus have been divided into grammatical categories:

Adverbs

This is the case of the adverb “admitidamente”, which has the same form as “admittedly”. Another adverb is the neologism “comprehensivamente”, a possible copy of the English adjective “comprehensive” meaning “complete” or “large”, since no corresponding adjective with this form exists in Spanish.

Adjectives

- The neologism “aleatorizado” may have been modelled on the English adjective “randomized”. The existence of the adjective “aleatorio”, “random” in English might have converted into the neologism “aleatorizado”, a very frequent term used in English, mainly in the expression “randomized trial” and “randomized controlled trials”.

- Another adjective appearing in the corpus not registered in the DRAE that might be a consequence of a copy of an English term is the adjective “focalizado” used both in masculine and feminine “focalizadas” (e.g. “*Por el contrario la escala Hostilidad del JAS contiene elementos que hacen referencia principalmente a conductas [estilo duro, dominante y competitiva] sobretudo focalizadas en el ámbito laboral*”). This adjective might have its origin in the English adjective “focused” or it might be a neologism based on the verb “focalizar”. In the medical dictionary only the adjective “focal” was registered. Despite the acceptance by the DRAE in 2001 of the anglicism “focalizar” Navarro (2000:384) suggests avoiding it and using “enfocar” in technical contexts and “centrar”. “concentrar” or “determinar” in other contexts.

- The adjective “predictivos” is probably based on the English adjective “predictive”, which is translated in the Collins Dictionary as “profético”, “que vale como pronóstico” (e.g. “*Los factores predictivos de fracaso del stenting directo en la población analizada son la tortuosidad excesiva del vaso...*”). A synonym of this adjective is “predictor”, but neither “predictivo”

nor “predictor” were registered in the DRAE or in the medical dictionary “Diccionario Médico Océano Mosby” (e.g. “...*los grados elevados de glucemia y el recuento de leucocitos en el momento del ingreso fueron predictores independientes de mortalidad tardía*”). Navarro (2000) suggests avoiding the anglicism “predictor” and gives 3 possible equivalents in medical Spanish: a) “factor predisponente” or “valor pronóstico”, which correspond to the abbreviation of “predictor factor”, b) “prueba diagnóstica”, which corresponds to the abbreviation “predictor test” and c) “variable independiente”, which corresponds to the abbreviation “predictor variable”. Therefore, in the example sentence of the corpus “factores predisponentes” should have replaced “factores predictivos”, but as Navarro points out, the use of the former is very widely spread in medical texts and this kind of mistake is no longer contemplated when correcting for publication.

- As for the adjective “intrusiva” instead of the Spanish “intrusa”, it can be originated in the English term “intrusive (e.g. “...*seleccionar una intervención mínimamente intrusiva y que cause los mínimos efectos negativos en el mismo*”).

- The adjective “invasiva” in the corpus is probably a copy of the English “invasive”, for in Spanish the adjective “invasor” is the one that is registered in the dictionary (e.g. “*Hay que resaltar que ambas estimaciones pueden ser monitorizadas de manera no invasiva en el laboratorio de ecocardiografía*” or “*De ahí el interés por desarrollar modalidades diagnósticas que permitan, de forma fiable, la valoración no invasiva de los injertos*”). According to Navarro (2000:523) “invasiva” appears

as a consequence of the pressure of the English language on the Spanish language. He underlines that by 2000 its use is so widely spread in medical texts, that it is no longer corrected by linguists; therefore he concludes it might end up being included in the DRAE. Navarro gives two possible translations for this word: a) “infeccioso”, when referring to an illness or tumour or b) “traumático”, “agresivo”, “lesivo” (e.g. “invasive procedure” for “técnica traumática”).

- Another direct translation of the English language is that of the adjective “trained” in the expression “manos entrenadas” referring to “trained hands” that should have been translated with “especializadas” or “expertas”.

- According to Navarro (1994:145) the English adjective “ineffective” does not correspond to “inefectivo” in Spanish but to “ineficaz”, “inútil”.

Nouns

- The noun “comparabilidad”, not registered in the DRAE, appears as the translation for “comparability” in the Collins Spanish Dictionary dated from 1994, thus its origin can probably be in the English noun. Alternatively, this neologism can also be based on the adjective “comparable”.

- The same happens with “conceptualización” that is registered as the translation for “conceptualization” in the same dictionary. Despite the acceptance of the verb “conceptualizar” by the DRAE in 2001, Navarro (2000:211) states his preference

for other terms like “esclarecer”, “definir”, “expresar”, “imaginar”, etc.

- Another expression registered in the Collins Dictionary in Spanish is “en suma” as the translation for “in sum”; however, “en suma” was not included in the DRAE. Only the expressions “a lo sumo” and “de sumo” were registered, but they do not have an equivalent meaning to “in sum”. A possible Spanish equivalent to “in sum” is “en definitiva”.

- Another noun that is not registered in the dictionary is “deprivación”, which might be a copy of the English noun “deprivation”. The noun was in the medical dictionary with a different meaning (see above)

The noun “responsividad” is a possible loan translation of the English term “responsiveness” and should have been substituted in Spanish by “sensibilidad”, “interés”, “reactividad” (e.g. “bronchial responsivity”; “reactividad bronquial”) or “grado de respuesta” (Navarro, 2000:867), (e.g. “*Globalmente, hubiéramos esperado muchas más diferencias entre los diversos grupos clínicos, por lo que a la responsividad sexual se refiere...*”).

- Another noun that emerges as a possible semantic loan from English is the noun “significancia” from the English “significance”, which is translated as “significación”, “significado” or “trascendencia”. In Navarro’s words (2000:921) “significancia” cannot be used, the corresponding Spanish terms are “significación estadística” or in the common language “importante” (e.g. “*Ampliar la muestra de los subgrupos de*

pacientes con TCA para poder aplicar tests de significancia estadística”).

Verbs

The neologism “elicitar” is a possible loan from the English verb “to elicit” for which several Spanish words could be used instead like “obtener”, “provocar” or “sacar”. Navarro (2000:318) recommends avoiding this verb and using “obtener”, “producir”, “determinar” depending on the context.

Hyphenated medical words in English

Regarding English vocabulary, it is worth mentioning the many hyphenated items that take place in the English corpus. The most common Spanish translation for these terms is not reached by means of hyphens but Spanish needs longer expressions that contain a preposition. This enables the English language to concentrate information in few words and leads Spanish to longer sentences.

Additionally, it is difficult for Spanish speakers to become familiar with all hyphenated items, but learning the most typical ones in medicine is possible, and in fact advisable, since they are difficult to use and sometimes to interpret if not known. To illustrate this with an example, the term “all-cause mortality” corresponds to “mortalidad total” in Spanish. For this reason a list is provided with the hyphenated terms present in the corpus at the end of the research work (see appendix number

15). Some examples of the most common medical hyphenated terms are: combinations with “high” like “high-risk”, combinations with “low” such as “low-dose”, “low-risk”, combinations with “long” and “short”, especially “long-term” and “short-term”, combinations with the prefixes “pre-“ and “post-“ such as “pre- and postoperative care”, combinations with “self” like “self-efficacy”, hyphens with the noun “patient” like “patient-centered”, combinations with time expressions like “two-year period” and others like “follow-up”, etc. (e.g. “a 5 year survival”, “patient-related factors”, etc).

