

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**FACULDADE DE LETRAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS**

**DANILO DUARTE COSTA**

**LINKING ADVERBIALS IN APPLIED LINGUISTICS RESEARCH**  
**ARTICLES: A CORPUS-BASED STUDY**

**BELO HORIZONTE**  
**FACULDADE DE LETRAS DA UFMG**  
**2015**

**Danilo Duarte Costa**

**LINKING ADVERBIALS IN APPLIED LINGUISTICS RESEARCH**  
**ARTICLES: a corpus-based study**

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Linguística Aplicada.

Área de Concentração: Linguística

Aplicada Linha de Pesquisa: Ensino/  
Aprendizagem de Línguas Estrangeiras

Orientadora: Profa. Dra. Deise Prina Dutra

Belo Horizonte  
Faculdade de Letras da UFMG  
2015

**LINKING ADVERBIALS IN APPLIED LINGUISTICS RESEARCH**  
**ARTICLES: a corpus-based study**

**DANILO DUARTE COSTA**

Dissertação defendida e aprovada em 06, de março de 2015., pela Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Estudos Linguísticos da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do título de "**Mestre em Linguística Aplicada**", na área de concentração de "**Linguística Aplicada**".

---

**Profa. Dra. Deise Prina Dutra - (UFMG) – Orientadora**

---

**Profa. Dra. Patricia Bértoli Dutra - (UERJ) – Examinadora**

---

**Profa. Dra. Bárbara Malveira Orfanó - (UERJ) – Examinadora**

---

**Prof. Dr. Ricardo Augusto de Souza - (UFMG) – Suplente**

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

C837I

Costa, Danilo Duarte.

Linking adverbials in applied linguistics research articles  
[manuscrito] : a corpus-based study / Danilo Duarte Costa. –  
2015.

91 f., enc. : il., grafs., tabs., color.

Orientadora: Deise Prina Dutra.

Área de concentração: Linguística Aplicada.

1. Língua inglesa – Estudo e ensino – Teses. 2.  
Redação acadêmica – Teses. 3. Corpora (Linguística)  
– Teses. 4. Artigos científicos – Teses. 5. Linguística  
– Processamento de dados – Teses. I. Dutra, Deise

CDD : 418

## ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor, Professor Deise Prina Dutra, without whom my academic life in the Master's Program would not even have started. Professor Deise was the first to encourage me to pursue larger steps towards a seemingly very far away goal, and the first to believe that it was possible. I am also extremely grateful for all her comments on this study throughout the whole writing process, which has certainly made it much more meaningful and coherent.

A big thank you to my parents, who always believed that this journey could be possible - even when I myself came to think otherwise. Thank you for always being by my side. A big thank you to my grandparents Juarez (*in memoriam*) and Sônia, who helped my parents finance my English language studies when, in hard times, we needed it. Another big thank you to my godmother Mônica, who has been an inspiration to my studies in Linguistics as well as an outstanding teacher who I most admire. A big thank you to my beloved Cacá for supporting me and for being so understanding and patient when I could not make myself present.

A special thank you to Professor Ana Larissa Adorno Marciotto for sharing her great expertise with me, for always being so supportive, and for always giving me great advice.

I would also like to thank CAPES for the financial support given these two past years.

Last but not least, a huge thank you to my sweetest mates Carolina Bohórquez and Janaína Henriques for always walking along this path hand in hand with me, for the great support, advice, reflections, chats, and, above all, their true friendship.

## AGRADECIMENTOS

Antes de tudo, gostaria de agradecer à minha orientadora, Professora Dra. Deise Prina Dutra, sem a qual minha vida acadêmica no programa de mestrado não teria sequer sido iniciada. Obrigado, Professora, por ser primeira a me incentivar a galgar passos mais longos em direção ao que parecia uma meta muito distante, e a primeira a acreditar que isso seria possível. Sou, também, imensamente grato por todos seus comentários durante todo o processo de escrita deste trabalho, tornando-o muito mais relevante e coerente.

