2-1-2010

A corpus analysis of stance marker use in international and Egyptian medical research articles

Sarah Ahmed Seleem

The American University in Cairo

School of Humanities and Social Sciences

A Corpus Analysis of Stance Marker Use in International and Egyptian Medical Research Articles

A Thesis Submitted to

The Department of Teaching English to Speakers of Other Languages (TESOL)

In Partial Fulfillment of the Requirements

For the Degree of Master of Arts

By

Sarah Ahmed Seleem

May 2013

The American University in Cairo

School of Humanities and Social Sciences

A Corpus Analysis of Stance Marker Use in International and Egyptian Medical Research Articles

A Thesis Submitted to

The Department of Teaching English to Speakers of Other Languages (TESOL)

May 2013

In Partial Fulfillment of the Requirements

For the Degree of Master of Arts

By

Sarah Ahmed Seleem

Under the Supervision of

Pr. Amira Agameya
Thesis Committee Advisor_____
Affiliation: The American University in Cairo

Dr. Lori Fredricks
Thesis Committee Reader/ Examiner _____
Affiliation: The American University in Cairo

Dr. Marilyn Plumlee
Thesis Committee Reader/ Examiner_____
Affiliation: The American University in Cairo

_____ _____ _____ _____

Dept. Chair Date Dean of HUSS Date

Acknowledgments

I would like to express my sincere appreciation to Dr. Amira Agameya for her valuable and constructive suggestions before and during the planning and development of this research work. Her willingness to give her time so generously has been very much appreciated. I would also like to thank her for being so keen to keep my progress on schedule. Dr. Amira, no words of gratitude can express how I feel. Working with you not only enriched my knowledge, but it also taught me how I should be with my future students.

I am also very thankful to Dr. Lori Fredricks whose feedback has enriched my study, starting from the feasibility stage. Your boosting words and encouraging feedback have always motivated me to work harder, hoping to meet your expectations. Your time and effort is really appreciated. I am also grateful to Dr. Marilyne Plumlee for her remarks on my several drafts. It is to her credit that my thesis is out in a neat form.

I would like to express my deep gratitude to Pr. Deena Boraie for helping me getting my feasibility paper accomplished. I cannot imagine myself writing a piece of research without her guidance. Dr. Deena, what you have taught me will remain for the years to come, and I shall be always grateful.

I would also like to sincerely thank Dr. Atta Gebril for his very helpful suggestions concerning the research design. The meticulous details that he directed my attention to rendered this study to be a well designed piece of research. Special thanks are extended to Dr. Robert Williams for his helpful feedback in the feasibility stage which taught me to be a cautious researcher.

I would like to extend my thanks to Ahmed Awwaad who helped me a lot with the material that helped me start my research, and the rest of my colleagues in the Proposal Writing course, Wessam Ibraheem, Rebecca Jones, Omar Abouelazm and Ghada Ghobara. Your feedback on my presentations helped me shape my thesis and I definitely learned from each one of you. Special thanks go to Saliha Arseven whose moral support in the times when I felt down kept me on track. I was lucky to have such a supportive colleague.

Finally, I wish to thank my parents for their support and encouragement throughout my study, especially my unique father and patron, Ahmad Seleem and my wonderful mother, Amani Farag, who has been such a nice and caring grandmother during my absence for classes. I wish also to express my special thanks to my dear husband, Khaled Badawi, for being such an understanding and an encouraging husband, and for bearing with me throughout my MA years.

Abstract

The need for *stance* expression and making the position of the writers clear about what they are reporting is well documented in the literature (Biber, 2006; Kelly & Bazerman, 2003; Hyland, 2005; 2008; Molino, 2010; Myers, 1989; Williams, 2006). Even *hard knowledge* disciplines that were traditionally expected to sound objective and detached employ *stance* projection strategies by which researchers express their opinions, degree of certainty or ownership of the different claims that are stated in their work (Harwood, 2005 ; Hyland, 2005).

Authors of research who are writing in English as their second language (L2) are reported to find difficulty in making their voice heard in their research articles (RAs) (Flowerdew, 2001). Differences in the frequency and the accuracy of using stance markers in RAs written by L2 writers of research and native speakers (NSs) have been detected in numerous studies (Hendersson & Barr, 2010; Hyland, 2002; Jafarpour & Taki, 2012; Maurannen, 1993 ; Molino, 2010 ; Orta, 2010).

Meeting the needs of almost 30% of research producers at Cairo University, this study detects the differences between *stance* markers in Egyptian RAs compared to internationally published RAs in the field of medicine. Differences are diagnosed and patterns of the "acceptable" use of stance markers are listed in order to help Egyptian doctors write more professionally and gain acceptance in international publishing.

In this corpus-based study, 47 RAs published in local Egyptian medical journals representing different medical schools and institutions across the country were examined and compared to the Corpus of Contemporary American English (COCA). The use of direct self-reference using first person pronouns *I, me, my, we us* and *our* compared to the more impersonal *"it…that"* structure was examined in both corpora.

The study showed a tendency in the Egyptian RAs to sound more distant and cautious. The use of first person pronouns in Egyptian RAs was generally less frequent. Egyptian medical researchers avoided using the singular first person pronoun in their RAs but they sometimes directly referred to themselves when they wrote in a group. Egyptian researchers also showed higher frequency of the more mitigated and impersonal structures such as *"it…that"* structures, the passive structure and doubt adverbs.

On the other hand, Egyptian researchers showed awareness with the preferred verb tenses collocating with the first person pronoun *we* as well as an awareness of the different structures of the *"it…that"* structures despite its complexity. The Egyptian Medical Research Articles corpus (EMRA) showed a lack of variety in some of the lexical collocates of the first person pronouns as well as *"it….that"* structures.

Contents

Chapter 1: Introduction

I. Introduction and Background…….1
II. Research Problem…….....3
III. Research Questions…….....7
IV. Definitions…….....8
Stance
Evaluation…..……..……8
a. Operational definition of the construct....…….9
b. Definitions of the Variables……..…....9
V. Delimitations of the Study…….9

Chapter 2: Literature Review……..…….....11

I. Genre Analysis……..…..11
a. What is Genre Analysis……..……...11
b. Genre Analysis in ESP/EAP Contexts…….…14
c. Stance in Academic Discourse…….…15
i. Self-mention……...…….....19
ii. Existential "*it*"…….....20
iii. Stance and Disciplinary Differences……..……..24
iv. Stance across Cultures……...28
v. The Problem of Writing in English as an L2….…….....37
vi. Conclusion…….....38

Chapter 3: Methodology……40

I. Design…….……40
II. The Corpora…….....40
III. Corpus Compilation Procedures……...……41
Compiling Process…….....41
IV. Answering the Research Questions…….……42
V. Procedures…….....45
VI. Instrumentation…….……..…….....45
VII. Methods of Reporting the Results…….....45

Chapter 4: Stance Markers in Medical RAs…….46
A. Use of First Person Pronoun in COCA……..47
I. First Person Singular Pronouns…….....47

a. The First Person Pronoun for the Singular Subject *I* ,*Me* and *My*…………………………………………………………...47
b. Tenses Collocating with *I*…………………………………..51
c. . The First person Singular Possessive *My*………...……...53
II. First Person Plural Pronouns *We, Us* and *Our*……...……..54
a. Plural Personal Pronouns for the Subject and Object *We* & *Us*……………………………………………………….……..54
b. Tenses Collocating with *we*.…………………………...…...............55
III. Functions of the First Person Pronouns in Medical RAs…………………………………………….................58

B. "*It…..that*" Structures…………………………………………….............67
I. "It * that"……………………………………...…...........70
II. "It * * that"…………………………………………….......73
III. "It…..that" Constructions with more than Two Slots…………………………………………………….......74

Chapter 5: Egyptian RAs and Internationally Published RAs Compared.................77
A. Comparing the Frequency and Use of Personal Pronouns in EMRA and COCA……………………………………………………….………….........77
I. First Person Singular *I, Me* and *My*.…………………………......……..….77
II. Plural Personal Pronoun *We, Us and Our*.……………………..…..............78
III. Collocates of First Person Pronoun in COCA and EMRA……..………...81
a. Nouns Collocating with *Our*.…………………………………………….....81
b. Collocates with *we*.………………………………………….....................83
IV. Functions of the First Person Pronouns in EMRA……………………...86

B. Existential "*It….that*" Structure……………………………………….…...90

I. Syntactic Structures of *"it…is"* inEMRA……………………….……...…92
C. Other Stance Markers……………………………………………….......94
I. Certainty and Doubt Adverbs……………………...…...................95
II. Passive Voice……………………………………………………...97

Chapter 6: Discussion……………………………………………………...…...99

- Patterns of Stance Marking in Standard Medical RAs COCA…………….99
- Differences between Egyptian Medical Research Articles EMRA and the Standard Production in COCA…………..………………………........101
- Implications for Teaching………………………...……………......…..103
- Limitations of the Study……………………………………...……..……104

- Suggestions for Further Research…….……...…..…….105

List of Tables:

Table 4.1: *Verb tenses used with "I"*……..51

Table 4.2 *Tenses collocating with "we" in COCA.*……..…56

Table 4.3: *Different structures of "it….that" matrix as occurred in COCA.*……...69

Table 5.1: *Comparing the use of first person pronouns in COCA and EMRA.*……...79

Table 5.2: *Comparing the top noun collocates with the pronoun "our" in COCA and EMRA.*……..82

Table 5.3: *Top verb-collocates with "we" in COCA and EMRA compared.*……....85

Table 5.4: *Frequencies of the tenses of the verbs collocating with "we"*…….86

Table 5.5: *Comparing the frequency of each structure of "it …that" in COCA and EMRA.*……....91

Table 5.6: *Adverbs expressing certainty of claims in COCA and EMRA.*…….....95

Table 5.7: *Adverbs expressing doubt in COCA and EMRA.*……..96

Table 5.8: *Frequencies of the passive structure in different tenses in both COCA and EMRA.*…….....97

List of Figures:

Figure 4.1: *The frequency of "I", "me" and "my" in the different academic disciplines in COCA*…….48

Figure 4.2: *First Person Singular pronoun "I" in COCA: Comparing different disciplines.*…….50

Figure 4.3: *We in medical RAs as compared to other disciplines in COCA.*……..…54

Figure 4.4: *The frequency of the one-slot "I * that" clause compared in the different sub-sections of COC.*……..71

Figure 4.5: *Frequencies of the "it * * that" construction in the different COCA sub-sections.*……....74

Figure 4.6: *The frequencies of the different constructions of the "it …that" matrix in COCA.*……..75

Figure 5.1: *Frequencies of the different structures of "**it …is"** in COCA and EMRA.*……....*94*

Abbreviations and Terms

RA/s: Research article/s.

COCA: Corpus of Contemporary American English.

EMRA: Corpus of Egyptian Medical Research Articles.

NS: Native speaker.

NNS/s: Non-native speaker/s.

L1: First language.

L2: Second Language.

ESP: English for specific purposes.

EAP: English for academic purposes.

Ppm: Parts per-million.

Chapter One

Introduction

I. Introduction and Background

The academic research article (RA) is a specific form of argumentation that has emerged in the modern world as a means of circulating information and sharing knowledge among the members of each scientific community. Academic writing was traditionally regarded as an objective genre where the writer's personal voice should be totally absent. It is thought that the academic research article's main function is to deliver knowledge and to represent findings reached through empirical evidence that is obtained by following flawless procedures. However, as it is documented in the literature, scientific research is a rhetorical genre and is meant to be persuasive (Kelly & Bazerman, 2003; Hyland, 2005; 2008; Molino, 2010; Myers, 1989; Williams, 2006).

The image of the writers of the research article playing the role of the transmitter soft the truth is no longer that of the detached " humble servants of the discipline" (Hyland, 2001, p.207). In fact, RA writers at the present time are expected to interact with their readers and make them feel the writers' presence and what position they are taking from what they are reporting. Biber (2006) concluded in his study that stance is expressed in all academic registers. Also, Flowerdew (2001) reported in his study that one of the journal editors he interviewed said that one of the problems he noticed among NNS researchers who submitted their papers to journals he reviewed was the inability to show an authorial voice in the text they write. NNS writing lacked the evaluative quality that made the whole work sound more credible and persuasive.

This phenomenon of not showing a voice in the written text could be explained by suggesting that citing other researchers' work in the RA is a face threatening act (FTA) that heightens the imposition on other members of the field. Challenging other researchers' work or the results of previous studies that are agreed upon is also threatening to the face of the writer (Myers, 1989). Sometimes, the new issue of a scientific journal might contain an article that renders the hard work of previous researchers useless because of findings that prove the opposite. Presenting this new knowledge to other researchers should be done skillfully, yet within the acceptable norms of the field. So, in research writing, a balance between a humble tone that does not threaten the well-established knowledge in the field and the novelty that the writer is offering and is trying to justify the urge to publish it is needed.

Stance marking, generally, does not denote lack of certainty or probability, as much as it is a strategy of social interaction in shared written texts (Myers, 1989). Therefore, it is important to give L2 writers the different pragmatic options that English offers them so that they can choose what suits their individual differences within a generally accepted framework (Eslami-Rasekh, 2005).

But what exactly decides the stance expression of the written product? In examining stance in academic writing, and more specifically RAs, researchers look at linguistic aspects that make the academic text revealing their stance. Myers (1989) argues that there are some linguistic features that are used as "rational strategies for dealing with the social interactions involved in publishing an article" (p.3). Aspects like using the first person pronoun (Harwood, 2005; Hyland, 2002b), reporting verbs (Bloch, 2010; Thopmson &Ye, 1991) and hedges and boosters (Hyland 1998; Lewin,

2005), pay credit to the writer as well as the other authors being cited and also give a hint on where exactly the writers stand from what they claim. Texts are persuasive only when they employ rhetorical conventions that colleagues find convincing (Hyland, 2008; Hyland &Tse, 2005).

The problem that faces researchers who are interested in studying stance markers in academic language is whether there are fit-for-all formulas to be applied in all contexts or not. The literature shows that each discipline has its own conventions of writing and projecting the degree of certainty and credibility, and novice writers of research must learn those conventions in order to be able to show skill in using the accepted norm of language among the society of researchers they are addressing. This awareness helps them gain acceptance in the discipline they belong to, that is, to be able to publish their work. Gaining acceptance in a disciplinary community implies that a researcher abides by the norms of this community. This helps them also show the gatekeepers of the specific discipline (editors and reviewers of journals) that they have a good potential to be members of the community (Hyland, 1998; 2006). As they write RAs in a language that is not their native language, Egyptian researchers may find difficulty in getting internationally published. As linguists working in Egypt, paying attention to the needs of Egyptian researchers will help boost publishing, hence upgrading Egyptian universities ranking.

II. Research Problem

Considering the Egyptian context, the medical school of Cairo University, Kasr EL-Ainy, and its specialized university hospital units and centers has over 3200

members of the faculty. The total number of faculty members in Cairo University is around 10,500 members, with Kasr El-Ainy members representing almost one third of the faculty members in Cairo University at large. According to the University bylaws, for those faculty members to get promoted, they need to publish a certain number of research papers, with each article published internationally counting for two papers of the ones published locally. This is a strategy that the University administration has started so that researchers become more eager to publish in international peer-reviewed journals. This adds to the credibility of Cairo University with regards to ranking among universities world-wide. To the best of my knowledge, there are no studies that were conducted to examine the language of medical research articles in the Egyptian context form an ESP or an EAP point of view. Meaning that there is no "visible pedagogy" available on how to write according to acceptable norms. The approach of diagnosing the problematic areas of L2 writers and comparing them to the mainstream or the standard norm to formulate material that could be taught is referred to as "visible pedagogy,'' (Johns, et al, 2006, p. 237). This study is an attempt to start a series of research projects to create data-driven material to the members of one of the widest populations in the society of Egyptian researchers. This data-driven material is intended to be used for training Egyptian medical researchers-to-be to better express their stance in the internationally acceptable norms. The results are to be either compiled in a research writing manual or implemented in a writing training course. Giving suggestions for how to use the findings of this study is recommended.

In an attempt to find studies that provide pedagogical implications for researchers who write in English as a second Language (L2) in the medical field to use, I conducted

a search on the AUC library database, Science Direct and Google Scholar. There were no results for a recent study analyzing the problem of evaluative markers in medical research articles in particular. The studies that I found that examined evaluative and personalized academic language used in medical research articles were mainly those of Salager-Meyer (2001), Skelton (1997) and Williams (2006). As seen from the years of publication, the only one that is relatively recent is Williams (2006). The missing aspect in the study of Williams (2006), however, is that it was not a data-driven study. It only presented tendencies in evaluation in RAs in general, which is a dangerous endeavor because of the disciplinary differences mentioned earlier. Given that evaluation and voice in medical research writing are evolutionary (i.e. changes in the strategies and frequencies of expressing the presence of the researchers change over time), the need for a more up-to-date study which directly addresses medical academic language is needed (Li & Ge, 2009; Tas, 2010).

As for the comparative dimension, it seems that there are no comparative studies comparing the medical RAs that are published locally with internationally published RAs. Only El Malik and Nesi (2008) tackled *hedging* and *nominalization* as a side point to the moves analysis they conducted on medical research papers. They detected a preference for *hedging* and *nominalization* among the British RAs which they were comparing the Sudanese written texts with. However, this finding was not elaborated in a clear statement of what pedagogical application could be extracted from the study as it was not the main focus of the study.

In an informal interview with two post-doctoral researchers at the Faculty of Medicine, Cairo University, they reported that the most popular way to learn about the

appropriate writing styles and conventions is through imitation. This is done based on individualistic assumptions and not based on research. MA and PhD students are advised to use the first person freely for instance. They are told that this is the way to add more credibility to what they present. One of the researchers said that "[they] have no set rules that tell us how". Theses and research writers pick some articles from prestigious journals and try to follow the same structural and evaluative patterns. This method of learning, i.e. imitation, was also reported in Cargill, O'Connor and Li (2011). This highlights the importance of creating a data-driven model of stance expression those novice researchers (or researchers who seek acceptance to the community of the international publishers) can follow (Hunston, 2002).

The problem with L2 writers is twofold as they not only have to convey the content; they also have to conform to the style and format that is expected (Hyland, 2005). This need is even more imperative with the nonnative speakers of English being driven by the "publish or perish" motive and having to publish in the language that is most used in circulating knowledge, namely; English (Mauranen, 1993; Swales,1997).

The results of this study are intended to be used for a pedagogical purpose. That is, documenting data from authentic texts to use them as teaching tips or implement them in technical writing courses in order to help researchers in Egypt to be up to the standards of publishing internationally is intended. For the purpose of this study, I examine samples from medical articles published in international peer-reviewed journals and samples of medical articles that are published in Egypt in locally published journals. A comparison between the two corpora (the Egyptian and the international) was done both qualitatively and quantitatively.