Sometimes the use by a non-native is more complex because more than one hyphen takes place within the same expression (e.g. “3-year failure-free survival rates”, “12-month follow-up study”). Occasionally hyphens are difficult because more words precede the noun and not only the hyphenated expression (e.g. “space-occupying conus medullaris mass”). Finally, the last factors that turns hyphens into a difficult aspect in medical English and in general English as well is the lack of consistency, i.e. sometimes the same combination of words is hyphenated and others it is not (e.g: “self-referential artefact” versus “selfreferential world”).

Another inconsistency present in the English corpus is the presence of terms used sometimes in English and sometimes in Latin. This is the case of the plural form of “sequel” that is sometimes taken from Latin, “sequelae”, and sometimes a normal English plural is presented “sequels.

CONCLUSIONS

IV. Conclusions

The starting point of this research work was the study of the situation faced by doctors who are not native speakers of English when writing RA for international journals. It was assumed that Spanish doctors would need linguistic support since theirs is a Romance rather than a Germanic language.

Firstly, a study was conducted to establish the state of the art of academic writing, corpus linguistics and contrastive rhetoric. Once the literature had been reviewed and constraints in the above fields were determined, carrying out research in the field of medicine seemed a sensible decision. Therefore, this research work was divided into two main parts: on the one hand a corpus was created containing an even number of Conclusion Sections of medical RAs in English and Spanish with the aim of establishing a comparison between the linguistic structures of the two languages. On the other hand it was necessary to conduct a survey to confirm the assumption that Spanish doctors need linguistic guidance and to consider the importance of carrying out research in this field.

The findings obtained in the survey for the investigation were the following: regarding the objective of assessing real needs on the part of medical doctors, it can be concluded that Spanish doctors in the disciplines of Psychiatry, Paediatrics and Cardiology aged 43+-14 years require support to publish scientific papers in English due to a lack of sufficient linguistic skills in that language.

Comparatively, German doctors in the same specialties aged 44+- 11 years need less help when writing an English paper due to a higher command of the English language.

The differences concerning linguistic proficiency can be attributed to two main factors: a) the lower percentage of Spanish doctors learning English at school and high-school b) the lower tendency to work in a foreign country for periods over 6 months and the lower selection of English speaking countries for that purpose.

As a consequence of the comparably lower command of the language, Spanish doctors have more difficulties than German doctors to read and write in English. Therefore, fewer doctors consult English journals as frequently as those written in their mother tongue. However, the most significant difference between the two countries is that Spanish participation in English publications is less than half that of German doctors, namely 41.2 % versus 90.9 %.

The second remarkable factor that accounts for Spanish doctors need for support to write English papers is that whereas 90% of German doctors write directly in English, only 10.7 % of the Spanish do so, since the tendency is to write first in their mother tongue. As a consequence, Spanish doctors often rely on professional translators for the translation. Due to the extreme translation activity in Spain, doctors would consider it useful to be taught how to use a simple language in their original draft, which would facilitate translation tasks.

With regard to the compilation of the corpora, it soon became clear that the assemblage of Conclusion Sections, not

merging with other sections but as an independent section containing the word “Conclusion”, was not a simple task for it is very common both in Spanish and English RAs to find the Conclusion and Discussion section under the same heading.

The semantic content of the medical RA Conclusion analysis shows that their structure is not clearly defined and divided into moves; in fact on many occasions, the Conclusion Section only provides the reader with a short report in the form of a summary. As a matter of fact, only 20% of the Conclusion Sections present all the five moves selected for genre analysis on RAs (Ricart, 2004) (i.e. *Background*, *Summarizing*, *Advantage*, *Limitation* and *Further Research*). Normally, only up to four moves can be recognized and on some occasions only two or three moves can be identified (80%).

The occurrence of recurrent lexico-grammatical structures was verified by means of a significant number of common collocations in both languages (e.g. “high-risk patients”, “long-term treatment” in English or “diferencias significativas” in Spanish).

The patent presence of the above mentioned lexico-grammatical structures made it possible to group recurrent structures under common semantic fields including word families and collocations in the different Moves. For instance, some common linguistic fields established for the move *Background* are: a) time expressions, b) comparison with previous studies written by other authors c) review of previous literature, etc.

Some common semantic fields could be compared in the 4 moves of the two languages (4 in the move *Background*, 7 in the moves *Summarizing* and *Limitation* and 11 in the move *Further Research*). On many occasions Conclusions showed literal equivalences, on others, no literal equivalent was found but expressions were equivalent in terms of meaning (e.g. no literal equivalent was found for “concordar”, “estar en concordancia con”, “ser coherente con” or “coincidir” but they are equivalent in meaning to “to be in agreement with” or “to be in line with”).

Additionally some words appear more often in one language than in another (e.g. “to elucidate” or “to study”, which is used more frequently than the noun “paper”). This fact should be taken into account by non-native writers when choosing English words if they want their writing to sound natural in medical English. In order to achieve an appropriate English style, Spanish doctors should simplify the language in terms of vocabulary and syntax and use scientific terminology in accordance with the academic context.

On some occasions, expressions had no equivalent in the other language (e.g. “finalizar el tratamiento”, “complicaciones postoperatorias”, “methodological problem”, etc.). Evidence of non-equivalence is shown in the *Move Further Research* in the section devoted to expressions that occur only in one of the two languages (e.g. Spanish expressions such as “habrá que” or “será conveniente” or English expressions such as “prospective studies” or “warrants further investigation”).

Another interesting example in relation to “non-equivalence” is the different effect produced by the presence of the author as the subject of the grammatical sentences. When they appear in English their presence is normally more positive than in Spanish (e.g. “we have clearly shown”, “we are the first to report”, etc). Spanish doctors resort to expressions such as “podemos decir” to express some findings. However, these kinds of expressions are often preceded by the adverb “no” converting the sentence into a limitation rather than into a finding.

Finally, the Spanish corpus showed evidence that medical language continues to evolve, growing especially in the lexical area. The present research work revises the different areas related to the creation and influence of neologisms in the Spanish language in the last ten years. It can be stated that Anglicisms, borrowings and derivation are the main ways for the creation of neologisms. Despite the critique of different authors, Spanish medical language continues to adapt false friends and Anglicisms incessantly.

The use of Anglicisms has to do with the capacity of the English language to condense a lot of information into a few words. This is reflected both in the use of certain Anglicisms such as “stent” (endoprótesis vascular) as well as in hyphenated words (e.g. “self-efficacy”, “catheter-related”, “failure-free”). The translation of these hyphenated terms into Spanish would require a higher number of words.

Additionally, although Anglicisms do not always imply shortness, their spread use in English medical texts leads

Spanish medical doctors familiar with English medical literature to incorporate them in their Spanish publications (e.g. “estrategias de coping”, “test-retest”, and “screening”).

Another manifestation of the aforementioned linguistic contamination of medical English in Spanish medical RAs is evident in terms such as “hipotetizar”, “randomizado” or “comportamental”. These words are not included in the DRAE but they come up in Spanish medical texts. In this sense, the plasticity of the language is reflected in the use of words not accepted by the DRAE at the beginning but most being included in the dictionary with time.

Spanish medical language reproduces syntactic structures from other languages, mainly from English and French. Due to the predominance of the English language in medical RAs, there is a marked influence on the Spanish language regarding the position of the adjectives and nouns (“futuros estudios” versus “estudios futuros”). The use of passive, gerund and ellipsis is also present in the Spanish corpus despite making the Spanish language less comprehensible and less clear.