Um muito obrigado a meus pais, que sempre acreditaram que esta jornada seria possível, mesmo quando eu mesmo cheguei a pensar que não. Obrigado por sempre estarem ao meu lado. Um grande obrigado a meus avós Juarez (*in memoriam*) e Sônia, que ajudaram meus pais a custear meu curso de inglês quando, em momentos difíceis, muito precisamos. Outro muito obrigado à minha madrinha Mônica, que foi uma inspiração para meus estudos em linguística, bem como uma exímia professora em quem sempre me espelharei. Um muito obrigado ao meu amado Cacá pelo apoio e por ser tão compreensivo e paciente nos momentos em que não estive presente.

Um obrigado especial à Professora Ana Larissa Adorno Marciotto por compartilhar comigo sua expertise, pelo grande apoio e pelos ótimos conselhos.

Gostaria, ainda, de agradecer à CAPES pela bolsa de estudos concedida nestes dois últimos anos.

Por último, mas não menos importante, um enorme obrigado às minhas queridas Carolina Bohórquez e Janaína Henriques por sempre andarem de mãos dadas comigo por este caminho, pelo apoio, conselhos, reflexões, conversas e, sobretudo, pela amizade verdadeira.

*“If there is some end of the things we do, which we desire for its own sake, clearly this must be the good. Will not knowledge of it, then, have a great influence on life? Shall we not, like archers who have a mark to aim at, be more likely to hit upon what we should? If so, we must try, in outline at least, to determine what it is.”*

Aristotle. *Nicomachean Ethics*. (1.1094a18). c. 325 BC.

## ABSTRACT

This study aims at investigating the use of linking adverbials (Biber et al., 1999) in research articles written by Brazilian English L2 applied linguists in comparison to those written by English L1 professionals of the same field. Two comparable corpora have been compiled for this study, namely CRAB (Corpus of Research Articles written by Brazilians) and CRAN (Corpus of Research Articles written by Natives), both containing more than 300,000 tokens. The corpora compilation process followed strict methodological procedures based on Biber (1993) and McEnery et al. (2006).

The data, after undergoing the Log-Likelihood statistical test, were analysed using the software AntConc 3.4.2 for a qualitative examination. Seven different semantic categories of linking adverbials were investigated so as to find similarities and differences in the use of those linguistic elements between the two corpora.

The results show that there are significant differences in the use of linking adverbials in Brazilian academic writing in comparison to native speakers'. These differences are in both frequency of use (over and underuse of some forms) and the way in which the linking adverbials are employed in texts. In addition, we have found that there are adverbials which are, at times, misused by Brazilian writers.

**Key-words:** linking adverbials, academic writing, corpus-linguistics

## RESUMO

Este estudo se propõe a investigar o uso de *linking adverbials* (Biber et al., 1999) em artigos científicos de linguística aplicada escritos em inglês por brasileiros, em comparação com aqueles escritos por falantes nativos de inglês. Dois corpora comparáveis foram compilados para este estudo, a saber, CRAB (*Corpus of Research Articles written by Brazilians*) e CRAN (*Corpus of Research Articles written by Natives*), ambos com mais de 300.000 palavras. O processo de compilação dos corpora seguiu rigorosos procedimentos metodológicos embasados em Biber (1993) e McEnery et al. (2006).

Os dados, depois de submetidos ao teste estatístico *Log-Likelihood*, foram analisados utilizando o software *AntConc* 3.4.2 para uma análise qualitativa. Sete diferentes categorias semânticas dos *linking adverbials* foram investigados de forma encontrar semelhanças e diferenças na utilização desses elementos linguísticos nos dois corpora.

Os resultados mostram que existem diferenças significativas no uso de *linking adverbials* na escrita acadêmica dos brasileiros em comparação à dos falantes nativos. Essas diferenças dizem respeito tanto à frequência de uso (sobre e sub-uso de algumas formas), quanto à maneira pela qual tais elementos são empregados em textos. Além disso, foi observado que existem *linking adverbials*, por vezes, mal utilizados nos textos escritos pelos profissionais brasileiros.