Based on the literature, this study is expected to identify the areas of difference between medical research articles written by Egyptian researchers published in Egyptian journals and other internationally published research articles. The novelty of this study is that it explores an area that has never been explored in the Egyptian context. The differences in frequencies and in uses of stance markers that are examined in this study are expected to add to the literature. It is noteworthy that in piloting stance markers on a sample of Egyptian RAs and International RAs in Spring 2012, the hypothesis that there are differences between both samples was confirmed. That is, differences in the frequency of each stance marker that was studied were evident in the sample. The Egyptian locally published sample had more hedges than the international sample. The locally published sample also contained less variety in the reporting verbs that were used (*showed, demonstrated and reported),* with *showed* and *reported* used recurrently throughout the article.

III- Research Questions

This study aims to address the following questions:

1- What are the techniques of expressing stance that are specific to medical research articles?

2- What are the differences in the use of stance expressions between research articles published in international peer-reviewed journals and locally published Egyptian research articles?

III. Definitions

a. Theoretical Definition of the Construct *Stance*

A definition of the main construct under investigation is in order here. The main construct in the study is *stance. Stance* is the viewpoint or the opinion of the researcher regarding what they report in their writing, whether it is other researcher's work, facts that are generally agreed upon or the results of their own work. This opinion or point of view has been referred to in the literature by different terms e.g. "appraisal", "sentiment", "attitude" and "evaluation". However, the most recurrent term that I have encountered in the literature is the one used by Hyland, (2005) which is *stance.* Hyland (2005) says about stance:

> They express a textual 'voice' or community recognized personality which, following others, I shall call *stance*. This can be seen as an attitudinal dimension and includes features which refer to the ways writers present themselves and convey their judgments, opinions, and commitments. It is the ways that writers intrude to stamp their personal authority onto their arguments or step back and disguise their involvement. (p.176)

Evaluation

The terms *evaluation* and *stance* have been used interchangeably and seem to overlap in most of the cases. They are also described as "confusing" (Hunston & Thompson, 2000, p.2). Whereas Hyland (2005) recurrently used *evaluation* as an alternative for *stance*, both Hunston and Thompson argue that the term *evaluation* is more super-ordinate and comprehensive (Hunston & Thompson, 2002; Hunston, 2011). Since the linguistic markers that are under investigation in this study are closer to the term

stance, defined as the "personal feelings or assessments" (Biber, 2004, p.107), the term used in the current study is *"stance"* rather than the more general term "evaluation".

a. Operational Definition of the Construct

Lexical items that reveal the attitude of the writers or the evaluation tendency such as *self-mention* and *existential "it"* as used in internationally published RAs and Egyptian locally published RAs were examined. Those terms were collected from RAs sampling peer-reviewed journals, and RAs sampling locally published Egyptian journals.

b. Definition of the Variables

- **Self-mention**: Refers to the use of the personal pronoun, either in the singular or in the plural forms to project one's own opinion or to share interpersonal information (Hyland, 2005).

- **Existential *"it"***: Refers to the construction where "the theme position of the matrix clause is filled by an anticipatory 'it', which refers to a clausal item occurring in the postponed extraposed clause" (Boyd, Gegg-Harrison & Byron, 2005, p. 12).

VI. Delimitations of the Study

This study aims at investigating strategies of projecting evaluation in medical research articles. It only analyzes papers published in international medical journals that are compiled in (COCA) compared to medical research articles that are published in Egypt (EMRA). It will only focus on the evaluative language categories stated in the instrumentation section. The whole articles were uploaded on a concordancing program;

MonoConc Pro 2.2. Therefore, the whole articles were examined and not specific sections.

Chapter Two

Literature Review

Organization of the Literature Review

In this literature review, previous research relating to the same topic is presented, ordered thematically. The review starts with setting the study in its broader theoretical framework, namely, *genre analysis*. In this section a definition of *genre* is stated as well as how this theory was realized in examining academic writing. The following section examines the notion of *stance* and *evaluation*. The final two sections present previous studies on the two core concepts in this study, namely; *disciplinary differences* and *differences across cultures.*

I. Genre Analysis

a. What is Genre Analysis?

For the past two decades, *genre analysis* has been the dominant approach in dealing with academic texts in what is known as English for Academic Purposes (EAP) (Dudley-Evans, 2000). It started with the study of Swales (1990) in which he defines *genre* as a class of communicative events in a community whose members share the same communicative purposes. These purposes, as well as the rationale of the communicative events, are decided by the expert members of these communities.

The rationale of the events decides the choice of the structures and styles as well as the content of those *genres.* Dudley-Evans and St John (1998) describe *genre analysis* as "the study of the structural and linguistic regularities of particular genres or text types and the role they play within a discourse community" (1998: xv) as cited in Paltridge (2004).

Paltridge, in Jonhs et. al. (2006), defines genre as "ways in which people 'get things done' through their use of language in particular contexts." (Johns et al., 2006, p.235). Paltridge adds that in following a specific genre, the writers adjust their use of the language to fit the aim and purpose of this genre, suiting also the relation between the writers and readers that is normally accepted in this specific genre. "The way writers use language in a genre also depends on the expectations of the context in which the genre is being produced. Writers also draw on their previous experiences with the genre to produce a new text" (Johns, et al. 2006, p.235). A *genre* is simply described as "a kind of text" that has certain functions and purposes and employs specific conventions (Paltridge, 2006, p. 84)

Therefore, *genre analysis* assumes that texts of the same genre have certain characteristics that distinguish them from other texts of other genres and disciplines (Dudley-Evans, 2000). The implementation of *genre analysis* is an approach to academic writing that examines the context in which written texts were created. This approach to written texts is important as context is one of the factors that contribute to shaping the final product (Hyon, 1996). Researchers are "driven to look for relationships between language and what is not language" (Paltridge, 1994, p. 296). For example, the type of language of research articles in a certain discipline will definitely be different from the type of language used for textbooks in the same discipline (Paxton, Pletzen, Archer, Arend & Chihota, 2008).The example given by Dudley-Evans (2000) to make the idea of genre clearer was that of the difference between the uses of hedges used in two different genres, despite the fact that they were both communicating medical information. The two genres compared were prestigious academic research articles versus the popular scientific magazines such as *Scientific American* and *New Scientist.* Analyzing both samples revealed that differences in

the use of hedges occurred between the two texts. The source of these differences is that they do not share the same purpose nor do they have the same intended audience.

However, *genre analysis* is not as rigid as it might sound. On the contrary, the application of Swales's (1990) model of *genre analysis* was not expected to be very confining. That is, *genre analysis* just sets a prototypical model for each genre in each discipline, such as Create a Research Space (CARS) as a model of organizing the texts of the research articles for instance (Dudley-Evans, 2000). The individual realizations are measured by how much they adhere to the model. This means that there is room for limited individual divergence from the set model (Dudley-Evans, 2000; Hyland, 2005; Thompson and Ye, 1991).

Following the foundation of the genre analysis approach was the *moves analysis* tradition that was founded by Swales (Paltridge, 2004). This model for approaching academic writing was dominant and received a great deal of the researchers' attention in the field of English for Academic Purposes (EAP), starting from the Nineties onwards. This approach was employed to identify the sections and subsections of research articles like the studies of Basterkmen (2011), Kanoksilapatham (2005), Nwogu (1997) and Williams (1999). In these studies, samples from research articles in dentistry and medicine were examined. *Moves analysis* was applied and the authors concluded that there were distinct patterns that govern certain moves in the RAs in each discipline. This approach is concerned more with the formal dimension of RAs. It describes how researchers should organize information inside a research article in a way that makes it adhere to the conventions of this particular discipline in this genre. However, fewer studies were done on analyzing the language that the researchers use in negotiating their arguments with the

academic community. Of these language devices are the use of tenses, cohesive markers and evaluative markers.

b. Genre Analysis in ESP/EAP Contexts

The trend of *genre analysis* has been of much interest to the field of EAP and ESP. It is mainly directed to identifying the norms of writing in each discipline that is expected from writers who want to belong to a certain research community (Paltridge, 2004). This makes the process of making the writing conventions of a certain genre salient for writers of academic English as an L2 an easier task. Huang (2010) held interviews with Taiwanese PhD students from eight scientific disciplines and ten different institutions. The participants reported that they found themselves less privileged compared to their English-speaking counterparts. They also reported that they feel an extra burden in writing and communicating the content they want to share in the process of publishing their work internationally because they have to write it in English (Huang, 2010). They complained about their lack of English repertoire which makes writing in English for an international audience-and possibly a native speaker of English-a challenging task (Huang, 2011). Paltridge (2004) cites Prior (1995) who said that writers of academic texts are usually given advice about the positioning as well as organization of the texts they are producing. This typically reflects the Egyptian situation where researchers who do not receive academic writing training are asked, nevertheless, to be up to the writing standards of prestigious international peer-reviewed journals, a problem that is to be described in detail later. So, according to Paltridge (2004), the needs of the researchers who write in English as an L2 have to be acknowledged in the research of EAP and ESP if they are to be asked to meet certain standards. The stylistic pitfalls that they fall into must be addressed and the more

internationally accepted writing conventions should be made salient for them. This is where the importance of applying *genre analysis* for teaching English for academia lies.

As Hyon (1996) puts it, the study of genre is meant to empower students with better writing skills in order to be more successful. Although he speaks of students receiving instruction on writing in general, it can be well applied to the situation of NNS or writers whose first language is not English. Also, Hyland (2002a) states clearly that the study of *genre* or *specificity* as he names it, adds to the success of ESP pedagogy. Hinkel (1994) even says that the lack of familiarity with the conventions of writing in English and following what is regarded as the acceptable norms (speaking specifically of pragmatic conventions) may threaten the academic and professional opportunities of the L2 researchers.

c. **Stance in Academic Discourse**

In any discourse situation, we not only convey content, but we also convey our views towards that content and our relations to the audience we are addressing, either explicitly or implicitly (Riellya, Zamora & McGiverna, 2004; Precht, 2004). The argument that academic language expresses stance might sound paradoxical to those who regard academic language as “objective” and “distanced”. The current situation is that academic writing is no longer regarded as a mechanical process in which the writers only present knowledge, with no reference to their views or where they stand with respect to what they share (Hyland, 2005, 2008; MacIntyre, 2010; Tas, 2010).

Hyland, as cited in Johns et.al. (2006), defines stance as “the ways that writers intrude into texts to stamp their personal authority onto their arguments or, alternatively, to step back from their discourse. This writer’s disciplinary ‘‘voice’’ or personality includes

the ways writers present themselves to convey their judgments, opinions, and commitments to what they say" (Johns et. al., 2006, p. 238). Stance in academic writing also means the expression of "personal feelings and assessments" towards the knowledge that the researchers write about in their work (Biber, 2004, p.107). Speaking of stance here means the dimension of *evidentiality* of stance as described by Biber and Finnegan (1989). *Evidentiality* refers to the writers' expression of their attitudes towards what they write/report and its reliability. The twofold dimension of *discourse stance* that is studied in this research project is also named *epistemic stance* and *deontic stance.* Epistemic attitude expresses a relation between the speaker and the writer and the possibility, certainty, or evidence for the researchers' belief about the truth of what they reports. *Deontic attitude* entails a judgment or an evaluation of what is being reported (Berman, 2004).

Therefore, writers are now expected to represent themselves in their research papers, and self-representation strategies have become the subject of a great deal of research. They are also expected to interact with their readers in the appropriate manner that makes them sound credible and make them accepted in the disciplinary community they want to be admitted to. Hyland (2005) says:

> Successful academic writing thus depends on the individual writer's projection of a shared professional context. That is, in pursuing their personal and disciplinary goals, writers seek to create a recognizable social world through claims for the significance, originality and plausibility of their work rhetorical choices which allow them to conduct interpersonal negotiations and balance against the convictions and expectations of their readers (p.178).

So, the traditional vision of academic RAs ,especially scientific research articles, as only talking about facts and dry science has been proved to be misleading. "They are not just

talking about garlic proteins, stress fractures or brains in vats. Instead, they use language to acknowledge, construct and negotiate social relations" (Hyland, 2008, p.4).

According to the literature, the term *stance* is not the only term that is used to refer to the point of view of the writer. Other terms are used to refer to the same construct and to achieve the same goal; determining the writers' position (Jafarpoure & Taki, 2012). The two terms; *Stance* and *evaluations* particularly have been used interchangeably and seem to overlap in most of the cases (Marc-Silver, 2003; Hyland, 2008). They were also described by Hunston and Thompson (2006) as "confusing" (p.2). Whereas Hyland, (2005) recurrently used *stance* as an alternative for *evaluation*, Hunston and Thompson (2006) describe *evaluation* as the super-ordinate term for the term *stance* (Hunston & Thompson, 2006; Hunston, 2011).

Evaluation is the positioning of opinion or the point of view of the writer. Thompson and Hunston (2006) define *evaluation* as the broader term for the expression of writers' attitude, stance or feelings about entities or propositions that they are talking about in their written discourse. *Evaluation* is important as it servers three functions*:* "(a) to express the writer's opinion, and in doing so to reflect the value system of that person and their community, (b) to construct and maintain relations between the writer and the reader, and (c) to organize the discourse" (Thompson & Hunston, 2006, p.6). The notion of evaluation and stance in academic writing also refers to the interaction between the writers and their readers. This interaction has its origins in the notion of intertextuality of Bakhtin (1986) (as cited in Berkenkotter, 20011; Hyland, 2005). This theory views the writers and the readers of texts as part of a complex web which culture and shared beliefs are parts of. However, for the purpose of this study, I will use the narrower term *stance* as

the scope of the study only covers the dimension of *stance* in the sense that is mentioned earlier in this section.

The study of *stance* as a theoretical construct is not the main focus of the prior research produced on this topic as much as it was on the markers that signaled this attitude in written texts. The ideologies and points of view expressed in academic writing are usually delivered subtly through some linguistic signals that the researchers have to trace in any written text in order to be able to determine its stance. Stance can take many forms; it can be expressed lexically or grammatically and it can be implicit or explicit (Vasquez, 2007). Since Thompson and Ye's (1991) work on evaluative tendencies in reporting verbs, numerous studies have been conducted to identify evaluation through different stance markers. To the best of my knowledge, Hyland (2005) offers the most comprehensive model classifying interaction in RAs into both *stance* and *engagement.* Hyland (2005) highlights the importance of "positioning" or expressing where researchers stand from the information they are communicating or the persons they are reporting in their research articles. This expression of point of view can be achieved through certain structural categories that Hyland (2005) mentioned in his research which are *self-mention, hedges, boosters* and *attitude markers.* Other stance markers (either lexical or syntactic) were also examined in previous research, of which I investigate the first person pronouns (Harwood, 2005; Hyland, 2002b) and the *"it…..that"* structure (Berman, 2004; Biber & Finnegan, 1989; Carter-Thomas & Rowley-Jolivet, 2001; Myers, 1989; Reillya, et. al, 2004).

i. Self-mention

Self-mention is the use of first person pronouns and possessive adjectives to present propositional, affective and interpersonal information (Hyland, 2005, p.181). As subjective as it might seem in written academic context that is traditionally regarded as selfless, self-mention is important in academic writing as it demarcates the role of the researcher in his/her work (Taş, 2010; Hyland, 2005; Myers, 1989). This demarcation is important in highlighting an innovation to a conventional method that the researchers are following or a new finding that they are introducing (Martinez, 2001).

It might even be more surprising to know that the personal pronouns *I* and *we* are not only used to boost the researchers' presence in their work, but they are also used as a strategy to show politeness in the research article. It is sometimes used to mitigate the imposition on the reader of what is being said, by tracing it back to the writer, or by showing solidarity with the reader. The Personal pronoun *we* refers not only to multiple authors attributing what they say to themselves, but it also may refer to the discipline as a whole, creating a sense of solidarity between the writer/s and the reader by stressing the fact that they all belong to the same field (Myers, 1989).

Self-mention was the focus of Harwood's (2005) study in which he examined the use of personal pronouns in academic RAs. This study is one of the most cited articles in the studies published addressing the use of first person pronouns and self-mention. Harwood (2005) attempted to prove in his corpus-based study that even the "author-evacuated" hard sciences use self-reference *I* and *we* (p.1207). In his analysis, he built on Hyland's categorization of *soft* and *hard* sciences, but he added to them a sub-categorization; soft/ hard-pure and soft/ hard-applied. The sample analyzed for this study

consisted of 40 RAs, 10 from each discipline, and the corpus ran on 325,000 words. The four disciplines were Business and Management, Computing Science, Economics, and Physics. Harwood's (2005) study is designed more on the qualitative end of the qualitative/quantitative continuum, examining the personal pronouns in context and identifying their usage. He concluded that researchers most frequently use personal pronouns to promote their own contributions in the research work, at the beginning of a research article, or to self-cite. They can also be used to report other claims (either accepting or refuting these claims). Researchers also use personal pronouns to show procedural innovations, methodological pitfalls or to self-cite. The list of functions of the use of personal pronouns that Harwood (2005) stated in his findings is important on two levels; first, other researchers may refer to it as a model in their research. Secondly, they may be provided as a guide for novice researchers in the process of familiarizing them with these techniques as a way of empowering them (Hyon, 1996). For the purpose of my research, the list of functions is compared with the findings of my sample. More functions may be added if found.

ii. Existential *"it"*

Another device for expressing stance is the existential *"it"*. This device was first referred to in the most cited research work done on stance markers by Biber and Finnegan (1989). In their study that was based on 500 spoken and written texts, one of the common structures among the stance styles they detected was the "*it…..that"* structure. It was employed in six of twelve different marking styles that were listed at the end of their study. According to Biber and Finnegan (1989), the use of this structure is usually associated with *doubt verbs, certainty verbs, certainty adjectives, doubt*

adjectives and *affect expressions.* The structure is sometimes used as an alternative to the direct reference to one's feelings or assessment for something as in the difference between *"It was frightening"* and "*I am frightened"*. The *"it…..that"* structure, according to Myers (1989), is the "coy" structure in which the stance is not expressed directly (p.22). The effect of using this structure in academic writing is that of "hedging" the statement. Myers (1989; 1991) uses the word "hedging" to describe the mitigating effect of using this structure and not in the sense of being a *hedge*, which is a syntactic category that is also associated with the expression of stance. The reference is similar to that of Lewin (2005) who generalizes the term "hedging" to all *downtoners*, or devices that help mitigate the effect of the statements whether they are genuine hedges or other grammatical structures having the same effect. It is also reported that the impersonal *"it….that"* is the vague counterpart of "we" or "people" that has a clear agent (Berman, 2004; Reillya, et. al, 2004). It is also compared to the French impersonal subject pronoun *on*, since it has no English counterpart (Reillya, et. al., 2004).