If linguists do not strive in their task of analyzing and limiting the acceptance of certain foreign terms and structures, Spanish medical language will turn into a hybrid language as the term “inmortales-like” shows.

One possible line of study could be to provide Spanish doctors with a list of equivalent structures in the English language to support them when writing the Conclusion Section of a medical Research Article.

Further Research could extend this analysis of equivalent recurrent structures in medical Conclusion Sections to other sections of RAs. Additionally, the influence of English on Spanish production in terms of Anglicisms, syntax, etc. could also be observed in the other sections.

This work also contributes to the field of English for Specific Purposes with specific linguistic features and language patterns useful for teaching English to students of health sciences or medicine.

Another possible line of study could consist in the same analysis carried out in the present research work but taking English and other languages other than Spanish.

Further Research could also aim at comparing the Conclusion Section structure of RAs in other scientific disciplines to determine whether it is possible to generalize about a specific structure for this section and about recurrent expressions.

Finally, this research highlights the possible combination of translation studies with technical writing for an area of study called writing for translation, which encourages the use of a plain and concise language to aid in the faster and easier dissemination of research in other languages.

CONCLUSIONES

V. Conclusiones

El punto de partida de esta investigación ha sido el estudio de la situación a la que se enfrentan los médicos – no nativos de habla inglesa- cuando escriben un artículo de investigación para revistas internacionales. Se asume que especialmente el médico español va a necesitar apoyo lingüístico para su redacción en inglés, dada la naturaleza románica y no germánica del español.

En primer lugar, se llevó a cabo un estudio para determinar el estado del arte de la redacción académica, la lingüística de corpus y la retórica contrastiva.

Una vez revisada la literatura y establecidos los límites en los campos mencionados, la investigación se centró en el campo de la medicina. Como consecuencia, la presente investigación se dividió en dos partes principales: por un lado se creó un corpus que contenía un número parecido de secciones de Conclusión de artículos de investigación médica en español y en inglés. El objetivo era establecer una comparación entre las estructuras lingüísticas de los dos idiomas. Por otro lado fue necesaria la implementación de una encuesta para confirmar la hipótesis de que los médicos españoles necesitan una mayor dirección lingüística y considerar la importancia de llevar a cabo la investigación.

Los resultados obtenidos de la encuesta para la presente investigación son los expuestos a continuación:

Respecto al objetivo de evaluar las necesidades reales de los médicos españoles a la hora de escribir artículos en inglés médico, se puede concluir que los especialistas en Psiquiatría, Pediatría y Cardiología con edades comprendidas entre 29 - 57 años (edad media 43 ± 14) necesitan ayuda, debido a la falta de conocimientos lingüísticos.

Comparativamente, los médicos alemanes de las mencionadas especialidades con edades comprendidas entre 33-55 años (edad media 44 ± 11), necesitan menos ayuda para escribir artículos médicos en inglés, debido a un mayor dominio del idioma.

Las diferencias con respecto al dominio lingüístico, pueden atribuirse principalmente a dos factores: a) al menor porcentaje de médicos españoles que han estudiado inglés durante su formación secundaria y bachillerato respecto a los médicos alemanes b) la menor tendencia a trabajar en un país extranjero durante períodos superiores a 6 meses así como a que eligen con menor frecuencia países de habla inglesa para este propósito.

Como consecuencia del menor dominio del idioma inglés por parte de los médicos españoles en relación a los alemanes, los primeros tienen mayores dificultades para leer y

escribir en inglés. Por lo tanto, el porcentaje de médicos españoles que consulta revistas inglesas para estar al día en su especialidad, es menor que el de los alemanes, consultando los primeros con más frecuencia revistas en su propia lengua. En este sentido, la diferencia más relevante entre ambos países es que los médicos españoles publican en inglés menos de la mitad de artículos de investigación que los alemanes, es decir, 41,2% versus 90,9%.

Un segundo factor a destacar que justifica la necesidad de apoyo por parte de los médicos españoles para escribir artículos médicos en inglés, es que mientras el 90% de los médicos alemanes escribe directamente en inglés cuando publica un trabajo, solo el 10,7% de los españoles lo hace. Consecuentemente, los médicos españoles recurren con mucha mayor frecuencia a traductores profesionales.

Por este motivo, los médicos españoles considerarían útil el aprendizaje de ciertas estructuras simples a la hora de escribir el texto en español que facilitase su posterior traducción.

En lo referente a la elaboración del corpus, podemos concluir que la obtención de las secciones de *Conclusión* no fue tarea fácil ya que en la mayoría de los casos las secciones *Conclusión* y *Discusión* de los trabajos médicos aparecen en un apartado común.

Respecto al contenido semántico de la *Conclusión* de los artículos de investigación, su análisis mostró que su estructura no está claramente definida ni dividida en los distintos *Moves*. De hecho, muchas veces la sección *Conclusión* sólo proporciona al lector un breve informe, a modo de resumen del artículo. En este sentido, sólo un 20% de las secciones de *Conclusión* presentaban los cinco *Moves* seleccionados para el análisis de género de los artículos de investigación (Ricart, 2004), es decir: *Background*, *Summarizing*, *Advantage*, *Limitation* y *Further Research*). En el 80% de los casos sólo podía identificarse un máximo de dos o tres *Moves*.

En lo referente a las estructuras lexico-gramaticales recurrentes, su presencia queda manifiesta a través de un gran número de colocaciones en ambas lenguas (e.g. “high-risk patients”, “long-term treatment” en inglés, o “diferencias significativas” en español).

La presencia de las mencionadas estructuras lexico-gramaticales repetitivas, posibilita su agrupación en campos semánticos comunes, incluyendo familias de palabras y colocaciones en los diferentes *Moves*. Por ejemplo, algunos de los campos lingüísticos comunes encontrados en el *Move Background* en ambas lenguas son: a) expresiones de tiempo, b) la comparación de estudios previos llevados a cabo por otros autores, c) la revisión de la literatura previa, etc.

Respecto a las equivalencias en los cuatro *Moves* en las dos lenguas, podemos concluir que existen campos semánticos comunes, que pudieron compararse (4 en el *Move Background*, 7 en los *Moves Summarizing* y *Limitation* y 11 en el *Move Further Research*). En muchos casos se observaron equivalencias literales en las Conclusiones, y en ocasiones aunque no se trataba de una equivalencia literal, existía una equivalencia semántica (e.g. no se observó equivalencia literal en “concordar”, “estar en concordancia con”, “ser coherente con”, o “coincidir”, pero aunque con sutiles diferencias semánticas, todas ellas tenían un significado común equivalente a “to be in agreement with” o “to be in line with”). Adicionalmente, hay que destacar que ciertas palabras o expresiones aparecen con mayor frecuencia en un idioma que en el otro, (e.g. “to elucidate” o “study” cuyo uso es mucho más frecuente que el de “paper”). Este hecho debería ser tenido en cuenta por el autor no nativo a la hora de elegir las palabras en inglés, si quiere que su redacción “suene natural” en inglés médico. Para una adecuada redacción en inglés, un médico español deberá simplificar la lengua en cuanto a vocabulario y sintaxis y buscar la terminología científica adecuada al contexto académico.