**Palavras-chave:** *linking adverbials*, escrita acadêmica, linguística de corpus.

## LIST OF ABBREVIATIONS

BNC	British National Corpus
CL	Corpus Linguistics
COCA	Corpus of Contemporary American English
CRAB	Corpus of Research Articles written by Brazilians
CRAN	Corpus of Research Articles Written by Natives
EFL	English as a Foreign Language
ESL	English as a second Language
L1	First language
L2	Second Language
LL	Log-Likelihood
NNS	Non- native speakers of English
NS	Native speakers of English

## LIST OF GRAPHS

<b>Graph 1:</b> Overall distribution of linking adverbials in CRAB, CRAN, and Biber et al. (1999).....	52
<b>Graph 2:</b> Overall distribution of linking adverbials by semantic category in three corpora: CRAB, CRAN, and Biber et al (1999).....	54
<b>Graph 3:</b> Distribution of linking adverbials among different semantic domains in CRAB and CRAN.....	55
<b>Graph 4:</b> Overuse and underuse of linking adverbials in CRAB in comparison to CRAN.....	56
<b>Graph 5:</b> Distribution of linking adverbials belonging to the semantic category of enumeration in CRAB and CRAN.....	58
<b>Graph 6:</b> Distribution of linking adverbials belonging to the semantic category of addition in CRAB and CRAN .....	60
<b>Graph 7:</b> Distribution of linking adverbials belonging to the semantic category of apposition in CRAB and CRAN. ....	63
<b>Graph 8:</b> Distribution of linking adverbials belonging to the semantic category of result/inference in CRAB and CRAN. ....	66
<b>Graph 9:</b> Distribution of linking adverbials belonging to the semantic category of contrast/ concession in CRAB and CRAN.....	69
<b>Graph 10:</b> Syntactic realisation of linking adverbials in CRAB and CRAN.....	72

## LIST OF TABLES AND BOXES

<b>Table 1:</b> Summary of linking adverbials' form, meaning, and use in three grammar books.....	35
<b>Box 1:</b> Quirk's subdivision of semantic categories for linking adverbials.....	36
<b>Table 2:</b> Biber et al.'s classification of linking adverbials according to semantic categories. ....	38
<b>Table 3:</b> Biber et al.'s vs. Huddleston & Pullum's classification of linking adverbials according to semantic categories.....	40
<b>Table 4:</b> Journals used as a source of texts for the compilation of CRAB and CRAN.....	46
<b>Table 5:</b> Steps involved in the clean-up process of CRAB and CRAN. ....	48
<b>Table 6:</b> Total count per million words of therefore and thereby in the BNC and COCA. ....	67

## LIST OF FIGURES

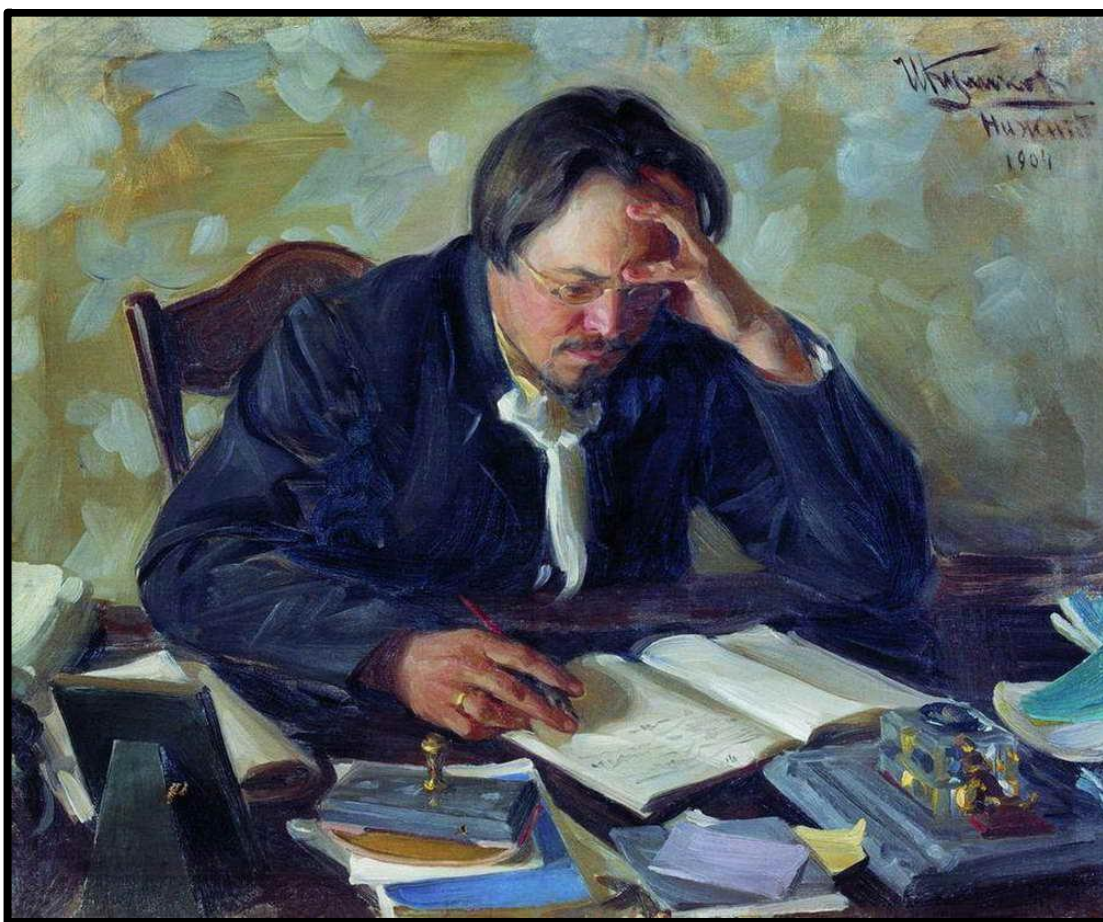
<b>Figure 1:</b> Portait of E.N Chirikov, by Russian Painter Ivan Kulikov, 1904.....	16
<b>Figure 2:</b> Illustration of the Principle of Lateral Continuity.....	18
<b>Figure 3:</b> The of conjunction .....	30
<b>Figure 4:</b> Syntactic realisations of linking adverbials in conversation and academic prose .....	73