The existential impersonal "*it…that"* seems not to be in the focus of attention in the literature of stance markers in academic language. This is possibly because of its relative complexity as it is formed of more than one lexical item belonging to different syntactic categories. Looking at it from a purely structural point of view, the pronoun *'it"* takes the position of the subject and is followed by a verb. This structure is also known as the *anticipatory "it"* (Celce-Murcia & Larsen-Freeman, 1999, p. 446; Carter-Thomas & Rowley-Jolivet, 2001, p.12). However, what takes the verb slot might not be a

single verb, and after the verb, *"that"* might have alternatives such as *"to"* (Biber, 2006), or other as shown from Boyd, Gegg-Harrison and Byron (2005, p.44):

1-*It* + *v*erb+ adjective + *that*

2- *It*+ *verb*+ *that*

In the different structures of the *existential "it"* a variety of verbs may be used in the construction. However, the *be-verbs* are not as common as other verbs such as *seem* and *suggest* (Myers, 1989).

As I mentioned earlier, unlike other stance markers, there are very few studies that investigate the existential *"it"* construction as a stance marker. The pragmatic function of this construction is not examined thoroughly as was done with its form in grammar books, let alone examining its function in academic writing. This gap is, in fact, surprising as the *existential "it"* construction is the second top structure in the list of structures reported in the work that was conducted by Carter-Thomas and Rowley-Jolivet (2001) on stance markers in scientific texts. For this reason, I will discuss the section on *extraposition* and their conclusions about this structure in more detail.

In their study, Carter-Thomas and Rowley-Jolivet (2001) compared between the structures used in expressing the scientists' attitude in two different discursive contexts. The innovation of their study lies in that they compared the expression of the same content in different genres. They compared between orally delivered papers in an international physics conference and the corresponding proceedings on the same lectures to guarantee that the producers and the content of knowledge are the same in both situations. The corpus was complied of nine lectures delivered by native speakers of

English (NSs) and their nine corresponding proceedings. In the study, the structures that were most used in both sub-corpora were *passives, extrapositions, existential "there', inversion, "it' cleft* and *"wh" cleft.* In terms of frequency, *exitrapositioned "it"* was the second-top structure used in research articles after the *passive* construction but was ranked fifth in the oral presentations. This highlights the significance of this structure in the writing of RAs.

Carter-Thomas and Rowley-Jolivet (2001) reported that extraposition allows the writer to distribute the information load that might cause an imposition on the reader on different parts of the sentence. By doing so, the writer avoids putting complex structures (or meaning) in an initial position in the sentence. They also reported from their analysis that this construction was used as a hedging tactic to avoid expressing their opinions overtly. This was obvious to Carter-Thomas and Rowley-Jolivet (2001) after the comparison they made between the utterances and their equivalents in the proceeding. For example, *"it is shown that…."*was the written counterpart of *"we've demonstrated…."* And *"it seemed that…."* was the counterpart of *"we've developed an opinion"*. Considering each pair from a pragmatic point of view, it is clear that the spoken form might sound face-threatening if expressed in the same way in the written form. Carter-Thomas and Rowley-Jolivet (2001) concluded this section of the study that the *genre* of each corpus decided the variation in the form and frequencies of the stance markers that were analyzed. This conclusion echoes the notion that the difference in the medium of delivery may follow that the structures and the degree of explicitness and spontaneity vary as well (Paltridge, 2006).

iii. Stance and Disciplinary Differences

The degree of frequency, salience and decisiveness of the point of view of the writer varies depending on the discipline it is expressed in (Hyland, 2005; Williams, 2006). As Biber and Finnegan (1989) put it, "[s]pecialization in semantic field and grammatical devices thus seem to be highly correlated, presumably because of the different requirements of these styles" (p.117). This is akin to Connor's (2004) reference that both "big" and "small" culture, with "small" referring to the register, discipline and academic institution that the writer is affiliated to, and "big" referring to the nationality, cultural background or the ethnic group (p.292). But what is a discipline? According to the literature, the term "discipline" as related to academic writing was first introduced by Becher (1989). He first initiated the famous metaphor of disciplines as "academic tribes" or "tribes of academe". Each discipline is unified, or rather shaped according to a shared epistemological ground, common aims and similar concepts and methods. He gives an example with statistics as an epistemological field, and whether academic communities regard it as a separate discipline or not. He adds that the independence of statistics from mathematics, for example, depends on how different from mathematics its followed methods and conventions are. The Institutional independence and representation in independent academic journals also count in considering statistics as a stand-alone discipline. In other words, the definition of discipline does not strictly apply to different epistemological areas, as much as it is done with the agreement of academic community at large, i.e. it does not follow that every department is an independent discipline.

According to Becher (1989), being admitted to a certain academic discipline, follows that the new member adheres to its rules and follows its conventions that are highlighted and articulated through the use of the language. The linguistic and formal representation of content in different communities is one of the gateways of admitting new members of the society as well as expelling "illegal immigrants". This idea of acceptance in a certain community requires that the researchers who want to be introduced as members to the discipline show the gatekeepers of the discipline (reviewers in journals or supervisors for dissertations for example) that they are aware of the "[c]ommunicative conventions" of that new discipline (Hyland, 2002a, p. 389). Learning about a discipline requires the learning of not only the methodological and epistemological conventions, but also the linguistic strategies that are preferred in this discipline.

> Disciplines have different views of knowledge, Successful academic writing thus depends on the individual writer's projection of a shared professional context. That is, in pursuing their personal and disciplinary goals, writers seek to create a recognizable social world through claims for the significance, originality and plausibility of their work rhetorical choices which allow them to conduct interpersonal negotiations and balance against the convictions and expectations of their readers. (Hyland, 2005, p.178)

The fact that different disciplinary groups employ different strategies of evaluation to give different degrees of effect in an RA is well documented in the literature (Stotisbery, 2003). According to Hyland (2002a), "This directs us to the ways disciplinary texts vary, not only in their content but in different appeals to background knowledge, different means of establishing truth, and different ways of engaging with

readers"(p. 391). The following section shows how evaluation was realized in different disciplines in comparative studies between different disciplines.

Perhaps one of the broadest studies in terms of the disciplinary scope is Hyland (2005). Hyland analyses 240 research articles from eight different disciplines. He selected the top ten journals of each discipline to take three from each. The articles were compiled in an annotated corpus of 1.4 million words. The study instrument consisted of 320 search terms from previous literature. The study had two sub-categories to be examined; the first was stance, which is the focus of the present thesis, and the second was engagement. Stance, according to Hyland (2005), was an umbrella term for the categories *hedges, boosters, self-mention* and *attitude markers.* The purpose of his study was to identify the characteristics of stance and engagement markers in each discipline, and how frequently they are used. The results of his research show that an average of 200 evaluative words occured in every research article, which is a word every 28 words. These figures refute the claims that academic language is an objective language that is void of personal notes. The study also aimed at proving the interdisciplinary differences that is part of the focus of this paper. The results of the Hyland (2005) study showed that the soft sciences as humanities and applied linguistics employ evaluative language as much as three times as hard knowledge disciplines. The highest discipline in employing stance markers was philosophy and the lowest was mechanics. Hyland's study (2005) was mainly quantitative and general in terms of showing the directions of the high frequencies of evaluation markers across disciplines, narrower studies that pinpoint more specific items in context remain missing.

Auría (2008) examines the lexico-grammar of evaluative adjectives in two different disciplines by examining the introductions of 20 research articles taken from two journals; *Journal of the American Society of Information Science and Technology (JASIST)* and *The Journal of Applied Linguistics (AL).* A contextual analysis of stance adverbials was done manually and intuitively as the researcher reported that this category of stance markers is highly context-sensitive. The hypothesis that the researcher built his study on was that both disciplines would behave similarly in applying the use of adjectival stance markers. A total of 823 adverbials were collected from the corpus, and were checked for their category manually. The lexico-grammatical analysis revealed that the stance adverbials that occurred in the sample shared certain structures; this was a common aspect between both disciplines. However, there were some discipline-specific structures that appeared to be attributed to one discipline and not the other. For example, researchers in *AL* seemed to prefer adjectives that positively evaluate what comes later in the introduction section, while *JASIST* used adjectives to seek centrality of findings and negative assessment. Although my study does not examine stance adverbials as one of the markers of evaluation to be studied, the importance of this study is that it supports the hypothesis that any two independent disciplines might have a shared ground of persuasive devices, as well as peculiarities, and this has to be considered by EAP instructors in their curriculum design (Stotesbury, 2003). The differences that appeared in the previous study were expected as one belonged to the computational field while the other was from the social sciences. The following study sheds light on the fact that even the hybrid or mixed disciplines

(linguistics vs. computational linguistics) had different usage of stance markers from the disciplines they emerge from.

Degaetano and Teich (2011) in their study investigated the lexico-grammatical features of expressing stance in scientific research articles. Their study was based on a corpus analysis of published research articles in the fields of bioinformatics, computer science, biology and applied linguistics. Differences in the preferences of stance markers were detected across disciplines as well as comparing some disciplines with the ones that they emerged from. For example, mixed disciplines seemed to generally prefer *importance structures* (*it is important…..)* at the expense of the *complexity group* (*it is difficult, easy, hard….).* Comparing computational linguistics to the mother discipline -if one can call it so- which is linguistics, computational linguistics showed less use of the *possibility group.* On the other hand, microelectronics differs from electrical engineering in that it makes more use of the *importance group.* The results from this study highlight the danger of generalizing super-ordinate disciplines, writing conventions to their hybrid disciplines. Applying these findings to the purpose of this study, samples from pure medical papers (medicine, dentistry and physical therapy) were collected and examined and not samples from veterinary medicine or pharmacology for example.

iv. Stance across Cultures

With the introduction of *contrastive rhetoric* and *intercultural rhetoric*, attention started to be drawn towards the importance of comparing language production across cultures and the implication of this in understanding the differences between texts (Paltridge, 2006). Using these approaches, the notion that stance expression varies across

cultures was also proved (Precht, 2003).Thompson and Hunston (2006) reported in their book that *evaluation* in discourse is affected by the ideological backgrounds of different groups and subgroups. Therefore, if the argumentation and evaluative strategies are not established on a shared ground among different disciplinary and cultural groups, communication is very likely to fail. In his study that intended to elicit problematic issues that face NNS researchers when they submit their research papers to peer-reviewed journals, Flowerdew (2001) reported complaints about NNS researchers as being, in many instances, distant from their arguments and not showing the voice of the author. The responses were those of editors of peer-reviewed journals who noticed this problem based on their experience in reviewing and refereeing RAs. Orta (2010) also reported that the Spanish researchers she was sampling had difficulties in establishing a proper "tenor" when they wrote in English. (p.78). Also, Thompson and Ye (1991) said that from their experience, NNS students usually show difficultly in using reporting verbs effectively, and it was "common" that teachers find difficulty in identifying what views these students have towards their intended speaker (p.366). For the time being, we are using the term NNS following the tradition of studies that are to be mentioned later, although there is a more accurate reference term that could be used to describe the target sample for this study. This is going to be described and justified in detail in Chapter 3.

The next section is the most extensive section in this literature review. It is an attempt to pinpoint how intercultural differences affect the evaluative writing of NNS writers. The purpose of the following section is to provide examples of studies from different cultures and different geographical settings. Some of these studies are significant in terms of the results, others are important mainly for their methodology.

The first study in this section is reporting the Finnish setting (Maurannen, 1993) and then comes the study of Hyland (2002) about students from Hong Kong. Hendersson and Barr (2010) were reporting the French situation, Orta (2010) the Spanish, Molino (2010) on the Italian and Jafarpour and Taki (2012) on the Iranian. Thus, the attempt is to give evidence from different parts of the world, ranging from Scandinavian and European cultures, to Middle-Eastern and Asian, offering a different first language L1 background (Persian, Finnish, Spanish, French, Italian and Chinese).

Intercultural Differences

Perhaps one of the early studies on comparing academic written English across cultures that the researcher came across was the study of Maurannen (1993). In her study, Maurannen (1993) provides an early evidence for what Connor (2004) later on explained saying that writing is a cultural activity that researchers practice in order to communicate and circulate their knowledge. Therefore, linguists cannot escape analyzing the academic writing (represented in research articles) within a defined cultural context. In her comparative text-analysis study, Maurannen compared two excerpts from academic texts written by Finnish and Anglo-American researchers in terms of the use of metadiscourse. *Metadiscourse* is the term used by Maurannen to refer to the linguistic segments that writers use to organize and comment on the text. It is “meta-” in the sense that it is a text about the text. In this context, it is the broader term for evaluation markers that are employed in academic texts. Hyland (2004) also used the term *metadiscourse* in the same sense.

The strength of this study is in the choice of the sample. In her study, she drew her conclusions from a sample of documents that belonged to the same genre (academic RAs) and the same discipline (economics). For each pair, one article was written by a Finnish researcher and the other was written by an economist who is a native speaker of English (NS). Each pair also shared the same topic. Two articles were on *forest economics* models, and two on *taxation models*. Both the Finnish sample and its Anglo-American counterpart had some common features. This shared ground is referred to by Connor (2004) as *equivalence* or *tertium comparationis*, citing Chesterman (1998). Connor (2004) states that it is one of the criteria of the success and representativeness of contrastive rhetoric as it makes the differences in the language of the similar content and purpose more salient.

The analysis of the texts showed that the Finnish writers used significantly less metatext in their essays (22.6 % compared to the Anglo-American that had 54.2% of the whole text). Their arguments seemed to be more implicit and lacking elaboration. The presence of the writers was not felt in the texts. The discourse of the Finnish studies sounded more polite in the sense that there was less interaction between the reader and the writer. On the other hand, the Anglo-American texts were more elaborate, less obscure and a sense of the reader's presence in the mind of the writers was felt more. The style employed by the Anglo American researchers was more of the marketing-type rhetoric as opposed to the more poetic style employed by the Finnish researchers. This difference could be traced back, as Maurannen suggests, to the Finnish culture of respecting other readers by treating them as intelligent readers who do not need to be patronized. Such obscurity might be regarded by their NSs counterparts as arrogant or

impolite. It also finds its origins in the Finnish rhetoric in general; Finnish rhetoric can be described as implicit as well as poetic. This idea of the preference of discourse vagueness in L1 that affects the writing in L2 was also mentioned in Hinkel (2005).

In spite of the relatively small size of the sample, Maurannen's study gives rich qualitative information about the samples that she examined, which makes it insightful for similar studies. What is interesting about Maurannen's (1993) study is that she related the differences in practice to the geographical and anthropological context. The type of written discourse that appeared in the Finnish sample reflected the isolated and limited Finnish speaking community. The homogeneous nature of the society leaves little room for misunderstanding or obscurity. This is contrary to the English-speaking communities in which many cultures and people from different ethnic and linguistic backgrounds have to communicate. Hence, appears the need for elaboration and providing details and explanations on the discourse. However, these reasons sounded more as speculations on the part of Maurannen. She did not refer to any interviews or procedures by which she could get these inferences. In this case, Connor (2004) suggests that ethnographic procedures (interviews with the discourse producers for example) are to be held to probe for the real reasons for these differences.

This ethnographic dimension that Maurannen's study (1993) lacked was considered by Hyland (2002b) in his project in which he compared the use of personal references in the writing of students from Hong Kong with published RAs. He compiled a corpus of the 64 final project reports of undergraduate students. The reports were collected from biology (Bio), mechanical engineering (ME), information systems (IS), business studies (Bus), TESL, economics (Econ), public administration (PA), and

social sciences (SS). In that context, Hyland (2002b) highlighted that the projects written by the NNS students were to follow the typical conventions of a research article, and each paper ranged from 8000 to 13000 words. The corpus compiled contained 630,000 words and was searched for the personal pronouns *we, I, me, my, us* and *our*. Findings were compared to similar searches in his academic corpus that consisted of 240 research articles and 1.4 million words. Results of the study revealed that the personal referential occurred once every 1000 words, an average of 10 times per project report. Analyses for the frequencies lead the researchers to the conclusion that the writers of these reports prefer the anonymous passive form over self-reference. Recurrent as they might sound in this corpus, first person reference was represented in the Hong Kong students' projects as much as quarter as it was represented in the published articles. In the interviews Hyland held with some of the students as well as their supervisors, they reported that they regarded self-mention as individuality that is not preferred in academic writing. Therefore, they downplayed their role in their research projects and used fewer stance markers in their writing. He concluded his research project with a pedagogical note that teachers need to be aware of the conventions of self and stance expression to be able to teach them to novice writers of research. Like Maurannen (1993), he traced back this gap between the two corpora to the fact that the Asian culture does not prefer self-assertiveness. In making this claim, he echoes the complaint of editors about the inability of some NNSs of English (and especially Asian) researchers to show assertiveness that was documented in Flowerdew (2001).

Henderson and Barr (2010) also investigated this problem in their corpus-based research on undergraduate French students of psychology in their final year of University. The study was done by comparing three corpora; the first was a collection of papers written by the French students (NNS sample), Research introductions in the Directory of Open Access Journals (DOAJ) which again represents a NNS sample and the British Academic Written English corpus (BAWE). The size of the sample of this study was relatively small; 46,084 words for the NNS, 41,454 for the DOAJ and only 12,837 for the BAWE. This was justified by the authors who stated that the study was more of a pilot study and the target population is narrow and specified, hence, no need for a large corpus.

The results of the study supported the hypothesis that there are differences in the degree of using evaluative language among different groups of different cultural backgrounds. Although these students have received some training on the use of evaluative language before the samples were taken from them for the corpus compilation, there still appeared some differences between the two groups' practice. To begin with, the French students showed less use of grading verbs. As for the variety, the researchers reported that although there was a variety in the usage of evaluative items in the NNS corpus, a problem occurred in the accuracy of the meanings of those verbs. The defect with the NNS writing was that the evaluative markers were less accurate than the researcher had hoped for in preparing them with the teaching task that was given to them. Another difference was that NNS writers showed inability to evaluate works of other researchers. Moreover, their use of first person pronoun was limited to the function of guiding the reader to the different roles in the project they

were having in their papers, but they seemed to lack the confidence to use it as the voice of the expert analyzer of other reports. However, Henderson and Barr (2010) did not trace this problem to the possibility that they still do not have the confidence of experienced writers, as they are still university students who have not yet experienced writing research. Possibly, it is a skill that only experienced, and more confident writers learn. In this case, the expertise of the students would be acting as an extraneous variable in this study. In my study, the target sample is of RAs of researchers who are, in most of the cases, PhD and post-PhD researchers. The pieces examined are not representing "novice" writers, as much as they are genuinely reflecting researchers writing in English as their L2.

The methodological pitfall that the previous two studies have fallen into is the absence of "equivalence" in their sampling (Connor, 2004, p.298). The NNS and the published samples were not highly comparable in terms of the "small culture" (Connor, 2004, p. 292). The results of the two studies must be taken with reservation as the difference in practice might not find its origins in the culture or L1. Comparing undergraduate students who are new to the academic realm as well as new to the discipline they are writing in might cause them to be less confident in their self-presentation (Paxton, et al, 2008). The expertise and familiarity of the writers with the disciplinary writing conventions of RAs in Hyland (2002b) and PhD students in Henderson and Barr's (2010) , act as extraneous variables that may weaken the possibility of generalizing the findings of both studies. This can only be justified by the main objectives of the two studies; finding the model represented in authentic production that can be used as pedagogical material.