En algunas ocasiones las expresiones no tenían equivalencia en la otra lengua (e.g. “finalizar el tratamiento”,

“complicaciones postoperatorias”, “methodological problem”, “escasos estudios”, etc.). El hecho más destacado de la falta de correspondencia se refleja en el apartado que se ha dedicado a las expresiones exclusivas de cada idioma por separado en el *Move Further Research* de la presente investigación (e.g. expresiones como “habrá que” o “será conveniente” en español y expresiones como “prospective studies” o “warrants further investigation” en inglés).

Otro ejemplo a destacar relacionado con la “no equivalencia” es el diferente efecto que produce la presencia del autor como sujeto de las oraciones gramaticales. Su presencia en inglés tiene normalmente un efecto más positivo que en español (e.g. “we have clearly shown”, “we are the first to report”, etc”). El médico español para expresar hallazgos recurre a expresiones del tipo: “podemos decir”. De hecho, las expresiones como “hemos encontrado” suelen ir precedidas del adverbio “no” y expresan una limitación y no un hallazgo.

Finalmente, el corpus español evidenció que el lenguaje médico continúa evolucionando y creciendo. El presente trabajo revisa las distintas áreas cubiertas por expertos en los últimos 10 años en lo referente a la creación e influencia de neologismos en la lengua española y en lo relacionado con problemas de traducción y verifica que los anglicismos, préstamos y derivados lingüísticos siguen siendo las principales fuentes para la

creación de neologismos. A pesar de la crítica de diversos autores, el español médico sigue adoptando falsos amigos y anglicismos incesantemente.

El uso del anglicismo está condicionado por la capacidad de la lengua inglesa de condensar mucha información en pocas palabras. Esto se refleja tanto en el uso de ciertos anglicismos como “stent” (endoprótesis vascular) como también en las denominadas “hyphenated words” (e.g. “self-efficacy”, “catheter-related”, “failure-free”), cuya traducción en ocasiones requeriría un mayor número de palabras para expresar el mismo concepto en español.

Adicionalmente, aunque el anglicismo no implique siempre una mayor brevedad de la palabra, su extendido uso en los textos de inglés médico, conlleva a que los médicos españoles familiarizados con la literatura médica inglesa los incorporen en las publicaciones en español (e.g. “estrategias de coping”, “test-retest”, “screening”).

Otras manifestaciones de la contaminación lingüística del inglés médico en los artículos escritos en español se hace evidente en términos tales como “hipotetizar”, “randomizado” o “comportamental”, palabras que no aparecen en el DRAE pero que encontramos en dichos textos. En este sentido la plasticidad de la lengua se refleja en el uso de palabras no aceptadas por el

DRAE en sus inicios, pero que finalmente acaban siendo incluidas en el diccionario.

El español médico continúa copiando estructuras sintácticas de otras lenguas, fundamentalmente del inglés y del francés. Dado el predominio de la lengua inglesa en los artículos médicos, se observa una marcada influencia en la lengua española, en lo referente a la posición del adjetivo respecto del nombre (e.g. “futuros estudios” vs. “estudios futuros”). El uso de la voz pasiva, del gerundio y de la elipsis también está presente en el corpus español, a pesar de que este hecho dificulta su comprensión.

Si los lingüistas no se esfuerzan en analizar y evaluar la aceptación de ciertos términos y estructuras extranjeras, el español médico corre el riesgo de transformarse en un lenguaje híbrido tal como lo demuestra el término “inmortales-like”.

El presente trabajo puede servir de guía a los médicos españoles si se agrupan las estructuras equivalentes en ambos idiomas, lo que permitiría la elaboración de un manual de estilo, a la hora de redactar las Conclusiones de los artículos de investigación en medicina.

Los resultados del presente trabajo deberían despertar el interés de los expertos lingüistas para tratar de extrapolar los hallazgos obtenidos en estructuras recurrentes en la sección Conclusiones a otras secciones de los artículos de investigación

médica. Adicionalmente, la influencia del inglés en la lengua española en lo referente a anglicismos, sintaxis, etc., también podría evaluarse en otras secciones de los artículos de investigación, lo que sin duda podría constituir un futuro objetivo de indudable interés.

La presente investigación supone una pequeña contribución al campo de IFE (Inglés para fines específicos) aportando características lingüísticas específicas y patrones lingüísticos útiles para la enseñanza del inglés dirigido a estudiantes de ciencias de la salud o de medicina.

Otra posible línea de investigación consistiría en llevar a cabo el análisis realizado en el presente trabajo de investigación, comparando el inglés con otras lenguas distintas del español.

Finalmente, también cabría comparar la estructura de las Conclusiones de los artículos de investigación de medicina con la de otras disciplinas científicas, para determinar si es posible la generalización de una estructura determinada común para la sección Conclusiones y al mismo tiempo observar si se confirma también la repetición de expresiones recurrentes.

Por último, este trabajo abre el camino hacia una nueva área de estudio que combina la traducción y la redacción científica: “writing for translation” (redacción de textos que van a ser traducidos). Ésta apoya el uso de un lenguaje sencillo y

conciso en la redacción de los trabajos de investigación para agilizar y facilitar su diseminación en otras lenguas.

BIBLIOGRAPHY

VI. Bibliography

- Abad García, F. et al. (2006) Acceso abierto a revistas médicas españolas. *Medicina clínica (Barc)*, 0, 1-15.
- Aijmer, K. & Altenberg, B. (1991) *English Corpus Linguistics*. London: Longman.
- Alarcos Llorach, E. (1980) *Estudios de Gramática Funcional del Español*. Madrid: Gredos.
- Alarcos Llorach, E. (1995) *Gramática de la Lengua Española*. Madrid: Espasa Calpe.
- Alberch, P. (1996) Language in contemporary science: the tool and the cultural icon. In: Chartier R, Corsi P, dirs. *Sciences et langues en Europe*. Actas del coloquio celebrado en París del 14al 16 de noviembre de 1994. París: Centre Alexandre Koyré, 257-64.
- Álvarez, Álvarez, S. (2005) *Retórica contrastiva y calidad de la traducción: los patrones organizativos de las lenguas como motor principal del proceso traductor*. Downloaded from:
www.uem.es/web/fil/invest/publicaciones/web/autores/alvarez_art.htm on 11-12-2007.
- Amador Iscla, A. and Aleixandre, R. (2003) Los problemas del lenguaje técnico en la literatura médica actual. *El Médico Interactivo*. *Diario Electrónico de la Sanidad (Valencia)* nº 892-10. Downloaded from:

<http://www.medynet.com/elmedico/informes/informe/lenguaje.htm> on 06-07-2007.

American Institute of Physics, eds. (1990) *Style Manual of AIP*. (fourth edition, 2nd printing).

American National Standards Institute (ANSI), eds. (1979, 1989): *American National Standard for the Preparation of Scientific Papers for Written and Oral Presentation*.

American Psychological Association: APA Style: downloaded from: <http://www.apastyle.org/electref.html> on 07-01-2007.

Askehave, I. and Swales, J.M. (2001) Genre identification and communicative purpose: A problem and a possible solution. *Applied Linguistics* Vol. 22 p. 195-211.

Martín Arribas, M.C. (2004) Diseño y Validación de Cuestionarios. *Matronas Profesión*. Vol. 5 nº 17 p. 23-29.

Baker, M., Francis, G., Tognini-Bonelli, E.(eds) (1993) *Text and Technology*. In Honour of John Sinclair. Amsterdam: John Benjamins.

Baker, M. (1992) *In Other Words, A Coursebook on Translation*. London: Routledge.