## INDEX

<b>1. INTRODUCTION</b> .....	17
1.1 The motivations for this research .....	17
1.2. Research aims .....	21
1.2.1 Research goals.....	22
1.2.2. Research Objectives.....	22
1.3. The research questions .....	22
1.4. Relevance of this study.....	23
<b>2. LITERATURE REVIEW</b> .....	25
2.1. Textual Cohesion .....	25
2.2. Understanding Linking Adverbials .....	31
2.2.1. The conceptualisation of linking adverbials .....	31
2.2.2. Usage patterns of linking adverbials.....	35
2.2.3. Related corpus-based work on linking adverbials .....	41
<b>3. METHODOLOGY</b> .....	44
3.1. The corpora used.....	44
3.1.1. The corpora compilation process.....	45
3.1.2. The cleaning-up process .....	47
3.2. Procedures for data analyses .....	49
<b>4. DATA ANALYSIS</b> .....	52
4.1. Data Overview .....	52
4.2. Distribution according to separate semantic categories .....	57
4.2.1. Enumeration .....	57
4.2.2. Addition.....	60
4.2.3. Apposition.....	63
4.2.4. Contrast / Concession .....	68
4.2.4. Summation.....	71
<b>5. CONCLUSION</b> .....	75
REFERENCE.....	79
APPENDIX A - Disciplines considered for the aila world congress 2014 .....	84
APPENDIX B - List of linking adverbials searched for in this study.....	86

APPENDIX C - Graphs containing linking adverbials according to semantic categories: statistically significant and non-significant adverbials .....87

Fig. 1: Portait of E.N Chirikov, by Russian Painter Ivan Kulikov, 1904



## 1. INTRODUCTION

### 1.1 The motivations for this research

The process of writing has always been a difficult one. And, by using *always*, I intend to include times when the mere notion of pen and pencil were seen as highly advanced tools of which writers disposed in order to mediate the complex course of putting ideas into paper. The painting we see in the previous page is an attempt to illustrate these lines. It depicts a man who seems immersed in a thoughtful silence, as though struggling with ideas that simply do not come by easily. The sombre shades of dark colours, as well as the man's resolute-looking mood, give us the impression that the intellectual venture he has embarked on will be a long one.

And this seems to be true for any type of writing; be it a simple and informal love letter, in which the wrong combination of words could result in a disastrous breaking of someone's heart, or even a highly technical scientific article, which requires that certain genre-based rules be obeyed so as to reach the largest audience possible with clarity and precision. Undeniably, putting words together in writing is extremely demanding when compared to simply uttering the very same proposition through speaking. As Emig (1977) puts it, "writing is an artificial process", besides being "stark, barren, [and] even naked as a medium" (p. 27). However, it is also known that writing can be bettered, for it may be learned through instruction and practice (RICHARDS, 1990).