This methodological pitfall was avoided by Orta (2010). He used two target samples that were highly comparable. As he stated, he insured that the samples were all written by university-affiliated researchers in order to ensure that they were all familiar with academic writing conventions. They were also all management papers. The corpus was compiled out of 48 research articles of over 390,000 words. The number of articles was distributed evenly among the two sub-corpora; American and Spanish. Orta (2010) reported that the analysis of corpus showed that Spanish researchers tend to use modal verbs as stance markers differently from the way they are used in NS research articles. The Spanish writers used also hedges and boosters differently. The importance of this study is that it does not only trace differences in frequencies to compare them, but it also draws the attention of the reader to the differences in usage of the same lexical items. For example, the modal *can* was overused and mistaken for *may* and *might*. Orta said that in the Spanish language, the equivalent of *can* is inherently obscure and that the Spanish writers use *can* believing that they are giving the effect of the more mitigated *may* and *might.* Another finding was that Spanish writers generally lacked the sense of modalization (with the exception of *can).* Orta (2010) concluded from his findings that, in the case of Spanish speakers of English as an L2, in writing research "English is conditioned by the writing conventions of their national culture" (p.24).

Two more studies that underscore the note of Orta (2010) on the effect of L1 and how the original culture has its effect on the linguistic practices the RA writers use (Lightbown & Spada, 2006).The two studies compared samples of academic RAs from two different languages, one of them was English. They highlighted the differences in evaluative practices in each sub-corpus. Synthesizing Lightbown and Spada (2006),

Orta (2010) and the following two studies, the hypothesis is that L2 writers may tend to copy the writing styles and conventions of their L1 when writing in the target language.

The first one is the study of Jafarpour and Taki (2012). In their study, they applied a corpus analysis to compare stance markers in 120 scientific research articles as expressed in Persian RAs and other published research articles in peer-reviewed journals. Persian texts showed higher rate of using stance markers in general. However, *self-mention* was significantly underrepresented in what was regarded as a preference for highlighting the subject being discussed rather that the researcher himself. This act is achieved by avoiding using self-reference such as the first person pronoun. The other study is Molino (2010) in which the researcher compared English linguistics RAs with Italian linguistics RAs. The findings of the study indicated that the Italian academic language represented in the corpus employs personal references less frequently; not only the first person pronoun, but she was also looking for first person inflections that are remarkable to the Italian language. As Molino states it, the differences in representation of a certain linguistic aspect are evidence that researchers from different cultural backgrounds go for different choices in their interactional strategies.

v. The Problem of Writing in English as an L2

From the previous section, one can anticipate a difference in the language use in the sample that is examined in this study as compared to a more standardized model that is represented in the internationally published RAs. The problem does not stop only at the difference level, but it may also cause problems to the researchers who are after publishing in international medical journals. Cultural differences make the process of

writing in L2 more challenging (Hinkel, 1994). And L2 writers -and teachers as well- have long acknowledged that the process of writing in English is a difficult one (Hinkel, 1994; Huang, 2010). Research also demonstrated that the pragmatics of learners and native speakers (NSs) are different (Bardovi-Harlig & Dornyie, 1998). It is also expected that the NNS writers are going to go back to their writing paradigms while writing in English. This requires that they receive the amount of persistent and consistent training that will help them write in a manner that is within the acceptable norm (Hinkel, 1994; 2005).

Conclusion

Reviewing the literature, it becomes obvious that the cultural as well as the academic contexts have their stamp on the linguistic realization of the RAs written by researchers whose L1 is not English. Differences between published research and the RAs written by L2 scholars will guide instructors and course designers in the field of EAP to the areas that should be addressed. However, these differences are not identified to be copied completely, but researchers would rather use those conventions that are apparently representing the standard model. Researchers-to-be should learn to replace the writing strategies that they learn from their L1 with other more acceptable strategies that are generated from the model language (El Malik & Nesi, 2008; Sheldon, 2007). Comparative studies, as well as this study, are also useful in identifying the degree of "markedness" and salience of structures and rhetoric styles that are reported to make the process of their teaching/learning easier (Kelly & Bazerman, 2003). The problem with this comparative approach is that its results are not compared to responses of researchers

on what problems they actually face and what negative feedback they receive from members of editorial boards. If the comparison is done, more useful pedagogical implications could be derived from those analyses, instead of being merely descriptive.

Chapter Three

Methodology

I. Design

This study employs the methodology of corpus analysis. It compares two corpora: the medical sub-corpus of the Corpus of Contemporary American English (COCA) and a corpus of 47 Egyptian Medical Research Articles (EMRA), published in local journals. Being designed according to the mixed-exploratory method, this study was done in two stages. The first stage is the quantitative stage in which frequencies of each of the variables mentioned in Chapter 1 were compared. For example, the use of the personal pronoun *I* was traced in both corpora and was compared in terms of the number of occurrences in COCA and in EMRA. The number of occurrences for each search was converted into parts per million (ppm) to make it comparable to the frequencies in COCA. The second stage is the qualitative, linguistic analysis stage in which the samples of the dependent variables mentioned above were examined for meaning in context and compared in terms of what exact level of evaluation and stance expression they communicate.

II. The Corpora

A definition of what is meant by the word *corpus* is needed before explaining the data collection procedures. Traditionally, *corpus* refers to a body or a collection of naturally occurring language. This collection is used to observe certain patterns that may be related to certain meanings or specific to certain situations. A software program

referred to as “concordancer” is used for analyzing the corpus. Corpus analysis does not offer any new theories about the language; it only describes the language as it is authentically used. The importance of the *corpus* use in teaching lies in the fact that it offers evidence-based approach to the language that is to be taught (Hyland, 2006).

III. Corpus Compilation Procedures

The corpus from which the linguistic data for this study was collected has two main parts. The first is the previously compiled, ready-to-use medical section of the Corpus of Contemporary American English (COCA). It was complied of 6,700,848 words. Search results from this corpus were compared to another corpus of Egyptian Medical Research Articles (EMRA) that I have compiled. This corpus contained 99,467 words. For the purpose of compiling this corpus and before the compilation process, permission from the editorial board of each of the journals was obtained.

The Corpus Compiling Process

To make the corpus more similar to that of the medical section of COCA, I have chosen articles from the field of dental medicine and physical therapy as well as pure medical articles. The three Egyptian journals represented in this corpus are *the Journal of Advanced Research (JAR), the Journal of the Egyptian National Cancer Institute (ENCI),* and *the Journal of Egypt Public Health Association (EPHA).* The first two journals are Egyptian journals that are published in Cairo (more specifically, at Cairo University) and hosted by Elsevier; the third journal is published in Alexandria and hosted by Lippincott Williams & Wilkins. The articles used to compile the corpus for this study were published

between 2008 and 2013.*The Journal of Advanced Research* is represented by 18 articles, *the Egyptian National Cancer Institute* is represented by 20 articles whereas the *Journal of Egypt Public Health Association* is represented by 10 articles only. The reason that the number of articles from each journal is different is that I intended to choose articles whose authors are affiliated to national Universities and research institutes, for example, Cairo University, Suez Canal University, El-Fayoum University, Assiut University and the National Research Institute, to make sure the sample is representative of the writing of Egyptian researchers. Those institutions normally hire Egyptians only, which controls for the possibility that the researchers might not be Egyptian. Moreover, I intended to choose articles written by researchers who are affiliated to institutes that are directly related to medicine. I did not include articles written by researchers from nursing schools or from the High Institute of Public Health, for instance. To conclude, the two criteria that I mentioned previously limited my options and resulted in the discrepancy of representation between the journals.

In the process of compiling EMRA, the articles were transformed into text files so that they become readable by the concordancer. The corpus contained files for each journal with the articles numbered by the year of publication and its order in the issue of publication. For example, the first article I used from the 2013 issue was coded as 2013_1 and so on.

IV. Answering the Research Questions

In order to answer the research questions, a corpus-based analysis of 47 research papers that were actually published or in press from locally published

Egyptian RAs within the time span 2008-2012 was conducted. The articles for this research project were analyzed using a concordancing program (*MonoConc Pro 2.2*) that can be uploaded by electronic copies of the data. Both corpora are comparable, hence, observing any possible differences in the use of evaluative language and stance expression in both samples. For this project, it was possible for the researcher to use the corpora based on a training that was offered at the American University in Cairo, Spring 2012. The main instrument comprises of lists of two different linguistic structures expressing stance and evaluation.

The main reason for choosing published research articles in particular is that it is the actual medium of exchanging knowledge among members of the community (Peacock, 2002; Swales, 1990). The language that was analyzed is their authentic academic language produced by authentic participants and the language was examined in its authentic context. Hence, it is the researcher's premise that the data are representative samples of the language. Furthermore, corpus-based studies that appeared in the literature to analyze academic language regard published research articles as representation of the target population as those conducted by El Malik and Nesi, (2008), Huang, (2010), Hyland, (2005) and Bloch, (2010).

This data collection method has some pitfalls. First, it is very possible that the authors of some of the articles published in the Egyptian journals are not Egyptian, maybe other nationalities affiliated to Cairo University, although it is a very rare case (speaking from first-hand experience as a faculty member at Cairo University). It is also very possible that authors who are not from English-speaking backgrounds publish in international journals, which makes the "English" used not typically representing the

native use of the language. To avoid those pitfalls, the target sample is relatively large to make it easier to track the patterns that represent the mainstream as well as the outliers. Also, the comparison is done between high caliber peer-reviewed international journals to ensure, as much as possible, that the conventions followed are more of the "standard" and they represent the norms that are accepted internationally. Having pitfalls as it does, this method has been the only one employed in previous comparative studies that the researcher has come across (e.g., El Malik &Nesi, 2008; Harwood, 2005; Henderson & Barr, 2010; Hyland, 2002, 2005; Orta, 2010).

The two research questions in this study were answered according to the following procedures:

1- What are the techniques of expressing stance that are specific to medical research articles?

To answer the first research question, COCA was searched for the different aspects of stance expression mentioned above. International high-caliber journals are regarded as the model that represents the standard from which we make inferences about what L2 research writers should learn. This was done quantitatively by searching the frequencies of each search term as well as qualitatively by looking at the context of the most frequent and least frequent words and examining their meanings.

2- What are the differences in stance expression between research articles published in international peer-reviewed journals and locally published Egyptian research articles?

To answer this question, both corpora were searched for instances of each category and were compared in terms of frequency as well as use and meaning in context.

V. Procedures

The procedures for completing this study are mainly corpus-based. Each search term mentioned in the *instrumentation* section was searched for separately in both corpora (Hunston, 2002). Quantitative data is presented in tables and graphs containing the frequencies of each. Some lexical items were selected (according to the high/low frequency rate). In-context analysis was carried out with these terms in a search for any differences in meaning or use.

VI. Instrumentation

A list of potentially productive stance and evaluation markers was set based on the literature (Hyland, 2005; Harwood, 2005; Myers, 1989; Boyd, Gegg-Harrison & Byron, 2005). From the literature, I decided to search for the following linguistic structures as search terms for the corporal analysis.

1- Self-mention *(first person pronouns I, me, my, we, us* and *our).*

2- *"It…that"* structure.

3- Certainty and Doubt Adverbs.

4- Passive Voice.

VII. Methods of Reporting the Results

Results for the quantitative part are presented in tables showing frequencies, and numerically comparing samples from COCA with EMRA. As for the qualitative data, verbal description of the pragmatic meaning of each concordance line representing one

of the structures chosen by the researcher were analyzed contextually. Differences in meaning were reported verbally as well.

Chapter Four

Stance Markers in Medical RAs

This chapter examines stance expression using the different stance markers listed in Chapters 2 and 3. The frequencies, usage and collocates of the personal pronouns (singular and plural) and the impersonal "*It…..that"* structure that appeared in COCA were examined. The results stated in Chapter 4 directly answer the first research question "What are the techniques of expressing stance that are specific to medical RAs?"The results are also needed to set the framework for a comparison between the "standard" use represented in COCA and the use of researchers of the Egyptian medical RAs corpus EMRA. This is important for the process of diagnosing the main problematic areas that have to be tackled during the process of instruction.

This chapter will begin by examining the singular personal pronouns *I* and *me* and the possessive *my*. Then, their plural equivalents *we, us* and *our* will be examined. After reporting the findings of the personal expression of stance in the first section, the more impersonal form of presenting the writer's stance, namely existential *"it…that"* structure is reported. Whenever relevant, both structures were juxtaposed following the pattern Carter-Thomas and Rowley-Jolivet (2001) employed in their research as shown in Chapter 2.

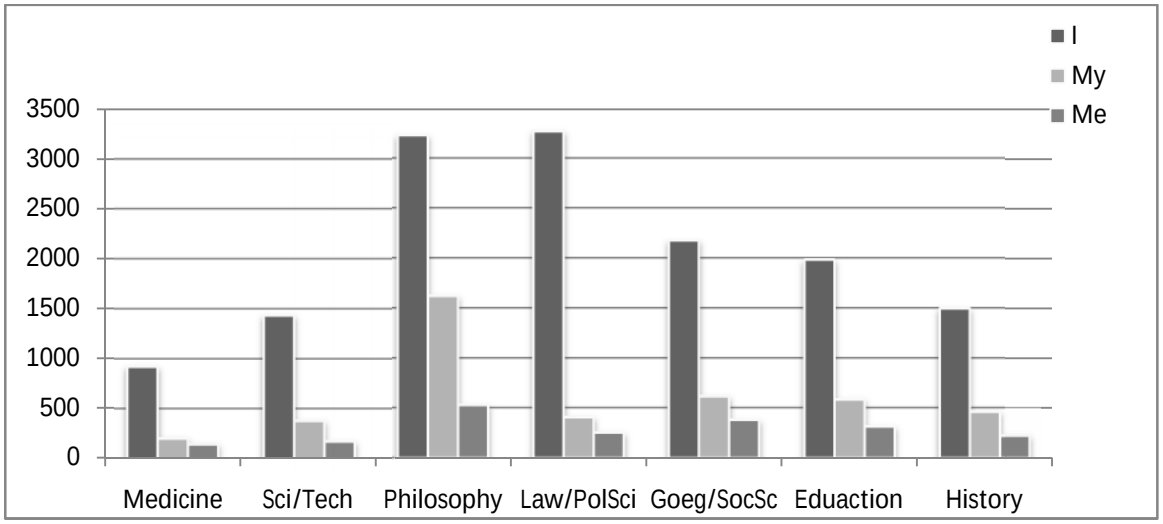
A. Use of First Person Pronouns in COCA

I. First Person Singular Pronouns *I, Me* and *My*

a. The First Person Pronoun for the Singular Subject *I*

Figure 4.1

The frequency of "I", "me" and "my" in the different academic disciplines in COCA



Before reporting the frequencies of each first person singular pronoun in isolation and what these frequencies imply, a global picture of the use of singular forms of the first person pronouns in the different disciplines is important. As shown in Figure 4.1, medicine is ranked least in terms of the frequency of the singular first person pronouns *I, me* and *my*. This finding implies that the use of singular forms of first person pronouns is marked in medical RAs compared to other academic disciplines. This implication is important in the course of instruction for two reasons. First, academic writing instructors should help the researchers-to-be use first person pronouns without either totally avoiding them or overusing them. That is, it is a form that is used in medical RAs, but is not very frequent and it has to be done with caution. Second, visualizing the frequencies of the

disciplines compared to medical language helps the instructors understand the source of the problem of Egyptian researchers in totally avoiding the use of these forms. It is its rarity in published medical RAs in general.

Looking at each pronoun in isolation, it became noticeable after conducting the first search attempts, that the personal pronoun *I* is not as recurrent in academic language compared to other types of linguistic discourse. With a 2,355.4 occurrences per million in the academic subsection of COCA, academic use of the pronoun was the least frequent, compared to spoken language which showed 18,117.8 occurances per million and magazines which contained 5,763.4 parts per-million. The frequency numbers, in fact, denote the markedness of this form in academic language. First, the medical sub-corpus of COCA was searched for the personal singular pronoun for the subject using the search key "*I*". The result was 6,177 occurrences. However, it was noticed that this search collected all instances of *I* including numerical references and *i.e.* instances. Therefore, other search keys were attempted to search for the instances where *I* is used as a first person reference. The first attempt was for the search *I + ve*rb and it collected 3648 hits. The *I + adverb* structure was also searched for which collected 272 instances. Narrowing down the search to examine the frequency of *I* in the medical corpus of COCA, again, medicine comes at the end of the list in terms of frequency compared to other disciplines that are included in the corpus with a frequency of 921.9 parts per-million. The highest frequency of *I* was in the humanities subsection with a frequency of 3,267.7 parts per-million, as well as the philosophy and religion subsection with 3,229.8 parts per million. Therefore, the first implication that we infer from these findings is that the use of the personal pronoun *I* is marked in medical RAs and is to be used with some caution when researchers come to write their research articles.

Figure 4.2

First person singular pronoun "I" in COCA: Comparing different disciplines.

Discipline	History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine
No of Occurrences in the Corpus	18316	18752	35255	18524	38972	21770	20208	6177
Parts per-million	**1,495.77**	**1,985.75**	**2,178.91**	**2,153.86**	**3,267.69**	**3,229.83**	**1,435.70**	**921.87**
Frequency								

The singular personal pronoun *I* occurred 6,177 times in the medical subsection in COCA, which is the lowest frequency among all the disciplines included in the subsection as shown in Figure 4.2. However, when I examined the concordancing lines of the search "*I*", I noticed that they were not all representative of the personal pronoun *I*. The search term "I" collected all the instances including numerical references and abbreviations. For example, the "I" in "stage I, II, II and IV….." had a numerical reference. *I* was part of an abbreviation in *"C.I for OR Upper-Lower Tumor…."* the abbreviation *"i.e."* was also collected and appeared in the results. Therefore, a search term that limited the results to those of the genuine personal pronoun references was needed. Looking at the concordance lines, it becomes noticeable that the personal reference *I* is usually followed by either a verb or an adverb. So the search *I + verb* as in the sentence "*Hereafter I will use the term' CFS/ME*" and "*One thing that I found very interesting is that none of these patients had ever had an audiogram."* and *I+ adverb* such as *"I also wanted the research to generate a substantive theory explaining how older people can feel"* were attempted as COCA is a tagged corpus. The results of this search was 3648 instances for the former and 272 for the latter, with a total of 3,920 constituting the instances of *I* as used as a personal pronoun and not as other indicators.

Both searches had a total of 405.92 parts per million across the medical subsection of COCA. This is particularly interesting knowing that the plural personal pronoun *we* had 2,222.67 parts per million.