Baker, P. et all. (2006) *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Ballesteros Fernández, .A. (2003) El lenguaje de los médicos. El médico interactivo. *Diario electrónico de la sanidad*. nº 892-10. Downloaded from:

<http://www.medynet.com/elmedico/informes/informe/lenguajemedico.htm> on 06-07-2007.

- Bassnett, S. (1988) *Translation Studies*. London: Routledge.
- Bazerman C. (1988) *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison: University of Wisconsin.
- Berg, E.C., Hult, F.M. and King, K.A. (2001). Shaping the climate for language shift? English in Sweden's elite domains. *World Englishes*, 20 (3), 305-319.
- Bernard Wilkinson, S. (1989). *Inglés Práctico para Médicos*.
- Bernard Wilkinson, S. (1992). *Diccionario Médico Práctico Inglés-Castellano*. Ed Benard, Valencia.
- Biber, D., Conrad, S., Reppen, R. (1998). *Corpus Linguistics: Investigation Language Structure and Use*: Cambridge University Press.
- Bhatia, V.K. (1997a) Applied genre analysis and ESP. In T. Miller (Ed.), *Functional approaches to written text: classroom application* (pp.134-149). Washington D.C: USIA.
- Bhatia, V.K (1997b). Introduction: Genre analysis and world Englishes. *World Englishes*, 16 (3), 313-319.
- Bhatia, V.K. (1997c) The power and politics of genre. *World Englishes*, 16 (3), 359-371.

- Buican, I., Hriscu, V. & Amador, M. (1993) Using International English to prepare technical text for translation. Los Álamos National Laboratory.
- Buican, I., Hriscu, V. & Amador, M. (1994) Preparing technical text for translation: A comparison between International English and methods for simplifying language. Los Álamos National Laboratory.
- Bratu, F. Guidelines for Designing Documents intended for translation. Downloaded from: <http://desktoppub.about.com/od/layout/a/translation.htm> on 10-11-2007.
- Candel Mora, M.A. (2003) *Terminología aplicada a la traducción técnica de informes sobre la gestión de la I+D*. Valencia: Editorial UPV.
- Ciapuscio, G. & Otañi, I. (2002) *Las conclusiones de los artículos de investigación desde una perspectiva contrastiva*. RILL, Vol. 15, p.117-133. Universidad de Buenos Aires-Conicet.
- Clyne M. (1987) Discourse structures and discourse expectations: Implications for Anglo-German academic communication in English. In: *Discourse across cultures: Strategies in world Englishes* (Smith LE, ed) pp 73-83.
- Collins Cobuild English Dictionary. Harper Collins Publishers. London: 1995.

- Collins Cobuild English Language Dictionary. London; 1987.
- Collins English Thesaurus in A Z form (1995). Glasgow: Harper Collins Publishers.
- Connor, U, & Kaplan, R.B. (1987). Writing across languages: Analysis of L2 text, Addison- Wesley.
- Connor, U. (2002) New Directions in Contrastive Rhetoric (TESOL Quarterly, Vol 36 p. 493-510. Indiana University-Purdue University Indianapolis.
- Connor, U. (2004) Intercultural rhetoric research: beyond texts. *Journal of English for Academic Purposes* 3 p. 291-304.
- Cotterall, S & Cohen, R. (2003) Scaffolding for second language writers: producing an academic essay. *ELT Journal*, Vol.57 p.158-166
- Crompton, P. (1997) Hedging in Academic Writing: Some Thoretical Problems. *English for Specific Purposes*, Vol 16 n°4, p.271-287.
- Crystal, D. (1997) *English as a Global Language*. Cambridge, U K. Cambridge University Press.
- Day, R. A. (1998) How to write and publish a scientific paper, 5th ed. Oryx Press.
- Diccionario de la Lengua Española (2001) vigésima segunda edición. Madrid: Espasa Calpe.

- Diccionario de medicina océano Mosby (Mosby's medical nursing and allied health dictionary) (1994). MMI, Sant Louis Missury.
- Diccionario Latino (1992) VOX, 19ª edición, Barcelona: Bibliograf.
- Diccionario latino-español, español-latino, (1969) 6ª edición. Barcelona: Biblograf.
- Dudley-Evans, T. (1991) Socialisation into academic community: Linguistic and stylistic expectations of a Ph. D. thesis as revealed by supervisor comments. In: *Sociocultural issues in English for academic purposes* (Adams P, Heaton B, Howarth P, ed) London: MacMillan, 41-51.
- Dudley-Evans, T. (1997) Genre Models of the Teaching of Academic Writing to Second Language Speakers: Advantages and Disadvantages. In T Miller (ed.) *Functional Approaches to Written Text: Classroom Applications* (p 150- 159). Washington, D.C, U.S.A.: USIA.
- Duque García, M. (1999), Manual de Estilo. El arte de escribir en Inglés científico-técnico. Madrid: Paraninfo.
- Dyke Ford, J. (2004) Knowledge Transfer across Disciplines: Tracking Rhetorical Strategies from a Technical Communication Classroom to an Engineering Classroom.

IEEE Transactions on Professional Communication. Vol. 47, nº4, 301-315.

- Fernandez, Fernández, M. (2004) La incorporación de neologismos al español actual. Extranjerismos y nuevas acepciones. Ávila: Universidad Católica de Ávila.
- Field, W. F. (1984) *Linguistic borrowing in bilingual contexts*. Amsterdam: John Benjamins.
- Flowerdew, L. (2000). Using a genre-based framework to teach organizational structure in academic writing. *ELT Journal* Vol. 54/4.; p.369-378.
- Fortanet Gómez, I. (2002) *Cómo escribir un artículo de investigación en inglés*. Madrid: Alianza Editorial, S.A.
- García Yebra V. La interferencia lingüística. En: Teoría y práctica de la traducción (2ª ed.) Vol I. Biblioteca Románica Hispánica, manuales, 53. Madrid: Dedos, 1989; 353-384.
- Garrido, M: (2003) Aplicaciones pedagógicas del informe empresarial en el aula del inglés de los negocios: implementación de recursos discursivos escritos en contextos de especialidad. (chapter 1: *El inglés como lengua internacional* and chapter 2: *Adquisición del inglés como lengua de especialidad*) Universidad Jaume I.
- Gili Gaya, S. (1994) *Curso Superior de Sintaxis Española*. Barcelona: Bibliograf.

- Gledhill, C. (2000) The discourse function of collocation in research article introductions. *English for Specific Purposes* 19 p. 115-135.
- Gómez Torrego, L. (2002) Gramática didáctica del español. Madrid: Ediciones SM. 18-25.
- Gosden, H. (1995) Success in Research Article Writing and Revision: A Social-Constructionist Perspective. *English for Specific Purposes*, Vol. 14. n°1 p.37-57.
- Graddol, D. (1997a) The future of English? London: The British Council.
- Graddol, D. (1997b) English: history, diversity and change. London: Routledge, p.171-179.
- Gross, Alan G. (2002) Communicating Science: The Scientific Article from the 17th Century to the Present. New York: Oxford University Press.
- Hackett, S. Think globally about your software and documentation. Downloaded from: <http://www.epictrends.com/resources/localization/Think%20globally.shtml> on 4-11-2007.
- Hall, P.T. (1995) Designing a Multilingual Document: Lessons for the Manager. STC Proceedings. Downloaded from: <http://www.stc.org/confproceed/1995/PDFs/PG133135PDF> on 05-05-2007.