But for any formal instruction to take place, it is first necessary that knowledge on the topic be produced. This knowledge production comes, among other places, from the academic environment, which hypothesises, tests, retests, and, finally, comes to sound conclusions on whatever it is that needs to be made public in benefit of a given population. Since we highly believe that writing can be made easier - insofar as research can cast light on characteristics of this life and academic skill - this study attempts to contribute to the development of the understanding of an explicit feature which is present in any type of academic prose: the so-called linking adverbials, words such as *therefore* and *moreover*.

We are aware of the fact that, if treated in isolation, such a linguistic component of the clause will not cause great impact on one's overall ability to write. Nevertheless, we see this study as a tiny piece that fits into a much larger and complex jigsaw. Indeed, linking adverbials are unnecessary for the achievement of clausal grammaticality (HALLIDAY & HASAN, 1976). Why, then, bother to investigate an item which, grammatically speaking, is of secondary significance? The answer to that question seems quite straightforward. Whereas linking adverbials do not display any syntactic importance in a text, they do play a dominant role in textual cohesion (HALLIDAY & HASAN, 1976; BIBER et. al, 1999). That is because they function as a means of bringing unity to a text, i.e., they enable separate parts of the text - be it clauses, sentences, or paragraphs - to be interpreted as belonging to the same whole.

Nevertheless, the idea of having separate parts form a unity is not limited to linguistics. It seems reasonable to say that, in the natural world, in which language is included for it is fundamentally natural, parts need to be gathered together to come to be or exist. An elementary example would be that of the formation of planet Earth. Mass particles of star dust attracted one another, and those separate parts eventually formed a huge mass; our planet. Similarly, for a text to come to exist, its sentences need to be held together; otherwise, they would just be a group of unrelated sentences. For a group of sentences to form a text, it is first and foremost necessary that cohesive elements be present, thus creating what Halliday and Hasan (1976) called texture. A more concrete analogy can be made between the absence of linking adverbials in a text and a natural geological process.

In geology, a well-known principle is that of lateral continuity, illustrated below.

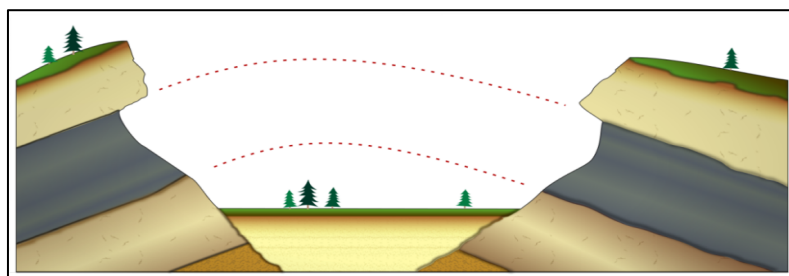


Fig. 2: Illustration of the Principle of Lateral Continuity

What we see in the picture is that the sides separated by the valley necessarily belong together. Because of some sort of erosion process, they are now separated, but, as they are laterally continuous, those sedimented rocks can be said to be part of the same totality, of the same whole. Now, going back to texts, what happens when linking adverbials are *not* present is very similar to what can be seen in the picture. We know that the messages should be understood as parts of the same whole; yet they are separate, as can be noticed in the examples below.

[1] “*Although the learners had not been explicitly taught GM, they successfully deployed it in this isolated writing task. **Therefore**, it is reasonable to expect that, with targeted instruction toward GM use and its effects on textual meaning-making, these learners should ...*” [CRAN, 2a]

[1a] *Although the learners had not been explicitly taught GM, they successfully deployed it in this isolated writing task. [Ø] It is reasonable to expect that, with targeted instruction toward GM use and its effects on textual meaning-making, these learners should ...*

Example [1] shows how *therefore* establishes a link between two sentences, explicitly marking a connection between them. On the other hand, in example [1a] the linking adverbial was intentionally left out. Despite the absence of *therefore*, the two sentences can still be interpreted as having some connection. However, it feels like there is something missing between them. A text without linking adverbials could, thus, look somewhat like a discontinuous lining of sedimented rocks; one whose parts can be identified as belonging to the same whole, but that presents a discontinuous connection. Hence, linking adverbials should be seen as linguistic elements which essentially serve the purpose of bringing a sense of semantic connection (i.e. cohesion) to messages of a text. Without them, a certain text could be rather difficult to be interpreted, for the ideas it presents would not logically follow one another, but rather, simply be sequenced in space.