The presence of the pronoun *I* in the medical corpus should not be overlooked, though. It does not follow that the low frequency of the pronoun as compared to other disciplines or compared to its plural form that the singular first person pronoun should not be given attention. One possible explanation for this rarity is that medical research writing tends to be done in groups. Individual researchers and authors in writing medical RAs is not the main trend. The frequencies of the use of the singular personal pronouns in COCA are significant when compared to EMRA. It can only be inferred that the personal singular pronoun for the subject should be used with caution in this genre and discipline.

b. Tenses Collocating with *I:*

Table 4.1
Verb tenses used with "I"

Tense	Number of Tokens	Part-per-million
Simple Past Tense	1366	203
Simple Present	932	139
Present Perfect	156	23.3
Present Continuous	30	4.5

The relation between the tense used with the degree of certainty, agency or responsibility for what is being reported is highlighted in this section. As shown in Table 4.1, the most frequent tense that was used with the single first person pronoun *I* in COCA

was the simple past tense with 203 parts per million . The second highest frequent tense was the simple present with 139 parts per million. In an attempt to understand the relation between the use of tense and the first person pronoun, the most frequently used verbs that occurred in the corpus with *I* were examined.

With the simple past tense, the top collocating verbs were *was, felt, found, knew* and *thought.* As can be inferred from the examples, the simple past tense used with *I* was mainly anecdotal (except for the first verb that is acknowledging someone and the third verb *found* that is procedural). It seems that the simple past tense with the singular pronoun is used to relate subjective information.

> -Acknowledgments: While working on this article, **I was** supported by a Harkness Fellowship from the Commonwealth Fund.
>
> " **I felt** confident because it was run by Crossroads, which meant that anyone on their books would be vetted to come into our home.
>
> **-I found** 2,256 articles that contained the keyword ototoxicity
>
> -I used to be very reliant on my line manager to lead the sessions and **I thought** it was just about going through my cases and taking direction.
>
> **-I knew** he was deteriorating but I didn't expect him to die too soon.

The simple present tense is less frequent compared to the simple past tense. The top verbs that collocated with *I* in the present were *think, feel, know, want* and *believe.* It can be easily noticed that three of the five verbs are verbs that denote uncertainty. As for the other two verbs *know* and *want*, the first one, *know* was mostly used in quotes such as *"I'm sorry, **I know** it tasted bad"* and the second one, *want* was used to state the

objectives of the researcher in doing something "*That is my objective in this column. **I want** to itemize some of the things we know about abuse"*.

c. The First Person Singular Possessive *My*

Looking for a more genuine personal pronoun to confirm the findings of the first person *I* and *me,* I conducted a search on the possessive pronoun *my*, following Harwood (2005), Hyland (2005) and Salager-Meyer (1999).

According to Celce-Murcia and Larsen-Freeman, (1999), the possessive pronoun (traditionally referred to as the genitive case) is a part of speech that denotes possession. They provide in their grammar book a list of functions for the use of the possessive pronouns of which I include *possession, agency,* and *representation* as they are related to the action of writing RAs. For the use that denotes *possession,* it may refer to the work that the researchers claim they own. For *agency*, it expresses a sense of responsibility for what is being reported as the agent or the doer is the researcher himself/herself. For *representation,* it functions as a device to present what the author/researcher has achieved, and what exactly is done by him/her and what is it that is done by other researchers s/he is reporting in the RA. According to Celce-Murcia and Larsen-Freeman (1999), the meaning and use of the possessive pronoun is highly genre and context dependent, which follows overlooking the other functions of the possessive as they are not applicable to the forms under investigation in this study (first person possessives) and they are not applicable to the genre under investigation (academic RAs) either.

Similar to the case with *I,* the possessive *my* was significantly less frequent in the medical sub-corpus compared to the other sub-corpora with only 1,367 instances (204.02

parts per-million). The use of possessive *my* in the medical corpus is almost used as one-fifth (1/5) of its use in humanities (1,006.58 parts per million) and less than education with less than 50% (616.68 parts per million). It may be assumed that the expression of ownership in medical RAs is not preferred.

I. First Person Plural Pronouns *We, Us* and *Our*

a. Plural Personal Pronouns for the Subject and Object *We* and *Us*

Figure 4.3

We in medical RAs as compared to other disciplines in COCA

Discipline	History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine
Number of Tokens in the Corpus	17691	18431	29710	19734	26428	25203	33159	15366
Parts per-Million	**1,444.73**	**1,951.76**	**1,836.21**	**2,294.55**	**2,215.91**	**3,739.16**	**2,355.83**	**2,293.27**
Frequency								

It seems that there is a paradigm shift in the use of the first person pronoun in the plural form. The frequency of the use of the pronoun *we* in the academic section of COCA is still the lowest among other discourse situations with 2,222.67 parts per-million compared to 9,615.29 in the spoken section and 3,228.88 in the newspapers section. However, comparing the different academic disciplines that are included in the different subsections in to the medical subsection, one can notice that the plural pronoun is not at the end of the list like the case with the singular pronoun *I*. As shown in Figure 4.3, researchers in medical RAs use *we* more frequently than those of the humanities, history and education, which was not the case with the personal pronoun *I*.

The relatively high frequency of the personal pronoun *we* in the medical sub-corpus may signify that the use of this pronoun is less marked than the singular pronoun and it is more accepted in RA writing , even more than other disciplines such as history, education and geology. The higher frequency of the plural pronoun compared to its singular equivalent suggests that in medical academic writing, it is more acceptable to make self-references when the authors are in a group. This hypothesis is confirmed by the relatively higher frequency of the first person object pronoun *us.* The pronoun *us* had 348.9 ppm in COCA, compared to 147.2 ppm for the singular pronoun for the object *me.* This could be traced back to two reasons; first, it could be due to the fact that multiple authors completing a single study is a trend in medical RAs. Second, the plural *we* down-tones the imposition of the reference to the self as referring to a plural self is more mitigated than the singular *I,* a hypothesis that is to be revisited by studying the choice of the Egyptian researchers in their work in the next chapter.

b. Tenses Collocating with *We*

It can be inferred from frequencies in Table 4.2 that the preferred simple tense to be used with the first person pronoun for the subject *we* is the simple past tense with 1007.2 parts per million in COCA. The present simple comes next with 526.38 parts per million. It is important to note that the simple past tense occurred as much as twice the instances of the simple present tense. This frequency sounds logical in light of the different functions of the personal pronoun *we* in the research article that is to be illustrated shortly.

Table 4.2

Tenses collocating with "we" in COCA:

Tense	Tokens	Parts Per-Million
Simple Present	3,527	526.38
Simple past	6,749	1,007.24
Present perfect	528	78.80
Present Continuous	257	38.36

Simple past tense was used with the personal pronoun *we* in medical research writing in describing the methods of the research, in stating the methodological innovations or pitfalls, reporting results and drawing conclusions and giving anecdotes.

1- The method **we used** builds on the strengths and seeks to avoid the pitfalls of previous environment and health plans from other countries**. We systematically reviewed** publicly available national environ mental health planning documentation from Europe, Australia, and China as summarized in Supplemental Material, Table 1 (http: **32;81772;TOOLONG).

2- Conclusions **We demonstrated** that partetravirus infection is widespread in monkeys, chimpanzees, and humans

3- **We established** six levels for the outcome variable: 0 = no dispositions, 1= one minor disposition, 2 = one serious disposition, 3 = two serious dispositions, 4 = three serious dispositions, and 5 = four or more serious dispositions

4- In the latter series, **we reported** a high number of complete responses restricted to tumours with mutated TP53

5- Discussion **We found** a high prevalence of diabetes in our study population.

The definition of one of the functions of the simple past tense that Celce-Murcia and Larsen- Freeman (1999) mentioned should be summoned here. They describe the simple past tense as "a definite single completed event/action in the past" (p. 114). The most frequent verbs that are used with *we* were as follows: *were, used, found, conducted* and *examine.* It is noticed that four of the five verbs are describing the accomplishments of the researchers that are completed and that they are claiming agency by using the pronoun *we*.

- **We conducted** a cross-sectional study in a population-based sample of 628 households in the UAE.
- **We examined** 2-way interactions among the following variables: self-rated health, body mass index.

The present perfect was used in the same contexts but with significantly less frequency in the corpus (78.80 parts per million). On the other hand, the simple present tense was used in acknowledgments, criticizing or commenting on other researchers' work and stating claims.

1- Moreover, the association between upward economic mobility over the life course and lower preterm birthrates does not extend to teenage African American mothers. **We suspect** that this reflects their short duration of residence in nonimpoverished neighborhoods and consequently limited exposure to protective factors over their life course. 7 Neighborhood affluence may take an extended period before it improves the reproductive health of formerly impoverished African American women.

2- Acknowledgments **We acknowledge** helpful suggestions on earlier drafts of the article by Laurence Grummer-Strawn, PhD.

3- **We assume** that the benefits of facial nerve surgery in patients with traumatic facial paralysis averages 1 or 2 House grades, (n2, n5) but this is debatable;

The present continuous was the least frequent tense collocating with the personal pronouns with a frequency rate of 38.36 parts per million.

1- So **we 're hoping** there will be little impact to the number and availability of short breaks.

2- **we are attempting** to determine whether the antagonist protects against ototoxicity

3- **We are investigating** whether the new mycoplasma has a role in human disease.

II. Functions of the First Person Pronouns in Medical RAs

In this section, the focus will be on the contexts in which the first person pronoun was used. In their studies, Harwood (2005) and Hyland and Tse (2002), generated two lists of functions of the use of the personal pronouns in RAs (*See Appendix A)*. Since none of them included medical RAs in their multi-disciplinary corpus, the attempt here is to apply the different functions on the medical sub-corpus of COCA using lines of the instances from COCA. The list of functions that will be generated is to be set as a model for a teaching material for novice medical researchers. It is also important as a basis for comparing the usage of *we* in the Egyptian corpus to the standard one.

i- Self Promotion at the Beginning/End of an RA

1 - Objectives# **We** investigated whether elevated risks of health disparities exist in Hispanic lesbians and bisexual women aged 18 years and older compared with non-Hispanic White lesbians and bisexual women and Hispanic heterosexual women. #

Methods. **We** analyzed population-based data from the Washington State Behavioral Risk Factor.

2- In this article, **I will** describe some of the pitfalls I (or other researchers) have encountered over the past 10 years of doing clinical research

3- Conclusions: We demonstrated that the agent of L-BSE can be transmitted by the oral route from cattle to mouse lemurs. As expected, orally inoculated animals survived longer than IC-inoculated animals**.**

4- In this review, **we** summarize documented cases of highland malaria that occurred in Ecuador during the early 20th.

As noticed from the examples, the first example seems to be the abstract of the article. The writers of this article are demarcating what they have actually done, drawing a line between what they did and what they are going to report as other researchers' work later in the article. In Example 3, the researchers are doing the same action at the end of the article, reminding the readers of the scope of what they did and the results they obtained. The final example (example 4) seems to be taken from the beginning of the literature review, also highlighting what the researchers intend to present.

ii. Self-citation

It is when a researcher mentions a previous article or research that s/he had conducted earlier, especially if it relates to the research being reported in the current research article, for example:

5- In the first article **I discussed** medical knowledge and decision-making, the sources that doctors currency use to find this knowledge, and some of the drawbacks

of these sources**. I also described** computer-based methods for searching bibliographic files. In this paper **I concentrate** on novel computer systems in which medical knowledge is not only stored and manipulated as text meaningful to human beings but is also encoded as symbols meaningful to computers.

As shown in the previous example, the writer of the research paper from which the extract was taken compares what s/he had done in the first article to what s/he did in the proposed one.

iii. Reporting on Other Researchers (Agreement/Disagreement)

When researchers are willing to agree or disagree with a procedure or a point of view, they can use the first person pronoun to show this objection. This is a strategy to show that this is an opinion that belongs to the researcher and not something that is agreed upon. In the following example, we notice that in the second sentence, the author seems to be starting a justification to his opinion. Giving the rationale behind his/her criticism mitigates the strength of the negative evaluation s/he states in the second sentence.

6- **I criticize** the concern over the politicization of public health policy as a justification for preferring a narrow to a broad model of public health. # My critique proceeds along 2 lines.

iv. Methodological Pitfalls

7- # Although **we** performed a thorough literature search via MEDLINE and PubMed, it is possible that **we** failed to document all conversions.

8- However, when we look more closely at this study, **we** see several serious flaws

Here, the researchers are admitting some pitfalls in their design or procedure and they are claiming responsibility for any mistakes that may threat the validity of their study.

v. Introducing Procedural Innovation

9- To facilitate the interpretation of results, **we** did not transform the dependent variable, but we conducted sensitivity analyses in which we restricted the distribution to values that were normally distributed

As example 9 shows, it could be inferred from the sentence that the course of action that is traditionally taken in this situation is supposed to be X. However, the researchers decide to take another route that will help them come up with the optimum statistical results and analysis. They state that they avoided transforming the dependent variable and that they found conducting a sensitivity analysis instead. They give the rationale for their choice together with using the first person pronoun to introduce the new method of calculation that they are proposing.

vi. Personalizing Claims

10- The implications, **I believe**, are about the future of environmental health.

11-**We** therefore presume that these were participants with high income.

In examples 10 and 11, the writers claim responsibility for what is deduced from the findings they are reporting. The phrase *"I believe…"* in example 10 implies that the researcher builds this claim upon his own belief, and not an implication to be taken for granted. In example 11, it seems that the researchers do not know for sure whether the participants that they are reporting were with high income or not. They are just building their assumption based on some given information. Using *we* as the agent of the claim,

decreases the imposition on the reader to accept this piece of information as a given. Personalizing the statements is an act of claiming the researchers' responsibility of what is being said.

vii- Including the Reader to Create Solidarity

> 12 - Do **we** recommend this type of arthroplasty to all of our patients? Suppose the patients **we** see in our hand clinics are young construction workers. How do **we** as surgeons move from specific evidence to clinical application?

This is a strategy of creating solidarity and a shared ground of agreement on a certain issue (in the case of example 12, it is a question and a hypothesis). In this example, the researcher claims solidarity with the readers of the RA by repeating the pronoun *we.* It is like a call for them to be included in the process of thinking with him. This action of engaging the reader is made clearer by defining the group he identifies himself and the readers among using "we surgeons".

viii- Stating the Purpose of the Research

It is acceptable that the researcher uses the first person pronouns in stating what the research paper aims at. It specifies the work that the researchers are going to do in the study drawing the line between what they do and what is reported from the accumulative knowledge that is available in the literature. Again, it is an act of claiming responsibility and an act of presentation of one's own accomplishment as Celce-Murcia and Larsen-Freeman (1999).

13 -**I explore** how human understandings of influenza have altered over the past 500 years and how public policy responses have shifted accordingly. **I trace** the progress in human understanding of causation from meteorological conditions to the microscopic, and how this has prompted changes in public policy to mitigate the disease's impact**. I also examine** the latest trend of viewing pandemic influenza as a security threat and how this has changed contemporary governance structures and power dynamics**.**

ix – Stating the Research Procedures

The use of the personal pronoun *I* in the process of describing the procedures that the researchers follow in their research was quite recurrent across the results lines of my search. This can occur in the methodology section and the abstract; it cannot be confirmed where these instances were located (abstract, methodology….). It is an observation that I deduced after looking at random samples of the concordance lines.

14 - I used individual-level data on adults from the National Longitudinal Survey of Youth 1979.

15- I included an observation in the sample if there was information on the respondent's current weight, zip code area food environment, and long-term family income. **I arranged** the data as a panel so that there were multiple observations per respondent and the unit of analysis was a person-year.

x- Acknowledgments

When researchers acknowledge those who helped them accomplish the study, it is common to use the first person pronoun.

16 - I am particularly grateful to the reviewers

17 - While working on this article, **I was** supported by a Harkness Fellowship from the Commonwealth Fund

18- *I thank* Prof Keith Peters who stimulated **me** to present a lecture to the Association of Physicians in Cambridge in 1989.

xi- Drawing Conclusions

In the process of reporting the findings that the researchers draw from their study, it is also useful that they claim responsibility and ownership of what they report by using the first person pronoun in this context. Using this strategy helps mitigate the imposition that other researchers might feel on accepting these results. Secondly, and more importantly, it gives the researchers credit for adding these findings to the body of research in their fields. It reinforces the senses of ownership that the researchers are claiming.

19 - On the basis of these findings, **we** estimated that if prion infectivity is present at all in sCJDMM1 urine, it is at most 0.38 IU/mL if the 20-fold infectivity loss is factored in.

The list of functions illustrated previously is the list of functions that I compiled from Harwood (2005) and Hyland and Tse (2002). There are two more functions that I drew from the corpus that I was working on. Those functions were not mentioned by Harwood (2005) and Hyland and Tse (2002). It is not confirmed if these functions are limited to medical language of RAs or not, but since they appeared, they are reported below:

xii- Stating a Hypothesis Based on Evidence

In some cases, the researchers are simply giving predictions or hypotheses that are not yet proved by means of research. These statements remain as predictions until they are proved or refuted. They are sometimes taken as ideas for further research. In this case, statements have to be expressed in a cautious attitude that makes the researcher bear responsibility of what s/he is reporting or else, its source might be confusing to the reader.

20- On the basis of these findings, **we** estimated that if prion infectivity is present at all in sCJDMM1 urine, it is at most 0.38 IU/mL if the 20-fold infectivity loss is factored in.

21- **We** therefore presume that these were participants with high income.

xiii -Anecdotal

This is when the researcher is speaking about some experience that s/he went through. It was rather recurrent in the medical subsection of COCA, especially in giving some anecdotes about events that took place during the process of data collection.

22 -On December 5 of 1992, **I visited** the Chernobyl Nuclear Power Station. The most exciting part of **my** five-day trip…..

23 -I do not speak Russian or Ukrainian, and **I did** not know anyone from the former Soviet Union before **my** trip.

IV- Rate of Repetition

It is also interesting to report in the findings that repeating the personal pronouns is acceptable whenever necessary. There does not seem to be constraints on the recurrence of the same pronoun in the same section or paragraph as long as it is justified. As shown in the examples below, in the first example, the researchers repeated the pronoun *we* nine times in 161 words, with an average of once every 17.9 words. In the second example, the pronoun *I* occurred five times in 90 words with an average of once every 18 words.

1- **We** obtained the ratios of the geometric means (with 95% CIs) by exponentiating the coefficients and standard errors from the linear regression models on log-transformed blood lead levels. # **We** performed the models with progressive levels of adjustment for relevant determinants of blood lead. First, **we** adjusted for gender, age, race/ethnicity, country of birth, BMI percentile, and survey year. Second, **we** further adjusted for household education and PIR. Third, **we** further adjusted for housing age (according to the year the family house was built). **We** conducted plots of model residuals to confirm the adequacy of the models. **We** conducted a test fo r trend in the association between increasing cotinine and blood lead concentrations by including cotinine medians corresponding to each quartile as continuous variables. **We** conducted a test for trend in the association between increasing number of smokers at home by including 0, 1, and 2 as continuous variables in the regression model. **We** used stratified analysis by participant characteristics.