- Hengl, T. & Gould, M., *Rules of thumb for writing Research Articles* (2002). Downloaded from: <http://www.itc.nl/personal/hengl/RT> on 03-02-2007.
- Henle J. (1871) *Handbuch der systematischen Anatomie*. Braunschweig: Vieweg.
- Hinds, J. (1987): "Reader versus Writer Responsibility. A New Typology" in *Writing Across Languages: Analysis of L2 Text*. Connor et al (eds.) Addison-Wesley. 141-152.
- Horacio, C. (2004) *Libelo contra el Inglés. Revista bibliográfica de Geografía y Ciencias Sociales* Vol. IX, nº 490.
- House, J. (1999): "Misunderstanding in intercultural communication: interactions in English as a lingua franca and the myth of mutual intelligibility" in C. Gnutzmann (ed.), 73-89.
- Hunston, S. and Francis, G. (1999). *Pattern Grammar*. Amsterdam: John Benjamins.
- Hyland, K. (1996) Talking to the academy: Forms of hedging in scientific research articles. *Written Communication* 13/2, 251-181.
- Hyland, K. (2002). *Teaching and researching writing*. Harlow, Essex: Longman.
- Huth, E.J. (1990) *How to write and publish papers in the Medical Sciences*. (Editor, *Annals of International Medicine* Williams and Wilkins. First published 1982).

- Jenkins, J. (2000) *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- Johansson, Sig. *Seeing Through Multilingual Corpora: On the Use of Corpora in Contrastive Studies*. Amsterdam: John Benjamins, 2007 p.23. *Using Multilingual Corpora*.
- Kachru B. (1995) *World Englishes: Approaches, issues, and resources*. In: *Readings on second language acquisition* (Brown HD, Gonzo S, eds) Englewood Cliffs, NJ: Prentice-Hall, pp 229-261.
- Karamanian, A.P. (2002) *Translation and Culture*. *Translation Journal*. Vol.6 n°1.
- Kei Matsuda, P. (2001) *On the origin of contrastive rhetoric: a response to H.G. Ying*. *International Journal of Applied Linguistics*, Vol. 11, n°2, p. 257.
- Kennedy, B. (1988) *Adult versus Child L2 Acquisition: An Information-Processing Approach*. *Language Learning* Vol. 38 (4) p 447-495
- Kerans, M.E. (2005) *Project Management and Quality Assurance in Cover-to Cover Translation of a Medical Journal*. *IEEE*. p.220-236.
- Lee, D.Y.W. (2001). *Genres, registers, text-types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle*. *Language Learning and Technology* 5, p.37-72).

- Leech, G. (1991) *The State of the Art in Corpus Linguistics*. In Aijmer & Altenberg (eds.), 8-29.
- Lewin, B. (2005) Hedging: an exploratory study of authors' and readers' identification of "toning down" in scientific texts. *Journal of English for Academic Purposes* 4 p.163-178.
- Luzón, Marco, M.J. (2000) Collocational frameworks in medical research papers: a genre-based study. *English for Specific Purposes* 19 p. 63-86.
- Manning, C. & Schütze, H. (1999) Term, Technical Term, Technological Phrase. Collocations. p.141-177.
- Markel, M. (1992): *Technical Writing: Situations and Strategies* (3rd ed.). St. Martin's Press, 2.
- Martín Arribas, M.C. (2004) Diseño y validación de cuestionarios *Matronas Profesión* Vol. 5 num. 17, p.23-29.
- Martin, J. 1985. "Process and Text: two aspects of semiosis" in J.D. Benson and W.S. Greaves (eds.) *Systemic Perspectives on Discourse, Vol. 1: selected theoretical papers from the 9th internacional Systemic Workshop* Norwood, NJ: Ablex.
- Martín Martín, P. (2003) A Genre Analysis of English and Spanish Research Paper Abstracts in Experimental Social Sciences. *English for Specific Purposes* 22, p. 25-43.

- Martínez Ramos, D. (2006) El lenguaje científico en cirugía. Una asignatura pendiente. *Cirugía española*. Vol. 79, nº 02 p. 83-88.
- Mauranen A. (1993) Contrastive ESP rhetoric: Metatext in Finnish-English economics texts. *English for Specific Purposes* 12/1, 3-22.
- Mathes, J. & Stevenson, D. (1976): *Designing Technical Reports: Writing for Audiences in Organizations*. Bobbs-Merrill Educational Publishing, 3.
- Matías-Guiu, J. (1996) Las publicaciones científicas en español. *Rev Neurol (Barc)*; 24, 506.
- Mayor Serrano, M^a B. (2002) La formación de los traductores médicos. Sugerencias didácticas. *Panacea*. Vol.3, n.9-10 p 83-89.
- Mayor Serrano, M^a B. (2006) Periodistas y traductores médicos: ¿dos mundos diferentes? *Panacea* Vol. VII, nº 23.
- Meierkord, C. (1996). *Englisch als Medium der interkulturellen Kommunikation. Untersuchungen zum non-native-/non-native speaker- Diskurs*. Frankfurt /Main: Lang.
- Montero, B. et al (2007) Discussion forum interactions: text and context. *System* 35/4. p. 566-582.
- Moreno, A. (2004) Retrospective labelling in premise-conclusion metatext: an English-Spanish contrastive study

of research articles on business and economics. *English for Academic Purposes* 3, p. 321-339.

Moreno, A. (1997) Genre Constraints across Languages: Causal metatext in Spanish and English RAs. *English for Specific Purposes*, Vol. 16, N°3 p. 161-179.

Mosby's Medical, Nursing & Allied Health Dictionary (1998), Mosby, Sant Louis.

Navarro, F.A. & Hernández, F. (1992) Palabras de traducción engañosa en el inglés médico. *Medicina Clínica Barcelona*. Vol.99 p. 575-580.

Navarro, F.A. (1993) El nuevo Diccionario de la Real Academia Española: su repercusión sobre el lenguaje médico. *Medicina Clínica Barcelona*. Vol. 101, p. 584-590.

Navarro, F.A. & Hernández, F. (1994) Nuevo listado de palabras de traducción engañosa en el inglés médico. *Medicina Clínica Barcelona*. Vol.102, p. 142-149.

Navarro, F.A. (1995) Tercer listado de palabras de traducción engañosa en el inglés médico. *Medicina Clínica Barcelona*. Vol.105, p. 504-514.

Navarro, F.A. & Barnes, J. (1996) Traducción de títulos al inglés en Medicina clínica: calidad e influencia del castellano. *Medicina Clínica Barcelona*. Vol. 106, p 298-303.

- Navarro, F.A. (2005) *Diccionario Crítico de dudas inglés-español de medicina* 2ª edición. MC. Graw-Hill.interamericana. Madrid.
- Oxford English Dictionary, Oxford: Oxford University Press. New York. 1994.
- Partington, A. (1998) *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. In: *What we can do with Corpora*. Amsterdam: John Benjamins.
- Peakock, M. (2002) *Communicative moves in the discussion sections of Research articles*.
- Petzold, R. & Berns, M. (2000). *Catching up with Europe: Speakers and functions of English in Hungary*. *World Englishes*, 19 (1), 113-124.
- Piqué-Angordans, J. & Posteguillo, S. (2006 a). *Peer positive and negative assessment in medical English written genres*, M. Gotti & F. Slager-Meyer eds, *Advances in Medical Discourse Analysis: Oral and Written Contexts*, 383-406. Bern: Peter Lang.
- Piqué-Angordans, J. & Posteguillo, S. (2006 b). *Medical Discourse and Academic Genres*. *Encyclopedia of Language & Linguistics*. Elsevier Ltd.
- Piqué-Angordans, J., Posteguillo, S., Melcion, L. (2006 c) *The development of a computer science dictionary, or how to help translate the untranslatable* in *Information*

Technology in Languages for Specific Purposes p.211-229
E. Arnó Macià, A. Soler Cervera and C.Rueda Ramos:
Springer.