In this sense, when reading a well written text, it seems quite easy to distinguish it from any other piece of writing which is not as successful in conveying clear and connected messages. More specifically, in the case of academic writing, where complex ideas and concepts need to be explored, learners tend to have a much harder time trying to make messages interact clearly and cohesively in comparison to expert writers (SHAW, 2009; LEI LEI, 2012). And that is even more evident when we deal with texts written in L2.

In the 1990's, as a student learning EFL (English as a Foreign Language), I experienced those difficulties myself. I would always write a text and have the feeling that what I was handing in to my teacher could be one of my masterpieces. It was not until the following week, when the same text was handed back to me with the teacher's corrections and comments, that I realised how confusing some of the ideas in the text were. Nevertheless, it is well known that reading plays a very important role in writing expertise, and by being exposed to a large number of academic texts throughout my undergraduate course, I realised that, to some extent, I managed to develop better writing skills.

Whenever reading academic texts in English, I observed that native speakers (NS) would often draw on a vast repertoire of linking devices, which, at times, I did not even know existed. Never before had I found words such as *notwithstanding* or *hence* in any other type of text. Words such as those would promptly catch my attention, and, every time a written assignment was given by a professor, I was eager to use one of those words in my tasks. But that was a naive mistake. What I initially believed could contribute to turning my text into something more academic-like, was, in fact, making it rather broken. That is because, as many inexperienced writers, I was overusing linking adverbials in order that the text could look like those written by mature professional writers. As asserted by Conrad (1999, p. 5), "overuse of adverbials can make a text sound awkward". Indeed, both overuse and underuse of linking adverbials negatively impact on the overall comprehension of a text.

Ultimately, I was intrigued to know whether professional non-native speakers (NNS) would use linking adverbials with the same ease and frequency in their

texts as their NS counterparts. A quick informal search for linking adverbials in texts written by NNS showed that indeed there are differences in the way they are used by expert Brazilian English-L2 writers in comparison to NS expert writers. Although the concept of professional writer is synonymous to expert writers, and their texts are seen as role models, there still seems to be room for improvement. Thus, I realized that it could be interesting to study the way these NNS professionals use linking adverbials in their texts, since identifying problematic usage - such as the overuse or the lack of variety of linking adverbials - could perhaps shed light to the way such a topic can be approached in the teaching of writing in the context of EFL academic writing courses.

Admittedly, any research dealing with large amounts of natural language data is required to make use of technological tools so as to be able to come to reliable results. In this regard, the knowledge produced by corpus linguists in the few past decades seems the most adequate to methodologically ground this study. After all, if one intends to pursue empirical research, they necessarily have to use empirical data, which, in this case, comes from texts that make up our corpora. In direct opposition to that view is the idea that conclusions about a language can be drawn by resorting to our intuition. However, when it comes to detailed quantitative data, intuition fails to be a consistent means of achieving reliable results. This view is shared by a number of contemporary linguistics, among whom Sampson (2002), who maintains that data that come from intuition are “hopelessly unreliable” (p.2).

## **1.2. Research aims**

It is known that learners of an L2 considerably develop their writing skills as they go on studying the target language. It is also true that once an individual reaches a high proficiency level, his linguistic performance closely compares to that of natives. However, it seems that even highly proficient L2 writers present some features in their writing that show some incompatibility to that of NS's. Thus, this research aims at better understanding the use of linking adverbials in research articles produced by proficient Brazilian professional English L2 writers

in the field of applied linguistics. Below we present the general aims (goals) and specific aims (objectives) of this research.

### **1.2.1. Research goals**

The goal of this research is to describe the use of linking adverbials in academic research articles written in English by Brazilian professionals in the field of applied linguistics. We wish to understand how those linguistic components of the clause are used by comparing NS' and NNS' written production. We believe that by learning more about the writing of L2 professionals, generalizations can be drawn so as to contribute to the teaching of writing skills to learners of academic L2 English.