2- **I included** an observation in the sample if there was information on the respondent's current weight, zip code area food environment, and long-term family income. **I arranged** the data as a panel so that there were multiple observations per respondent and the unit of analysis was a person-year. The analysis sample contained 14191 observations on 4202 women and 13634 observations on 4085 men. **I refer** to this sample as the main sample. **I created** a " change " sample from the main sample that included variables representing an individual.

I. *"It…that"* Structures

Moving to the more impersonal strategy of expressing stance, I will investigate the structures and use of the *"it…that"* structure in COCA in the coming section. As mentioned in Chapter 3, the evaluative value of this construction lies in the fact that it is more impersonal, so it gives the researcher the chance to withdraw and express stance more subtly. The second reason for being evaluative is the evaluative markers that are used inside the clause. In most of the cases that were encountered in examining the lines in COCA, the *"it…is"* clauses contained other stance markers that either heightened or minimized the stance expression of the statement as in the following examples:

1- Clinically, **it *suggests* that** hyperplasia of the remaining lymphoid tissue might be a concern later on,

2- According to the available animal data, **it *also appears* that** steroids cause little or no ototoxicity.

3- This finding has also been recognized by others. (n15-n17) In fact, **it** was ***reported*** **that** mandibular setback for the correction of mandibular prognathism has produced OSAS.

4- **It *also seems evident* that** psychiatric practitioners need to use their clinical experiences with stigmatization.

In the first example, the researcher seems to be commenting on some results, the use of the verb *suggests* implies that the researcher does not want to show absolute confidence in his statement. He rather precedes it with the word *clinically* so that he gives evidence to the truth of the statement by stating that it was examined clinically thus appearing as more modest.

In example 3, the statement shows a different tenor than that expressed in examples 1 and 2. The researcher uses a different reporting verb that implies more confidence and less modest attitude. The researcher directly states that this result is "reported" and that s/he does not claim responsibility for this reported claim. The use of the passive voice in this context may seemingly contradict the confidence of the author. However, the use of the passive might be justified if the researcher is going to refer directly to the researchers s/he is reporting and will cite their names. The sense of certainty is more accepted by the reader given the context of the clause. In the preceding sentence ***"This finding has also been recognized by others"*** the researcher states that the finding that he refers to had been confirmed by previous research as well. So, the confident tenor is reinforced by mentioning that the findings come in line with other findings.

In example 2, the lack of certainty expressed by the verb *appears* is explained in light of the previous statement *"according to the available animal data"*. The researcher shows a modest stance of uncertainty that arises from the sample he seems to be building on his conclusion.

In the fourth example, a rather paradoxical term is used in the "*it…that"* clause to create a balance in the stance that is being expressed. The researcher uses *seems* which reflects uncertainty to modify the adjective *evident*. The word *seems,* is a hedge that down-tones the imposition of the word *evident* leaving space for other researchers who would want to differ with his findings. It is as if the researcher is saying: "it is evident to me from what I see in the results, but if you do not find it convincing, you have the choice of not accepting it".

Table 4.3

Different structures of "it….that" matrix as occurred in COCA:

No. of Slots	**Structure**	**Example**
1	It + verb+ that	It **seems** that….
2	It + is/was+ verb+ that	It **was found** that…. It **is hypothesized** that….
	It + is+ adjective+ that	It **is desirable** that…. It **is clear** that
	It +adverbial+ verb+ that	It **also states** that… It **is likely** that…
	It + modal+ verb+ that	It **may be** that….
	It + became+ adjective + that	It **became apparent** that…
3	It + passive present perfect (3 slots)+ that	It **has been shown** that… It **has been estimated** that…..
	It+ is/was + adverbial + adjective+ that	It **is very important** that….. It **is theoretically possible** that…..
	It+ modal + passive (2 parts)+ that	It **should be noted y9**that….. It **can be concluded** that….
	It+ adverb+ verb+ adjective+ that	It **also seems peculiar** that….
	It +is/was + adverb+ verb+ that (passive)	It **is generally recognized** that….
	It+ is/was+ a + noun+ that	It **is a given** that….
	It +is/was + not+ adjective+ that	It **is not surprising** that….
4	It + is + adjective + to + verb+ that	It **is important to note** that….

	It + present perfect (2 parts) + pronoun+ experience + that	It **has been our experience** that…..
	It+ has/have + adverb + been + verb (passive) + that	It **has also been shown** that…..
	It + is/was + adverb + not + adjective+ that	It **is perhaps not surprising** that….
	It + modal + passive (2 parts)+ preposition + that	It **should be pointed out** that…..
	It + modal + adverb + be + verb (passive) + that	It **can also be concluded** that
	It+ is + preposition+ adjective + noun+ that	It **is of paramount importance** that……

As noticed from Table 4.3, the longer the construction of the *"it…is"* clause is, the more syntactically complex it gets, and the more variety of structures it allows. For the first structure, it only allows for a verb between the two ends of the clause. The second category, the search in the medical sub-corpus of COCA showed five different syntactic structures, utilizing verbs, auxiliaries, adverbs and adjectives. For the third and fourth categories, we could have as much as seven different structures using words from a broader spectrum of word classes such as prepositions or including both adjectives and adverbs in one clause.

*I. It * that*

It appears from the frequencies comparing the different disciplines that were included in the COCA academic corpus that the "*it *is"* structure is least used in

medicine, with 376 tokens out of the 6,700,484 words constituting the medical sub-corpus in COCA (56.1 parts per million). The asterisk located between "the" and "is" is the wildcard used in COCA to make the concordancer search for one slot between the two words that are searched for. For example, searching for "it * that" will only collect the occurrences that have only one word between "it" and "is", such as "it *appears* that" and "it *is* that". This relatively short structure, again, does not seem to be very frequent in medical articles.

Figure 4.4

*The frequency of the one-slot "I * that" clause compared in the different sub-sections of COCA*

	History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine
No. of Tokens	1266	878	1683	1181	1189	1048	936	376
Ppm	**103.39**	**92.98**	**104.02**	**137.32**	**99.69**	**155.48**	**66.50**	**56.12**
								

For this construction, the part of speech that is mainly used between *it* and *that* is a verb, with an exception of the adjective *clear* in the sentence *"make it clear that…."* that appeared 24 times in the corpus. The most frequent verbs used with this structure are *appears* (100 tokens), *seems* (38 tokens), *appeared* (15), *follows* (21), *suggests* (11) and *means* (11). This draws the attention to the significance of the use of the verb *appear* as it was used in the corpus in both the present and the past tenses for 115 times (30.6% of the *"it…is"* occurrences in the corpus). Together with *seem*, *appear* is a verb that reflects uncertainty, an attitude in writing that does not particularly denote that the researcher is

unsure about what s/he is saying but rather expressing that the extraposed clause is something that is open to other interpretations. The percentage of the use of both *seems* and *appear* in this construction was 42%, as they represented 148 tokens out of 376 tokens. This finding might suggest that medical RA writing tends to use less sure expressions and mitigated styles in writing.

> -From the boaters' comments, however, **it** *seems* **that** many of them do not believe that boaters are major contributors.
>
> **-it** *appears* **that** the elderly with functional impairment place more value on their stability.

As can be inferred from the previous examples, both instances are commenting on interviews or some sort of collected data from participants in the study. The researchers seem to be deducing information about the participants' motives to their answers/ behavior. This is different from the certainty that is expressed in the verb *mean* that is used in the following example:

> - There wasn't *any* difference between them HIV-positive and HIV-negative individuals, so **it** *means* **that** what a normal fellow can do, a person with AIDS can equally.

The esearcher in this statement shows a kind of certainty about what is being deduced from the results of the research tool. The source for this certainty may be the word *any* that is used in the first clause that negates any source for doubt that both groups of patients are not equal.

From the previous examples and the frequencies stated above, it can be noticed that part of the evaluative quality of the *"it ….is"* construction lies in what collocates with it. It is not only that this structure allows the evaluated object to be located towards the end of the sentence and not in the subject position; it is also the effect of the evaluative lexis that is inserted in the structure.

The choice of the tense of the verb used with the "*it….that" structure* with this structure also demarcates the stance of the writer. The predominant tense used with this construction is the present simple tense (243 times, 64.6%) then followed by the simple present tense (57 times, 15%), suggesting the preference for the incomplete verbs over the decisiveness that the simple past tense bears. The issue of the tense choice will be tackled again when COCA is compared to EMRA to investigate the preferences of the researchers.

II. "It * * that"

The difference between this structure and the previous one is the distance that separates between its two ends (*it* and *is)*. By distance here I mean the number of slots between the two ends. In this section, the search key used to collect instances for this structure was *"it * * that"* which collected the instances in which there are two slots between *it* and *that.*

It is important here to refer to the statistics of this structure compared to its shorter counterpart. The "*it * that"* structure that has only one slot in the academic section was ranked in the third position in terms of the frequency of the structure with 101.15 parts per-million. The spoken section came in the first position with 111.79 parts per million

followed by the fiction section with 107.55 parts per million. For the second structure " *it * * that*", the frequency raised from the third place among discourses to the second place, suggesting that the more distant the two poles are, the more formal and academic the language might sound. For this structure, the frequency was 290.05 in the medical section. The increase in the number of parts per million by almost three times also suggests that the language of the medical RA prefers the more distant one. This preference may be because the more distant the structure is, the more it allows for more tense markers, function words, modifiers, stance markers and mitigations/boosters to be inserted before the evaluated entity (Hyland & Tse, 2005). It is noteworthy that the spoken section contained 298.31 parts per million of the search term *"it * * that"*, whereas it only had 111.79 parts per million for one slot search term *"it *that.* This suggests that the longer structure is preferable in English in general.

Figure 4.5

*Frequencies of the "it * * that" construction in the different COCA sub-sections*

	History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine
No of tokens	3191	3489	5376	2464	3101	2556	3035	1958
Ppm	**260.59**	**369.47**	**332.26**	**286.50**	**260.01**	**379.21**	**215.63**	**292.22**
								

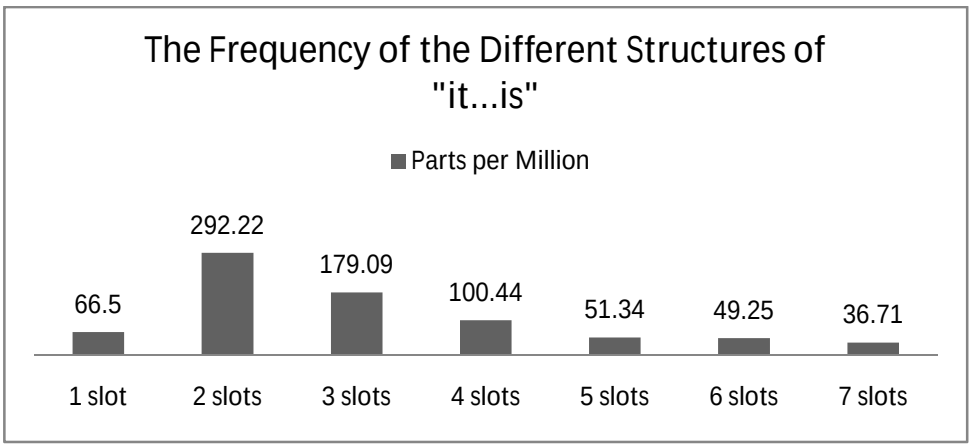
III. "*It …..that*" Constructions with more than Two Slots

As mentioned above, the existential *it* structure is least used in its one-slot form. This is confirmed by the fact that the structures with three slots as well as the structure with four slots still keep academic discourse in the second top rank among all six

discourse contexts that are included in the corpus. As mentioned above, the single-slot had only 101.15 parts per-million tokens in the academic language. Increasing the search terms to insert three and four asterisks in order to collect instances of the structure that have three (as in the sentence ***"It might be argued that the intervention did not succeed because it was not successfully applied***" and four words between *it* and *that* (as in the sentence ***"it is perhaps not surprising that we found no benefit of a community-based strategy for preschool")*** , the number of tokens per-million for each structure remained significantly higher than the short structure. For the three slots, there were 226.21 parts per million and for the four slots 162.37 parts per million.

Figure 4.6

The frequencies of the different constructions of the "it …that" matrix in COCA



As for the frequency of both structures, in the medical sub-section, the three slots showed a frequency of 179 parts per million and the four slots showed a frequency of 100.44 parts per million. However, as shown in Figure 4.6, there is a decline in the frequency of the parts per million starting from the structures with five slots upwards. The reason behind this decline might be the complexity of the longer constructions. Longer structures (five slots and above) are not to be discussed in detail in this study. I

will only focus on the first four categories as they show higher frequency in both COCA and EMRA.

Chapter Five

Egyptian RAs and Internationally Published

Articles Compared

Introduction

In an attempt to identify the differences between the Egyptian usage and the more "standard" one as Salager-Meyer (2001) describes it, the comparison between the search results from EMRA and those that appeared in COCA is established in this chapter (p.67). The first stance marker to be examined in this study is the first person pronoun and the first person possessives in both corpora.

A. Comparing the Frequency and Use of Personal Pronouns in EMRA and COCA

II. First Person Singular *I, Me* and *My*

One interesting finding was that of the use of the first person singular pronoun *I.* Searching for *I* in EMRA rendered 100 matches. However, by looking at the concordance lines carefully, it could be easily noticed that none of them indicated reference to the first person. *I* was used as a numeric character reference or as part of the abbreviation "*i.e.*". This finding is interesting given that 10 of the 47 articles that constituted EMRA (21.2%) are single-authored. Yet, none of those 10 different authors referred to their work using one of the three first person singular pronoun markers that are examined in this study (*I, me* or *my).* From this finding, one can infer that Egyptian researchers do not opt for overtly paying credit to themselves when they

are writing as single authors. They prefer to play it safe and not to give a direct reference to themselves.

II. Plural Personal Pronoun *We, Us and Our*

The use of the first person singular pronoun in both corpora becomes particularly significant when it is compared to the use of the first person plural pronoun *we*. For the singular personal pronoun (subject, object and possessive), there were no instances for the search in EMRA. However, the plural personal pronoun showed that Egyptian researchers are more comfortable referring to the self within a group, either a group of researchers or the researcher using the interactional *we* with the readers. It may be that they feel safer when each single author's self is being disguised in the whole group. There were a total of 118 instances of *we* out of the 99,647 words corpus (2.5 times per article). Nonetheless, the frequency of the plural personal pronoun in EMRA remained significantly lower than its frequency in COCA, with the number of tokens per-million in COCA signaling almost twice the number of tokens in EMRA. This suggests that, despite their acceptance to use the first person in RAs, Egyptian medical researchers do not prefer direct reference to themselves; hence they do it less frequently than the researchers of internationally published articles. This finding is to be confirmed when I report the findings for the search of the impersonal structure of the existential "*it….that"* clause.

The plural personal pronoun for the object *us* occurred less frequently in EMRA compared to its frequency in COCA. In EMRA, *us* appeared only twice, making 20.07 parts per-million. The same pronoun in COCA registered a frequency of 348.93 parts per-million. To be able to identify the difference between both corpora, I will calculate the ratio of *us* to *we* in each corpus. The ratio of the number of tokens of *us* to the number of tokens of *we* in COCA is

1: 6.6. This means that the pronoun *we* is used in medical RAs as much as 6.6 times more than the pronoun *us.* Using the same comparison between the same two pronouns in EMRA, we find that the ratio of the number of tokens of the pronoun *us* to the number of tokens *we* is 1:59, indicating that the pronoun *we* is used in the corpus almost 60 times more than *us.* The wide gap between the ratios of both corpora draws our attention to the tendency of the Egyptian researchers to avoid to overtly refere to themselves in the object position, and their preference for the self-reference in the subject position.

Table 5.1

Comparing the use of first person pronouns in COCA and EMRA

Pronoun	**COCA ppm**	**EMRA ppm**
I	405.92	0
Me	147.15	0
My	204.02	0
We	2293.27	1184.16
Us	348.93	20.07
Our	1157.38	2257.92

As shown in Table 5.1, the only instance of a personal pronoun where the number of the tokens in EMRA surpassed the number of tokens in COCA was the possessive *our.* The first person plural possessive pronoun *our* had 225 tokens in EMRA, with a frequency of 2257.92 parts per million compared to 1157.38 parts per-million in COCA. A possible explanation of this discrepancy is that possession in Arabic is expressed using and inflection "*na"* that is added at the end of the possessed object. In Arabic, an

inflection that bears less imposition than a complete word does. The phrase "our results" will simply be expressed as "nata'eg***na"*** as one word**.** Pragmatically, this word in Arabic bears a less sense of possession that the word *"our"* does.

A deeper pragmatic analysis of the pronoun *our* in context might provide better explanation for the reasons why this pronoun is used more frequently. The most frequent collocate with *our* was the noun *study.* The pronoun *our* collocated with *study* 82 times in EMRA. Examples of some of the concordance lines are:

1- No Kidney affection was recorded in ***our study.***

2- However, the prevalence in ***our study*** is higher than that recorded in Thiede et al.

3- ***Our study*** revealed a higher percentage of illiterate females.

In the first example, the strength of stating the findings using *our* is mitigated by the negative marker *no* at the beginning of the statement. It is as if the authors wanted to stress that it is a finding that is limited to their study. This is a strategy that makes the researcher show more caution about generalizing their findings .This strategy makes drawing conclusions sound more modest. For the second example, the use of the pronoun *our* comes also to down-tone the imposition of contradicting previous research findings. In this statement, the authors are reporting different results compared to one of the studies they seem to have mentioned in the literature. The use of *our* leaves the impression that the researchers are saying "we are not claiming that their findings are wrong, but we have a different figure for the same experiment in a different context". This makes it safer for their face to report their own results. The same applies to the third example as it limits the findings to the study being reported.

III. Collocates of Personal Pronouns in COCA and EMRA

a. Nouns Collocating with *Our*

This section is meant to examine the top collocates of the pronoun *our* in both corpora and set comparisons between them in terms of frequency, preference and use. This comparison is meant to test whether Egyptian researchers are aware of the correct use of the pronoun in the intended context or not. If not, what are the problems that the analysis of the data at hand will diagnose so that it becomes helpful for the process of instruction.

As mentioned in the previous section, the first top collocate to the first person possessive pronoun *our* in EMRA was the noun *study.* It occurred in the corpus 82 times, 822.9 parts per million. The same word was also a top collocate in COCA but with a number of tokens of 1053 (157.2 parts per million).The collocate of the second highest frequency collocate was the noun *results,* and here lies another difference between COCA and EMRA. Native speaking researchers seem to have a preference for the noun *findings* over the noun *results.* On the other hand, Egyptian researchers seem to have preference for the noun *results* (notice the gap between the frequencies of the two nouns in EMRA in Table 5.2). One possible explanation for this preference is that the noun *result* is the one that is usually used in Egyptian mathematics and science school books. The word *results* is also the literal translation to the Arabic word “nata’eg”. This hypothesis is reinforced by the fact that usually Egyptian students in governmental universities refer to any scores as the *results,* as the word “nateega” is used to refer to scores.