Posteguillo, S. 1999. The schematic structure of computer science research articles. *English for Specific Purposes* Vol. 18. p.139-160.

Quirk, R., Greenbaum, S. (1990) A Student's Grammar of the English Language. Harlow Essex, England: Longman.

Quirk, R., Greenbaum, S. (1972) A Grammar of Contemporary English. London: Longman.

Real Academia Española. Diccionario de la Lengua Española. 22.ª ed. Madrid: Espasa Calpe; 2001.

Ren, S., Liang, P., Zu, G. (1999) The challenge for Chinese scientific journals. *Science*, 286, 1683.

Reppen, R., Fitzmaurice, S.M., Biber, D. (2002). *Using Corpora to Explore Linguistic Variation*.

Ricart Vayá, A. (2004) A Genre rhetorical and lexicogrammatical analysis of a corpus of academic articles (unpublished).

Rimalower, G.P. (2000) Crossing Borders: Tips for Preparing your Writing for subsequent translation. STC Proceedings. Downloaded from:
<http://www.stc.org/confproceed/2000/PDFs/00019.PDF> on 15-07-2007.

- Riquelme, J. (1998) *Los Anglicismos: Anglicismos y Anglicismos huéspedes de la lengua*. Editorial Aguaclara. Valencia. Amsterdam: John Benjamins.
- Rodríguez González, F. (1997) *Nuevo diccionario de anglicismos*. Madrid: Gredos.
- Rubens, P. (2001) *Science and Technical Writing: A Manual of Style*. London, UK: Routledge.
- Ruiz Garrido, M. (2003): *Aplicaciones pedagógicas del informe empresarial en el aula del inglés de los negocios: implementación de recursos discursivos escritos en contextos de especialidad*. Castellón: Universidad Jaime I.
- Ruiying, Y. & Allison, D. (2003) Research articles in applied linguistics: moving from results to conclusions. *English for Specific Purposes* Vol. 22 p 365-385.
- Ruiying, Y. & Allison, D. (2004) Research articles in applied linguistics: structures from a functional perspective. *English for Specific Purposes* 23 p. 264-279.
- Salager-Mayer, F. (1990) Discoursal flaws in medical English abstracts: a genre analysis per research-and text-type. *Text*. Vol. 10, p 365-384.
- Sandelin, B & Sarafoglou, N. Language and Scientific Publication Statistics: a Note. Forthcoming in *Language Problems & Language Planning*.
- Sanchez Benedito, F. (2003) *Gramática Inglesa*. Madrid: Pearson Longman.

- Santos Marinas, E. (2004) Formación de calcos nominales en la traducción de los evangelios en antiguo eslavo. Tesis doctoral. Madrid.
- Scott, M., (2004-2006) *Oxford WordSmith Tools*. Oxford: Oxford University Press.
- Schramper Azar, B. (1989) *Understanding and using English Grammar*. New Jersey: Prentice Hall Regents.
- Seco, M, Andrés, O. and Ramos, G. Aguilar (2000) Diccionario Abreviado del Español Actual. Barcelona.
- Seidlhofer, B. (2004). Annual Review of Applied Linguistics, 24, p.209-39. Cambridge: Cambridge University Press.
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal* Vol. 59/4, p. 339-341.
- Sharon, D. & Stohl, S.M., Communicating with International and Multinational Audiences. Downloaded from: <http://www.stc.org/confproceed/1993/PDFs/P60710.PDF> on 2-09-2007
- Sinclair, J. (1987) Collins Cobuild English Language Dictionary. London and Glasgow: William Collins Sons & Co Ltd.
- Sionis, C. (1995) Communication Strategies in the Writing of Scientific Research Articles by Non-native Users of English. *English for Specific Purposes*. Vol.14 n°2 p. 99-113.

- Skelton, J.R. & Edwards, S. (2000) The function of the discussion section in academic medical writing. *BMJ*, Vol. 320, p. 1269-1270.
- Smith, C. (1994) *Collin Spanish Dictionary*: New York. Harper Collins Publishers.
- Smith, C. (1998) *Collins Spanish-English, English-Spanish Dictionary* Barcelona: Grijalbo.
- Spalink, K. (2000) Improving Cost-Effectiveness in the Documentation Development Process through Integrated Translation. In: *Managing global communication in science and technology*. New York: Edited by Peter, J Hager & H.J Scheiber. Wiley & Sons.
- Stig, J. (2007) Using Multilingual Corpora. In *Seeing through Multilingual Corpora: On the Use of Corpora in Contrastive Studies*. Amsterdam, NLD: John Benjamins p.23.
- Stubbs, M. (1996). *Text and Corpus Analysis*. London: Blackwell.
- Swales, J.M. (1990) *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales. J.M. (2006) *Corpus Linguistics and English for Academic Purposes* in Information Technology in

- Languages for Specific Purposes: p.19-33 E. Arnó Macià, A. Soler Cervera and C. Rueda Ramos: Springer.
- Swan, M. (1984). *Basic English Usage*. Oxford: Oxford University Press.
- Tardy, C. (2004) The role of English in scientific communication: lingua franca or Tyrannosaurus rex? *Journal of English for Academic Purposes* 3 p 247-269.
- Thomas, J. & Short, M. (1996) Using Corpora for Language Research. New York: Longman Publishing.
- Timo-Ilaria C. (1998) La catástrofe del monolingüismo anglófono. *Actas Dermosifilográficas* nº 89, 566-70.
- Trujillo Sáez, F. (2001) Escritura y cultura: La retórica contrastiva. *Pragmalingüística*, 8-9, p. 313-329.
- Valero-Garcés, C. (1996) Contrastive ESP rhetoric: Metatext in Spanish-English Economic Texts. *English for Specific Purposes* 15/4, 279-94.
- Vanderbroucke, J.P. (1989) On not being born a native speaker of English. *Br Med J*; 298, 1461-2.
- Vázquez y del Árbol, E. (2006) La Redacción y traducción biomédica (inglés-español). Granada: Universidad de Granada.
- Watson, J.D., Crick, F.H.C. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acids. *Nature* 171, 737-738.

- Webster's New Twentieth Century Dictionary. 1955. Cleveland, Ohio.
- Williams, I. (1999) Results Sections of Medical Research Articles: Analysis of Rhetorical Categories for Pedagogical Purposes. *English for Specific Purposes*, Vol. 18 N°4 p.347-366.
- Wood, A. (1997) International scientific English: Some thoughts on science, language and ownership. Science Tribune-Article. Downloaded from: <http://www.tribunes.com/tribune/art97/wooda.htm> on 13-04-2007.
- Zainqui, J.M. (1989) Diccionario razonado de sinónimos y contrarios. Barcelona. Editorial De Vecchi, S.A.