### **1.2.2. Research Objectives**

This research has three main objectives, all of which relate to the written production of scientific articles. They are as follows:

1. Investigate differences in the use of the various linking adverbials found in Brazilian professional academic writing in comparison to English L1 professional academic writers;
2. Describe how linking adverbials used by Brazilians differ in terms of semantic category and syntactic realisation;
3. Verify if any linking adverbial is inaccurately used by Brazilians.

## **1.3 The research questions**

The objectives just present can, thus, open into our research questions, which will help to present our results. They are as follows:

- A. To what extent do Brazilian academic writers' use of linking adverbials differ from English L1 professional academic writers?
- B. As far as semantic categories and syntactic forms of linking adverbials are concerned, how do Brazilians compare to English L1 professional writers?

C. Is there any linking adverbial which is inaccurately used by Brazilians?

#### 1.4 Relevance of this study

In Brazil, almost all theses and dissertations produced at the end of the master's and doctoral studies are written in Portuguese. However, for those who advance the academic-scientific career, the extensive work that easily exceeds 100 pages, gives way to smaller productions - though equally important - which have more defined objectives and are more far-reaching: the scientific articles. The publication of this kind of work by academics and researchers who are affiliated with research centres is an imperative demand, especially because it is through these articles that the intellectual exchanges that enable the advancement of scientific knowledge is materialized .

If, on the one hand, scientific publications are beneficial from the standpoint of knowledge, on the other hand they are very demanding for those who produce those publications. The pressure to publish suffered by the academics is not at all a recent one. In the early twentieth century, Coolidge (1932) coined the term *Publish or Perish*, suggesting the need to publish scientific papers to support a successful academic career. More recently, this pressure is something commonly talked about in university campuses and reported by researchers such as Flowerdew (2008) and Filho (2012), besides Disney (2001), who criticizes the fact that this pressure leads to research being produced in bulks, alluding to the mass production of a factory.

It occurs, however, that publishing research in national journals is not enough. The most important research agencies suggest that researchers publish their articles also in international journals in order to generate greater recognition of the studies, in addition to contributing to the increase in funding for their universities. And, if the exchange of ideas and findings on a global level is one of the goals of researchers, it is only natural that the work be written in English. As pointed out by Swales (1990, p.99), "there is no doubt that in the world English has become the predominant language of research."

However, there are a number of non-overt obstacles that hinder the publication of scientific papers written by researchers from developing countries (GIBBS, 1995; FLOWERDEW, 2008; CURRY & LILLIS, 2010). These obstacles imposed on nations that are not part of the so-called inner circle, i.e. countries whose first language is English (KACHRU, 1985), has resulted in prejudice which was symbolically inherited, as the so-called outer-circle countries do not have English as their first language, besides finding themselves in disadvantaged economic conditions. As stated by Flowerdew (1999), research written in a foreign language necessarily imposes limitations on those who write them.

Even if language-related issues are not the primary factor of rejection of an article by a journals' evaluation committee, work written by non-native speakers of English receive nearly four times more negative reviews than those written by native speakers, specifically in relation to components of the surface structure of the text, i.e. syntax/grammar, lexicon, cohesion, etc.(HEWINGS, 2006).

In addition, this research is, to our knowledge, the only study in this country that addresses the use of linking adverbials in professional academic writing. From what has been published, it seems that researchers tend to value, to a greater extent, learner written production over professional academic writing. In view of that, it seems of the utmost importance that academic writing be explored as this is an area of investigation from which L2 professional academic writers can highly benefit.

Ultimately, we expect that by apprehending the differences in L1 and L2 professional academic writing, we can shed light to the teaching of academic writing, as any incompatibilities found in mature writing of today may be avoided by professional writers of tomorrow.

## **2. LITERATURE REVIEW**

This chapter will address linking adverbials from a more general to a more specific discussion. We begin by dealing with aspects of textual cohesion, and then move on to discuss linking adverbials more explicitly, according to their meaning and form. Finally, in the end of the chapter, we bring some corpus-related work on the topic.

### **2.1. Textual Cohesion**

Biber et al. (1999) state that linking adverbials are important devices for the creation of textual cohesion. For this reason, the paragraphs that follow will briefly discuss both what it means to say a text is cohesive and present cohesive devices that writers commonly - consciously or not - resort to when producing texts.