Table 5.2

Comparing the top noun collocates with the pronoun "our" in COCA and EMRA:

Collocate	**Parts per million in COCA**	**Parts per million in EMRA**
Study	157.2	822.9
Results	43.7	260.9
Knowledge	29.8	100.4
Patient	50.4	80.3
Findings	58.4	40.1
Case	20.1	40.1
Data	23.7	30.1

Looking at the top collocates of the pronoun *our* in both corpora, it is interesting to note that among the top collocates for the pronoun *our*, the first seven top collocates in COCA were found among the top 10 collocates in EMRA, suggesting the Egyptian researchers' awareness of the collocation of the pronoun, the only difference was with the preference of *results* over *findings* that I have explained earlier. Another difference that can be noticed from the frequencies shown in Table 5.2 is the difference between the numbers of the top collocates in both corpora. For example the top collocate in both corpora is *study*, but the frequency of the word in EMRA is 822.9 parts per million whereas its frequency in COCA is 157.2 parts per million. The second top collocate in EMRA is *results* with 260.9 parts per million whereas the second top collocate in COCA is *findings* with only 58.4 parts per million. This brought to the spotlight the hypothesis

that non-native speakers of English might use a less variety of nouns for a certain purpose. The word was used in EMRA almost five times more than COCA. This hypothesis was more strengthened by the fact that EMRA had only 14 noun collocates after *our*, whereas COCA showed a list of 100 nouns. This claim, however, should be investigated further in a larger corpus as EMRA is a relatively small corpus compared to the size of COCA.

b. Collocates with *We*

, The noun collocates with the personal pronoun *our* were examined in the previous section. In this section, I will examine the collocations for the plural personal pronoun *we.* Searching for the main collocates after *we,* I noticed that the plural personal pronoun *we* is usually followed by either a verb or an adverb *(also, then, now, retrospectively…).* In this section, the frequencies of the collocates will be examined in both corpora and conclusions will be drawn from frequencies compared.

To begin with, when EMRA was examined for the top collocates for the pronoun *we,* the top 13 collocates were all verbs, except for the adverb *also* that was ranked fifth in the list of collocates and it had 5 tokens. Appearing in the list of collocates as the first (and only) adverb in the list triggers the question of whether the use of this adverb is close to the standard that is represented in COCA or not. Searching for *we + adverb* in COCA, the top adverb collocate was *also,* suggesting that Egyptian researchers are aware of the most frequent adverb collocate to this word. The following adverb collocates in COCA were *then, now, therefore and retrospectively.* Each one of them was searched for in EMRA but there were no results. It appears that the only adverb that Egyptian

researchers used following *we* is *also.* This observation, again, brings the question of the lack of variety in the style of Egyptian writers to the limelight.

The rest of the 12 top collocates that appeared in EMRA were verbs. Two aspects are to be discussed here; the top verb collocates in both corpora compared and the tenses preferences in both corpora. In this comparison, there appeared some dissimilarity between the two corpora. For this search, I picked the content verbs only. That is, any auxiliary verbs were ignored, and only the content verbs were counted. This is particularly important because the verb that was on top of the list of COCA was *have.* Checking the concordance lines, I found that it was mainly used as an auxiliary for the verbs in the past perfect tense. As shown in Table5.3 that shows he top ten verb-collocates in both corpora, the two lists agree in only 30% of the verbs (the shaded items in the table). That is, only the verbs *used, found* and *report/ed* were present in both lists. For the rest of the 70%, it seems that each population of research writers had their own preferences. In COCA, there was a preference for the verbs *did, believe, describe, conducted, thank, had* and *examined.* As for the researchers of the Egyptian RAs, they seem to prefer the verbs *were, employed, followed, aimed, compared, studied and recommended.*

Table 5.3

Top verb-collocates with "we" in COCA and EMRA compared:

	COCA	**EMRA**
1-	Used	Found
2-	Found	Reported
3-	Did	Used
4-	Believe	Were
5-	Report	Employed
6-	Describe	Followed
7-	Conducted	Aimed
8-	Thank	Compared
9-	Had	Studied
10-	Examined	Recommend

It is important at this point to discuss the tense of the verbs used with the pronoun *we* as it adds to the stance value of the statement. For the purpose of counting the instances of each tense, I used the tag options available in COCA. Since EMRA is not a tagged corpus, searching for the occurrences had to be done manually. I examined the 118 lines and counted the instances of each tense and then converted the number of instances under each category into parts per million to make it comparable to COCA.

As shown in Table 5.4, researchers of EMRA seem to have the same top preference for the simple past that researchers of COCA have. The simple past tense collocated with *we* 1007.24 parts per million in COCA whereas it collocated with *we* in

EMRA 742.62 times per million. The tokens in EMRA were only 30% less than they were in COCA. The percentage of the difference between the two figures in both corpora is significant for the comparison between the simple past tense and the present tense. The simple present tense was the second preferred tense in both corpora. It occurred in COCA 526.38 times per million and in EMRA, it occurred 190.67 times per million. The number of occurrences in EMRA was 64% less than COCA. This percentage, when compared to that of the simple past tense, denotes that EMRA researchers tend to be more comfortable with using the simple past tense.

Table 5.4

Frequencies of the tenses of the verbs collocating with "we":

Tense	**Parts per million in COCA**	**Part per million in EMRA**
Simple Present	526.38	190.67
Simple past	1007.24	742.62
Present perfect	78.80	20.07
Past perfect	8.06	0
Present continuous	38.36	10.04

The perfect tense in EMRA was almost not present, with only two occurrences for the present perfect and zero occurrences for the past perfect. This is an important remark about the influence of L1on the choices that NNSs of English make. It is important to note that Arabic has no present perfect tense. Therefore, Egyptian learners have problems

using perfect tenses, and it is understandable why they opt for simple present and past tenses; the ones they have equivalents for in their L1.

IV. Functions of the First Person Pronoun in EMRA

In this section, the function of the personal pronoun *we* in EMRA is examined, comparing my findings to the functions of *we* in COCA to see to what extent the writing of medical researchers adheres to the conventions of internationally published RAs. For the purpose of doing this search, a manual search was conducted, reading the 118 instances of sentences using *we* in EMRA in search for instances of the different categories mentioned in the previous chapter. The different categories are:

1- **Self promotion at the Beginning/End of the Article**

-In this research, **we** aimed to evaluate clinical and immunological features with patience with SLE.

2- **Reporting previous Research (Agreeing with/Refuting Previous Findings)**

-We confirmed the known literature data indicating the higher efficacy of marrow histological examination.

In this statement, the researchers agree with what they seemingly have included in the literature of the study. They use the first person pronoun to express this agreement.

3- **Personalizing Claims**

-We believe that preoperative percutaneous open or laparoscopic biopsy is discouraged.

In this statement, the negative evaluation for a procedure is “discouraged”. This negative statement is mitigating using a personal pronoun *we* in a manner of claiming

responsibility of the critique. It could be a face saving strategy that the researchers used to tone down their negative comment.

4- **Introducing Procedural Innovation**

-We suggest the following step-wise strategy that can be applied at least in developing countries….

Having done their research in a developing country themselves, the researchers seem to have come up with a new strategy of reaching data that they are sharing with other researchers. They use the first person pronoun *we* to claim ownership of this newly introduced strategy and to market for it.

5- **Including the Reader**

-From tables 2,and 3, we can note that the agreement between…..

Here, the researchers are engaging their readers by including them to the group using the pronoun *we* to invite them to look at the table in which the researchers are presenting the figures they are referring to. Harwood (2005) describes *we* in this context as acting as a discourse guide.

6- **Stating the Purpose of the Research**

-We aimed to compare the efficacy of alcohol-based hand-rubs versus traditional hand-rubs.

In this statement, the researchers are stating clearly what the purposes of the research they are conducting are. They assume agency of the whole process of finding the gap as well as conducting the research by using *we.*

7- **Describing Procedures**

-We used one-way AVOVA test to detect differences in clinic pathological variables.

The researchers are describing the statistical method that they use for this study. This is one of the procedures that the researchers went through and are claiming responsibility for using the personal pronoun *we.*

8- Reporting Conclusions

- In this study, we demonstrated that AM may be potentially useful peptide to counteract the deleterious effects of a diabetic state.

Here, the researchers are adding their contribution to this field of knowledge by reporting what they have concluded from their study. The use of *we* gives them credit for adding this new finding to the literature. However, the researchers seem to show some modesty by mitigating their statement using the modal verb *may* and the uncertainty adverb *possibly* to mitigate the confident tenor that the use of *we* adds.

9- Acknowledgements

- We would like to thank Dr. Doaa Tawfik….for her permission to conduct the research.

Following the conventional style of paying credit to others who contributed or helped in accomplishing the work, the researchers acknowledge the role of those who helped them do this research using the personal pronoun *we*.

10- Self-citing

- We have previously reported a high methylation frequency of APC….. Therefore, we sought to confirm this data in a larger cohortof HVC genotype 4 infected patients.

This statement seems to refer to a previous research that was accomplished previously by the researchers, and needed to be revisited to fill a certain gap. The

researchers are specifying what they had done before and what they are doing now, using the pronoun *we* in both statements to claim ownership of both studies.

11- Stating Methodological Pitfalls

- Thus **we** could not assess the fraction that opted to permanently discontinue their medications as those would not be accessible.

In this example, the researchers are describing the pitfall in their data collection that might give way to extraneous variables. They are stating that they were unable to access a certain population due to some circumstances that they mention.

In light of the above lines, we can conclude that Egyptian medical researchers agreed with the researchers represented in COCA in 11 of the 13 functions that are stated in the previous chapter. There is 84.6% agreement between the two corpora in terms of the different functions for which the first person plural pronoun is employed. The two functions that were absent from EMRA were using *we* in giving an anecdote and using *we* in hypothesizing based on evidence in the research.

B. Existential "*It ……That*" Structure

The samples under investigation in this study show some similarities and some interesting differences in terms of the form and use of the existential *it* followed by *that* structure. To begin with, starting the search in EMRA following the same search pattern that was followed in COCA (of which the findings were reported in Chapter 4), the hypothesis was that Egyptian researchers who are writing in English as a foreign language will not use this structure as frequently as it is used in internationally published

RAs. This hypothesis is based on fact that this structure is one of the most complex English structures, as illustrated in Chapter 2. My hypothesis was also that Egyptian researchers will have preference for shorter structures of *"it …is"* given the syntactic complexity of the longer ones.

Table 5.5:

Comparing the frequency of each structure of "it …that" in COCA and EMRA

Structure **No of Slots**	**COCA** **ppm**	**EMRA** **ppm**
1 slot	66.50	50.18
2 slots	292.22	401.41
3 slots	179.09	491.74
4 slots	100.44	60.21
5 slots	51.34	50.18
6 slots	49.25	20.7
7 slots	36.71	80.25

The first finding that contradicted my hypothesis was that, whereas researchers of international RAs showed preference for the second type (containing two slots) over the rest of the of the structures with 292.22 instances per million, Egyptian researchers showed a significantly higher rate of utilizing the three-slots structure (see Figure 5.1). It is important to note that the Egyptian corpus showed a frequency of the three-slots construction of 491.74 parts per million whereas COCA showed a frequency of 179.05 parts per million, meaning that this construction is used as much as 64% more frequently

than EMRA compared to COCA. This high frequency makes us consider whether this relatively high frequency compensates the low frequency of using personal pronouns in the writing of Egyptian researchers or not. Confirming this assumption will follow that we conclude that Egyptian medical research writers prefer the more faceless way of showing their presence in their writing.

In addition to the previous remark, the Egyptian corpus showed a higher frequency of the seven-slot construction, with 80.25 parts per million, whereas COCA showed a frequency of 36.71 parts per million. This finding suggests that Egyptian researchers show competence in using complex structures. This could be traced back to the nature of the Arabic structure that tends to be, to an extent, elaborate, containing embedded clauses and modifiers. At this point, it might be useful to take a thorough look at the syntax and the most preferred structures of *"it….is"* in the Egyptian RAs. In the process of the linguistic analysis, the shorter constructions are examined first while constantly referring to the "standard" formats that appeared in the concordance lines in COCA.

I. Syntactic Structures of *"it…is"* in EMRA

The "*it * is"* structure (one slot) appeared in the corpus in five instances only, three of which utilized the verb *appear*. The preference of the use of the verb *appears* echoes the high frequency of the verb in COCA as reported in Chapter 4. One sentence used the verb *seem,* and another used *reveal.* Although results from a sample of five sentences cannot be generalized to a conclusion that Egyptian researchers do not show competence of using a variety of verbs in this structure, it remains significant when it is echoed by similar

remarks in other contexts. Further research using a larger sample would confirm or refute this hypothesis.

Investigating the two-slot construction, again, brings the limited variety of choices to the limelight. The search key *it * * that* collected 39 instances. The 39 lines were distributed between two main structures; "it *is* …that" which occurred 13 times in the corpus and "it *was* …that" which was utilized 26 times. To compare these results with those of the same search term in COCA, it is noticed that there are some other structures that were used other than *is/was + adjective/passive.* Although the first 25 structures of the collocation list (that contained 100 results) used *was* or *is* as the verb, the list also contained other verbs such as *it could be that, it would seem that, it became clear/apparent that,* and *it seems likely that.* This observation, again, raises the question of the lack of variety that the Egyptian corpus exhibits.

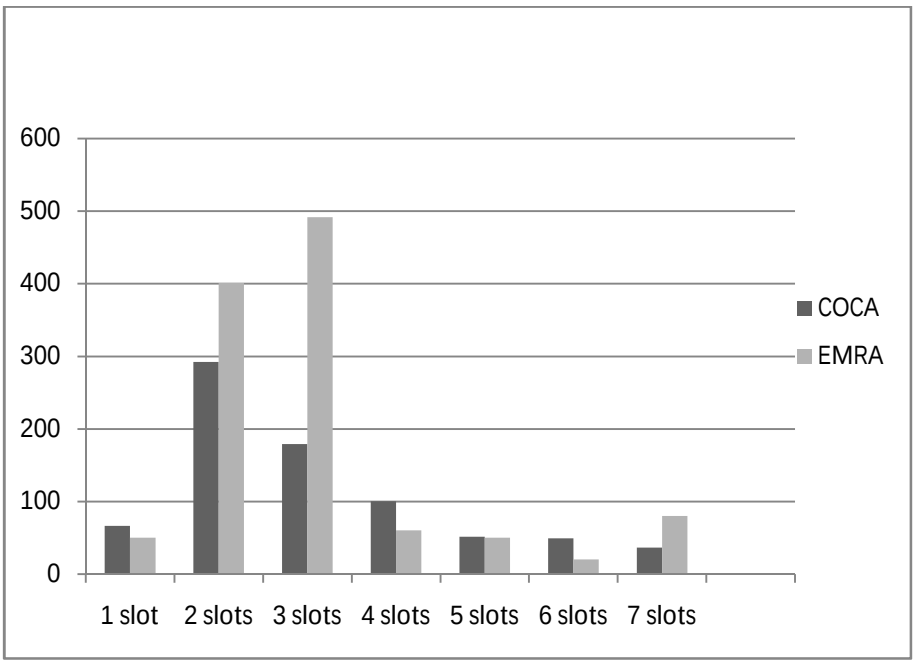
The longer construction, that contains three slots, was the one of highest frequency in EMRA. The search collected 48 lines. This construction was a little more varied. Beside the passive structure with an adverbial modifier *"is/was +adverb+ verb*", there also appeared the present perfect passive "*has +been + verb"* and the modal phrase *"can/could +be + verb"*. The most frequent of the three structures was the passive in the simple past and present tenses (with 19 occurrences) then came the present perfect passive with 15 tokens and finally were the construction with the modal verb passive with 7 tokens. It is noteworthy that it is in under this category that adverbs appeared in the corpus in 16 of the instances. The adverbs that were used were *well, also, generally, critically* and *now.* The construction of adjectives modified by an adverb was also used in 7 instances such as

"it is now clear that …." 3 tokens, *"it is well known that…."* 3 tokens *and "it is well established that…."*1 token.

It seems now more evident that Egyptian researchers choose from a limited variety of linguistic formats, either in structure or in the choice of words. This limitedness in constructions that Egyptians use in organizing the *it…is* clause, when contrasted to the variety of structures that is illustrated in Table 4.3 in the previous chapter pinpoints a possible defect in the style of the Egyptian researchers.

Figure 5.1:

*Frequencies of the different structures of "**it …is"** in COCA and EMRA*



C. Other Stance Markers Employed in EMRA

The findings of comparing the writing of Egyptian researchers to instances from COCA revealed differences in distancing and positing techniques between the two samples. The use of direct self-reference using personal pronouns is more frequent in COCA, while

researchers of EMRA seemed to prefer the more impersonal *"it…that"* structure. In an attempt to confirm the findings, few more stance markers that were tackled in previous studies were attempted. The selection of these markers was mainly based on their relatively high frequency in EMRA as well as the practicality of the search terms that will collect the target structures.

I. Certainty and Doubt Adverbs

Certainty and doubt adverbs are categories that Biber and Finnegan (1989) identified and generated a list of. They are adverbs that implicitly bear certainty or uncertainty about the statements they modify. They subtly reflect the attitude of the writers towards what they are writing. I use the research related adverbs that occurred in this list to compare them in both corpora (Table 5.6 & 5.7).

A. Certainty Adverbs

Table 5.6

Adverbs expressing certainty of claims in COCA and EMRA

Certainty adverbs	Frequency in COCA ppm	Frequency in EMRA ppm
Actually	82.08	20.07
Certainly	31.49	10.3
Evidently	1.64	0
Indeed	59.1	0
In fact	71.04	10.3
Obviously	21.34	10.3
Of course	31.94	0

B. Doubt Adverbs

Table 5.7
Adverbs expressing doubt in COCA and EMRA

Doubt Adverb	Frequency in COCA	Frequency in EMRA
Apparently	24.99	70.24
Perhaps	85.22	10.35
Possibly	53.28	70.24
Presumably	16.2	20.7
Reportedly	5.97	10.35
Seemingly	6.87	0
Likely	460.71	170.6
Technically	9.85	30.1
Theoretically	11.34	30.1

As noticed from tables 5.6 and 5.7, EMRA researchers showed a significantly lower frequency of the seven different certainty adverbs that were examined, indicating a tendency to avoid sounding certain about their claims. On the other hand, doubt adverbs showed higher frequency in six of the nine adverbs that were examined in EMRA (67%) suggesting the tendency to show a more modest stance. These findings are in line with the previous conclusions that Egyptian doctors prefer to sound more modest and prefer to use more subtle and faceless stance expression strategies.

II. Passive Voice

Carter-Thomas and Rowley-Jolivet (2001) reported in their study that compared between oral research presentations and their written proceedings in physics, that the structure that was most used in writing RAs was the passive voice. Biber and Finnegan (2001) also reported that the passive voice is one of the structures that are preferred in medical language. Therefore in an attempt to confirm the findings that suggest Egyptian researchers tend to be more distant in their writing, I also investigated the frequencies of the passive structures in both corpora.