ABBREVIATIONS

Abbreviations

L1: Mother tongue

L2: Second language

IMRD: Introduction-Method-Results-Discussion

RA.: Research Article

RAs: Research Articles

SSCI: Social Science Citation Index

ISI: Institute for Scientific Information

Contrastive Rhetoric: CR

JCR: Journal Citation Report

IF: Impact Factor

SPSS: Statistical Package for Social Sciences

Real Academia: Real Academia de la Lengua Española

DRAE: Diccionario de la Real Academia Española

SUMMARY OF THE DISSERTATION

Nowadays English is the means of communication and dissemination of scientific results, but not all the authors of scientific papers are native speakers. As a consequence, there is a need to analyse the characteristics of academic writing for research articles.

In order to limit the research and to be able to offer a thorough analysis, the hypothesis of the present research work is focused on Conclusion Sections and on the field of medicine.

The starting hypothesis has been to analyse the need on the part of medical doctors in order to reach linguistic skills when writing their papers in English. In the present research the lack of command of the English language on the part of Spanish medical doctors is taken for granted.

With the aim of corroborating the disadvantage of Spanish doctors with respect to German doctors when writing academic articles the first step was the elaboration of a survey addressed to Spanish and German researchers in the field of medicine. Once the need for a linguistic support on the part of the Spanish doctors was proven, two objectives were established:

- to present the general patterns for scientific academic writing in the field of medicine
- to analyse the most recurrent structures in the Conclusion section of medical research articles both in English and Spanish.
- to carry out a comparative study between the languages focusing on the structures at a semantic, grammatical and syntactical level supported by the tool WordSmith™. In this

way equivalences can be established: literal equivalences, non literal equivalences and lack of equivalence.

- additionally, the intrusion of neologisms in the field of medicine in the last ten years is analysed.

Literature on academic writing, contrastive rhetoric, corpus linguistics and technical translation has been used as the base for the analysis and processing of the corpus.

Once stated the state of the art several objectives were established. In order to reach the objectives a corpus was created which was made up of 311 Spanish Conclusion sections from medical research articles and 408 English Conclusion sections. All the articles belonged to the specialties of Paediatrics, Cardiology and Psychiatry.

The aim of the present research work is to establish the patterns to make the authors aware of them and consider them when writing about their results. In this way their articles will become linguistically appropriate in terms of scientific terminology and scientific structures. These patterns and structures should be in keeping with the international language in academic language, which is English.

RESUMEN DE LA TESIS

En la actualidad el inglés es el medio de comunicación y difusión de los resultados científicos, sin embargo los autores no necesariamente son nativos. Ante este panorama surge la necesidad de analizar las características de la redacción académica para los artículos científicos.

Para restringir el campo de estudio y poder ofrecer un análisis más exhaustivo la hipótesis de este trabajo se centra en la sección de las “Conclusiones” y en el campo de la medicina.

La hipótesis de partida del presente estudio es analizar las necesidades de los facultativos médicos para alcanzar suficiente destreza a la hora de escribir los artículos científicos en inglés. En este trabajo se asume una falta de dominio del idioma inglés por parte de los médicos españoles.

Con la finalidad de constatar la situación de desventaja de los médicos españoles respecto a los alemanes a la hora de redactar artículos de investigación el primer paso fue la elaboración de una encuesta dirigida a investigadores españoles y alemanes en el campo de la medicina. Tras observar la evidente necesidad de apoyo lingüístico de los médicos españoles se establecieron como objetivos los siguientes puntos:

- presentar las pautas generales para la redacción científico-técnica en medicina
- analizar las estructuras más recurrentes en el apartado de “Conclusiones” de los artículos de investigación médicos tanto en inglés como en español
- llevar a cabo un análisis comparativo a nivel semántico, gramatical y sintáctico de las estructuras de ambos idiomas con

la ayuda de la herramienta WordSmith™. De este modo se pueden establecer equivalencias en ambos idiomas tanto literales como no literales así como la falta de equivalencia.

- adicionalmente se lleva a cabo un estudio de la intrusión de neologismos en el lenguaje médico en los últimos años.

La literatura sobre redacción académica, retórica contrastiva, lingüística de corpus y traducción técnica se han usado como base para el análisis y procesado del corpus.

Una vez analizado el estado de la cuestión esta investigación se plantea distintos objetivos. Para llevar a cabo estos objetivos se parte de la elaboración de un corpus formado por 311 secciones de “Conclusión” en español extraídas de artículos de medicina de las especialidades de Pediatría, Cardiología y Psiquiatría y 408 secciones de “Conclusión” en inglés pertenecientes a las mismas áreas.

Con esta tesis se pretende extraer las pautas para una concienciación de los autores de cara a la difusión de sus resultados de manera que sus artículos sean lingüísticamente adecuados a la terminología científica y estructuras científicas características del idioma internacional, el inglés.

RESUM DE LA TESI

En l'actualitat l'anglés és el mitjà de comunicació i difusió dels resultats científics, no obstant això els autors no tots són nadius. Davant d'este panorama sorgix la necessitat d'analitzar les característiques de la redacció acadèmica per als articles científics.

Per a restringir el camp d'estudi i poder oferir una anàlisi més exhaustiva, la hipòtesi d'este treball se centra en la secció de les "Conclusions" i en el camp de la medicina.

La hipòtesi de partida del present estudi és analitzar les necessitats dels facultatius mèdics per a aconseguir la suficient destresa a l'hora d'escriure els articles científics en anglés. En este treball s'assumix una falta de domini de l'idioma anglés per part dels metges espanyols.

Amb la finalitat de constatar la situació de desavantatge dels metges espanyols respecte als alemanys a l'hora de redactar articles d'investigació, el primer pas fou l'elaboració d'una enquesta dirigida a investigadors espanyols i alemanys en el camp de la medicina. Després d'observar l'evident necessitat d'un suport lingüístic dels metges espanyols es van establir com a objectius els punts següents:

- presentar les pautes generals per a la redacció científic-tècnica en medicina
- analitzar les estructures més recurrents, en l'apartat de "Conclusions", dels articles d'investigació mèdics tant en anglés com en espanyol
- dur a terme una anàlisi comparativa a nivell semàntic, gramatical i sintàctic de les estructures d'ambdós idiomes amb

l'ajuda de la ferramenta WordSmith™. D'esta manera es poden establir equivalències en ambdós idiomes tant literals com no literals, així com la falta d'equivalència.

- addicionalment es du a terme un estudi de la intrusió de neologismes en el llenguatge mèdic en els últims anys.

La literatura sobre redacció acadèmica, retòrica contrastiva, lingüística de corpus i traducció tècnica s'han usat com a base per a l'anàlisi i processat del corpus.

Una vegada analitzat l'estat de la qüestió, esta investigació es planteja distints objectius. Per a dur a terme estos objectius es partix de l'elaboració d'un corpus format per 311 seccions de “Conclusió” en espanyol extretes d'articles de medicina de les especialitats de Pediatria, Cardiologia i Psiquiatria, i 408 seccions de “Conclusió” en anglés pertanyents a les mateixes àrees.

Amb esta tesi es pretén extraure les pautes per a una conscienciació dels autors de cara a la difusió dels seus resultats, de manera que els seus articles siguen lingüísticament adequats a la terminologia científica i estructures científiques característiques de l'idioma internacional, l'anglés.

Appendixes (attached CD)

I (Corpus)

II (Surveys)

III (Lists)