As shown n Table 5.8, the frequencies of using the passive voice from EMRA exceeded those from COCA in the three tenses and in both singular and plural. In the plural simple present for example, EMRA is only 31% higher than COCA, with 1846.5 instance for the former and 1472.9 for the latter. In other cases, the difference between the frequencies in COCA and EMRA is vast. For example, the simple present singular in EMRA was 67% more frequent than in COCA. The number of tokens per million in EMRA for this case was 10306.3, whereas COCA was 3411.99.

Table 5.8

Frequencies of the passive structure in different tenses in both COCA and EMRA

Passive tense/number	COCA	EMRA
Singular/Simple Present	1631.7	3181
Plural /Simple Present	1472.9	1846.5
Singular/Simple past	3411.99	10306.3
Plural/ Simple past	3126.3	8510.04
Present Perfect/ Singular	471.16	732
Present Perfect/ plural	472.8	602.1

For the purpose of this study, the passive structures and meaning are not to be examined any closer. The purpose of this elementary search was to find any further evidence before concluding that Egyptian medical researchers use a more detached writing style in their RAs than the internationally published researchers. The use of the passive voice so frequently, which disambiguates the role of the agent of the action by not mentioning him, is another evidence for the difference in demarcation attitudes of the two populations that were examined.

Chapter 6

Discussion

This study aimed at examining the usage of the first person pronoun and the *"it…that"* structures as stance markers in medical RAs. It was noticed from the literature that using the first person pronoun was not tackled from a medical RA point of view in any of the most cited research dealing with first person pronouns, namely; Harwood (2005) and Hyland and Tse (2002). Also, Salager-Meyer (1999), when she was dealing with the topic of referential behavior in general medical RAs, she only traced the change of using first person references over 185 years (1810-1995). Her analysis was not done from a pedagogical point of view. It was also noticed that the structure of the impersonal *it* plus the evaluative *that* was not addressed in the literature in spite of its frequent use in medical RAs especially in EMRA. So, I tried to address those gaps as well as examining the language of Egyptian researchers in the field of medicine as this topic was not tackled before, especially from a contrastive point of view. In this chapter, the main findings of this study are discussed, followed by discussing the implications for teachings, limitations of the study and suggestions for further research.

Patterns of Stance Marking in Standard Medical RAs COCA

As seen from the results reported in Chapters 4 and 5, the use of first person in medical RAs is an acknowledged strategy that researchers use to position themselves, give themselves credit as well as engage the readers in the arguments.

There were distinct contexts in which first person pronouns are usually located, and these contexts were first set by Harwood (2005) and Hyland and Tse (2002). Applying these situations to medical RAs, I found that both lists apply to medical RAs and two more functions were added based on my search results. From the search results, first person pronouns are used at the beginning or the end of an RA to promote for the researcher/s, they are also used to self-cite, cite the research of others, report procedural innovation, methodological pitfalls, claim responsibility for the statements and to engage the readers and to include them in the discussion. Personal pronouns are also used to state the purpose of the research, describing methods and research design and reporting results. Personal pronouns do not function in the articles in isolation. As shown in the qualitative analysis of the concordance lines, the evaluative quality and the stance of the writers are crystallized by means of the context of the personal pronoun itself. For example, research words like “reported” and “concluded” as well as mental verbs like “believe” and “presume”. These lexical items, determine (either heighten or soften) the degree of certainty or the strength of the claim that the researcher is reporting. These strategies of expressing stance are described by Biber and Finnegan (1989) as “faceless stance” (p.103).

The analysis of the concordance lines of COCA has shown that there are certain patterns that can be generated from the production of the internationally published research articles. There seem to be certain syntactic patterns that govern the formation of the existential *it… that* clause. These patterns were classified in this study by the number of slots between the two poles in the clause. These patterns can be used in teaching materials that can be included in academic writing courses.

Differences between Egyptian Medical Research Articles EMRA and the Standard Production in COCA

Perhaps the most remarkable conclusion that the findings of this study highlight is the avoidance of singular personal pronouns *I, me* and *my* by Egyptian research writers, although the corpus compiled for this study contained 10 single-authored articles out of a total number of 47 articles. This finding becomes particularly significant when it is compared to the use of the plural pronoun instances *we, us* and *our*. Although relatively less frequent than its frequency in COCA, the plural personal pronouns were used with an average of three times per-article. The conclusion is that Egyptian researchers, perhaps due to some cultural influences, or due to lack of linguistic awareness, are more comfortable feeling disguised within a larger group of contributors, hence, feeling less imposition in the direct self-reference. The tendency to use subtle strategies of stance expression is a conclusion that is strengthened by the results that showed that Egyptian researchers showed a higher frequency of using the impersonal *it* structure.

There is evidence, however, for some agreement between both corpora. For the preferences of the verb tense choice, Egyptian researchers showed awareness of the preference for simple past tense in reporting statements using *we*. They also showed an awareness of the second top preference, simple past tense. The only difference between the two corpora in terms of the use of the verb collocates with *we* was the rarity of the perfect tenses (only two instance in the Egyptian RAs corpus). This could be traced back to the Arabic grammatical system that does not have an equivalent to the perfect tense.

Another feature that also proves that Egyptian researchers are aware of complex structures is the use of the longer construction of the existential *it* clause. It is not only that they are acquainted with the structures, but they can also use them syntactically correct. This is evident from the finding that showed that Egyptian researchers used the *"it ….that"* matrix more frequently than the internationally published RAs.

One of the major observations about the writing of the Egyptian researchers is the lack of variety either in the choice of lexical items or syntactic structures. For example, Egyptian researchers also used only two constructions of the two-slots *existential it*; or rather, two different tenses of the same linguistic structure. Comparing this finding to the different constructions that were used in COCA, five different constructions appeared in COCA, utilizing different verbs like *became* instead of *is* or *was* and different parts of speech such as adverbs.

The effect of L1 in EMRA is not to be overlooked in this study. The use of the word *results* instead of *findings* more frequently in the corpus suggests that researchers might prefer the words that are closer to the literal translation of the word in their L1 as illustrated in Chapter 4. This is also reinforced by the fact that the present and past perfect tenses have only two instances in the whole corpus.

Looking at the findings of the study at large, it seems that we have enough evidence that Egyptian researchers prefer the more subtle positioning strategies, they would rather sound uncertain than think they are imposing their findings or abruptly evaluating statements. This conclusion about the stance of Egyptian researchers should be

taken a step further and be included in some pedagogical implications, helping researchers-to-be write better.

Implications for Teaching

This study was initially aimed at exploring problematic areas of the writing of the Egyptian researchers and to compare it to standard internationally accepted styles to generate meaningful teaching materials. It is part of a broader research project that aims at creating data-driven material that specifically fit Egyptian researchers in each discipline. The purpose of these materials is to help researchers who are writing in English as an L2 publish internationally using the accepted norms of RA writing, and hence, become more persuasive to the gatekeepers of the discipline.

a. Breaking the Barrier of Using the First Person Singular Pronoun

The first issue to be addressed is the use of the first person singular pronouns *I, me* and *my*. Researchers and researchers-to-be should be encouraged to use those pronouns in case they write as individual authors. It is useful to provide them with instances from the standard corpus to give them concrete examples of researchers who have actually published in prestigious journals (such as *Lancet* or *The American Journal of Public Health)* who are using the first person in certain contexts to serve a certain pragmatic function. Teaching the first person pronoun from a pragmatics point of view will be helpful in this respect, given that they are adult learners of language who can grasp the concept of pragmatics. The researcher-students could be given tasks in which they are asked to convert passive sentences into active statements using the first person

pronouns. These types of activities also address the problem of overusing the passive voice in Egyptian RAs.

b. Addressing the Problem of Lack of Variety

Overcoming the problem of limitedness of repertoire could be well dealt with by providing the researcher-students with the lists of the different compiled structures from COCA. For instance, lists of research process verbs that collocate with the first person pronouns that were generated from the top collocations of the search term in this study. Providing the researchers with a wider variety of syntactic structures of the *"it…is"* construction based on the taxonomy that is provided in Chapter 4 and applying it to some research papers of their specializations.

The teaching process could be content based. That is, the instructor could use actual research articles that are published internationally to take as a launching point for teaching. Teaching issues related to stance expression to novice researchers is recommended to be a part of an extensive academic writing for doctors course.

Limitations of the Study

As Connor (2004) states, attempting a high resemblance of corpora should be what researchers in contrastive rhetoric target. It is what she calls *tertia comparationes, or "equivalence"* (p.299).This resemblance makes the two samples comparable and controls any extraneous variables. As mentioned in Chapter 3, I used for this study two different corpora; one was a ready-for-use corpus that is available online (COCA) in

order to search for the standard production of language, the other one (EMRA) was compiled for the purpose of this study. The compiled corpus contained 47 articles with a total of 99,647 words whereas COCA contained 6,700484 words. For the purpose of this study, it would have been better to compare between two compiled corpora of the same size and choose articles from the same number of journals.

Suggestions for further Research

As mentioned earlier, including the points that have been discussed in this study in an extensive writing course follows that other problems in the researchers writing be addressed. Further research conducting moves analysis to address structural inconsistencies in Egyptian RAs. It is also useful to examine the use of tenses in the medical RAs in general.

References

Auría, L. (2008). Stance and academic promotionalism: a Cross-disciplinary comparison in the soft sciences. *Atlantis,* 30(1), 129-145.

Bardovi-Harlig, K. & Dornyei, Z. (1998).Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, 32 (2), 233-262.

Becher, T. (1989). *Academic tribes and territories: intellectual inquiry and the cultures of disciplines*. In Hyland, K. (2006) (Eds.).*English for Academic Purposes*, (2006), 143-146.

Basturkmen, H. (2011).A genre-based investigation of discussion sections of research articles in dentistry and disciplinary variation. *Journal of English for Academic Purposes*, 1-11.

Berkenkotter, C. (2001). Genre Systems at work: DSM-IV and rhetorical recontextualizationin psychotherapy paperwork. *Written Communication*, 18, 326-349.
DOI: 10.1177/0741088301018003004

Berman, A. (2004). Developing discourse stance in different text types and languages. *Journal of Pragmatics*, 37, 105-124.

Biber, D. and Finnegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect, *Text* 9(1), 93-124.

Biber, D., & Finnegan, E. (1994). Intra-textual variation within medical research articles. *Corpus-based research into language*, 12, 201-222.

Biber, D. (2004). Historical patterns for the grammatical marking of stance: A cross-register comparison. *Journal of Historical Pragmatics,* 5 (1), 107–136.

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes, 5(2),* 97-116.

Bloch, J. (2010). A concordance-based study of the use of reporting verbs as rhetorical devices in academic papers. *Journal of Writing Research, 2* (2), 219-244.

Cargill, M., O'Connor, P., & Li, Y. (2012). Educating Chinese scientists to write in international journals: Addressing the divide between science and technology education and English language teaching. *English for Specific Purposes,* 31(1), 60- 69.

Carter-Thomas, S. & Rowley-Jolivet, E. (2001).Syntactic differences in oral and written scientific discourse: The role of information structure" ASP: *La revue du* GERAS, no. 31-33, 19-38.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An EFL/ESL teacher's course.* Boston: Henile & Heinle.

Connor, U. (2004).Intercultural rhetoric research: Beyond texts. *Journal of English for Academic Purposes* 3, 291–304. www.elsevier.com/locate/jeap

Degaetano, S., &Teich, E. (2011).The lexico-grammar of stance. An exploratory analysis of scientific texts. *Bochumer Linguistische Arbeitsberichte,* 3, 57-66.

Dudley-Evans, T. (2000). Genre analysis: a key to a theory of ESP. *Iberica*, 2, 3-11

El Mamlik, A. T., & Nesi, H. (2008). Publishing in a second language: the case of Sudanese contributors to international medical journals. *Journal of English for Academic Purposes,* 7, 87-96.

Eslami-Rasekh, Z. (2005).Raising the pragmatic awareness of language learners Journal: *ELT Journal*, 59(3), 199-208.DOI: 10.1093/elt/cci039

Flowerdew, J. (2001). Attitudes of editors to non-native speakers' contributions. *TESOL Quarterly,* 35(1), 121-150.

Harwood, N. (2005). 'Nowhere has anyone attempted . . . In this article I aim to do just that': A corpus-based study of self-promotional *I* and *we* in academic writing across four disciplines. *Journal of Pragmatics,* 37, 1207-1231.

Henderson, A., & Barr, R.(2010). Comparing authorial stance in psychology students writing and published research. *Journal of Writing Research*, 2 (2), 245-264.

Hinkel, E. (1994). Native and nonnative speakers' pragmatic interpretations of English texts. *TESOL Quarterly*, 28, 353–76.

Huang, J. C. (2011). *Socializing NNES novice scholars into disciplinary discourse communities: Abstract writing course*. Retrieved from ProQuest Dissertations and Theses-Full Text (Legacy Platform).UMI Number: 3460759

Hunston, S. & Thompson, G. (eds.) (2006). *Evaluation in text*. Oxford: Oxford University Press.

Hunston, S. (2002).*Corpora in applied linguistics.* Cambridge: Cambridge University Press.

Hunston, S. (2011).*Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. N.Y: Routledge.

Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *TEXT*, 18 (3), 349-382.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20(3), 207-226.

Hyland, K. (2002a).Specificity revisited: how far should we go now? *English for Specific Purposes*, 21, 385–395.

Hyland, K. (2002b). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091-1112.

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing, 13* (2), 133-151.doi:10.1016

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192.

Hyland, K. (2006).*English for Academic Purposes: An Advanced Resource Book*. Taylor and Francis e-library.

Hyland, K. (2008).Persuasion, interaction and the construction of knowledge: Representing self and others in research writing. *International Journal of English Studies*, 8 (2), 1-23.

Hyon, S. (1996).Genre in Three Traditions: Implications for ESL.*TESOL Quarterly*, 30 (4), 693-722.

Jafarpour, F.& Taki, S., (2012). Engagement and stance in academic writing: A study of English and Persian research articles. *Mediterranean Journal of Social Sciences* 3 (1), 157-169.Doi: 10.5901/mjss.2012.03.01.157

Johns, A. M., Bawarshi, A., Coe, R. M., Hyland, K., Paltridge, B., & Reiff, M. J. (2006). Crossing the boundaries of genre studies: Commentaries by experts. *Journal of Second Language Writing*, 15(3), 234–249.

Kanoksilapatham, B. (2005).Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, 24, 269–292.

Kelley, G., & Bazerman.Ch. (2003). How students argue scientific claims: A rhetorical semantic analysis. *Applied Linguistics*. 28 (1), 28-55.

Lewin, B., (2005).Hedging: an exploratory study of authors' and readers' identification of 'toning down' in scientific texts. *Journal of English for Academic Purposes*, 4,163–178.www.elsevier.com/locate/jeap

Li, L. & Ge, G., (2009).Genre analysis: Structural and linguistic evolution of the English-medium medical research article (1985–2004). *English for Specific Purposes*, 28, 93–104.

Lightbown, P.M., & Spada, N. (2006).*How Languages are Learned*. London: Oxford University Press.

Nwogu, K.N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes, 16* (2), 119–138. Retrieved from http://www.sciencedirect.com/

Orta, I.V. (2010). A contrastive analysis of the use of modal verbs in expression of epistemic stance in business management research articles in English and Spanish. *Iberia,* 19, 77-96.

Martinez, I. (2001).Impersonality in the research articles as revealed by analysis of the transitivity structure. *English for Specific Purposes,* 20, 227-247.

Maurannen, A. (1993). 'Contrastive ESP rhetoric: Metatext in Finnish–English economics texts. In Hyland, K. (2006) (Eds.). *English for Academic Purposes,* (2006). 148-151.

Molino, A. (2010). Personal and impersonal authorial references: A contrastive study of English and Italian Linguistics research articles. *Journal of English for Academic Purposes*, 9, 86-101.

Myers, G, (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10 (1), 1–35.

Myers, G. (1991). Politeness and certainty: The language of collaboration in an AL project. *Social Studies of Science,* 21, 37-73.

Paxton, M., Pletzen, E., Archer, A., Arend, M., & Chihota, C. (2008). Writer's stance in disciplinary discourses: A developmental view, *Southern African Linguistics and Applied Language Studies*, 26(1), 107-118. http://dx.doi.org/10.2989/SALALS.2008.26.1.9.424

Paltridge, B., (1994). Genre analysis and the identification of textual boundaries. *Applied Linguistics*, 15 (3), 288-299.

Paltridge, B. (2004). Academic writing. *Language Teaching,* 37, 87-105.

Paltridge, B. (2006). *Discourse Analysis.* Continuum International Publishing Group. London.

Precht, K. (2003). Stance moods in spoken English: Evidentiality and affect in British and American conversation. *Text,* 23 (2), 239-357.

Reilly, J., Zamora, A. & Mcgivern, R. (2005). Acquiring perspective in English: the development of stance. *Journal of Pragmatics,* 37(2), 185-208.

DOI: 10.1016/S0378-2166(04)00191-2

Salager- Meyer, F. (2001). From self-highlightedness to self-effacement: A genre-based study of the socio-pragmatic function of criticism in medical discourse. *LSP and Professional Communication,* 1(2), 63–85.

Sheldon, E. (2011). Rhetorical differences in research articles written by English L1 and L2 and Castitalian Spanish L1 writers. *Journal of English for Academic Purposes,* 10, 238-251.

Silver, M. (2003).The stance of stance: a critical look at ways stance is expressed and modeled in Academic discourse. *Journal of English for Academic Purposes*, 2, 359–374.

Stotesbury, H. (2003).Evaluation in research article abstracts in the narrative and hard sciences. *Journal of English for Academic Purposes,* 2, 327–341.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tas, E. (2010). "In this paper I will discuss……..": Current trends in academic writing. *Procedia Social and Behavioral Sciences*, 3, 121-126.

Thompson, G., & Yiyun, Y. (1990). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*12 (4), 365-382.

Va´squez, C. (2007). Moral stance in the workplace narratives of novices. *Discourse Studies,* 9, 653–675.

Williams, I. (1999). Results section of medical research articles. Analysis rhetorical categories for pedagogical purpose. *English for Specific Purposes,* 18 (4), 347-366.

Williams, I. (2006). Move, voice and stance in biomedical research article discussions: a pedagogical perspective. In*: Proceedings of the 5th International AELFE Conference, University of Zaragoza*, 43-51

Appendix A:

A list of functions of using personal pronouns generated by Harwood (2005) and Hyland and Tse (2002):

1- Self promotion at the beginning/ closing of an RA
2- Self citation.
3- Reporting other researchers' Claims (accepting /refuting them).
4- Introducing Procedural Innovation.
5- Methodological pitfalls.
6- Personalizing claims (claiming authority and responsibility).
7- Including the reader to create solidarity.
8- Explaining a procedure
9- Stating results or claim
10- Elaborating an argument
11- Stating a goal/purpose.