On a preliminary account of cohesion, Halliday and Hasan (1976) claim that the concept is semantic rather than structural, and state that "it refers to the relation of meaning that exists within the text, and that define it as a text" (p.4). At first glance, this assertion appears to be in direct conflict to that exposed by Biber et al. (1999). The reasoning behind this assertion is that if cohesion can define what a text is, without the presence of linking adverbials a certain piece of writing would run the risk of not being considered a text. Nevertheless, a closer reading of Halliday and Hassan's principles reveals that this is not the case, and for a better understanding we should first know what a text is in Halliday and Hasan's terms.

For these authors, a text is language in use or language that is functional. This means that, to be a text, a certain piece of language (be it written or oral) needs to possess a sense of unity, as opposed to a group of isolated or unrelated sentences. These unrelated sentences can be found, for instance, in a textbook exercise aimed at teaching grammatical aspects of a foreign language, or on a classroom board, on which a teacher simply jotted down a few words about a given subject under discussion. A text can, therefore, be best viewed as "a unit of meaning" (HALLIDAY & HASAN, 1976, p.2), and despite the fact that "it looks

as though it is made of words or sentences, it is really made of meanings” (cf., 1989, p.10).

Another important concept in the definition of what a text is comes from the notion of *texture*. Texture is what distinguishes a text from something that is not a text and, adding to that, it can be said to be revealed by semantic relations between individual messages. Below is an example that illustrates what is being stated here, based on Halliday & Hasan (1976). It is worth mentioning that this next example, as well as all the others in this study, were extracted from our corpora<sup>1</sup>.

[2] “Moreover, we need to teach **our students** how to “consult” to enhance vocabulary learning. We can encourage **them** to repeat the word out loud alone...” [CRAN, 27a]

It is noticeable from the example that *them* in the second sentence refers back to *our students* in the first sentence. This means that the two items are identical in reference, i.e. they refer to the same thing, thus allowing for the interpretation of the two sentences as forming a unity. It is exactly this meaning relation that brings texture to a text. Therefore, the more cohesion there is in a text, the more texture it presents, which, in turn, makes it a meaningful piece. Notably, texture is perceived in the text as a whole and not only in separate parts of it. As described by Crane (2006, p.131) “a text without texture would just be a group of isolated sentences with no relationship to one another”. But example [2] is just a fragment of a larger text, as can be seen from the punctuation (the ellipsis) at the end of the second sentence. In order to judge whether the whole text indeed has texture, one would have to find more pairs of cohesively related items. These pairs, analysed individually, are called *cohesive ties*.

As Halliday & Hasan (1976, p.3) put it, a tie is “a single instance of cohesion”. In this sense, the relationship between *our students* and *them* above can be said to form a tie. Evidently, the formation of a tie presupposes the presence of two items which will necessarily have to be related, and such a relation should be viewed from the perspective of meaning for it is essentially semantic in nature.

---

<sup>1</sup> CRAB is the corpus of texts written by Brazilians, and CRAN the corpus of texts written by native speakers of English. More details about them are given in the next chapter.

Hence, if different messages follow one another forming a text, there needs to be something that will allow those messages to be interpreted as being part of the same whole. What makes that possible is the presence of these semantic ties, which form the basis for cohesion in a text. Based upon this logical reasoning, the authors give an interesting account of the concept of cohesion:

“Cohesion occurs where the interpretation of some element in the discourse is dependent on that of another. The one presupposes the other, in the sense that it cannot be effectively decoded except by recourse to it. When this happens, a relation of cohesion is set up, and the two elements, the presupposing and the presupposed, are thereby at least potentially integrated into a text.” (HALLIDAY & HASAN, 1976, p.4)

The cohesive relationship between the presupposing element and its presupposed counterpart may be, according to Halliday and Hasan (1976), classified into five distinct categories, namely reference, substitution, ellipses, lexical cohesion, and conjunction. With reference to example [2], cohesion is provided by the link established between *them*, which has no meaning alone, and *our students*, which fills it with meaning. This type of relation serves as an example for the first of the above listed categories: reference.

The next category in the list is substitution, which occurs when we replace an item for another. For instance, nominal groups can be replaced by the pronoun *one* (e.g. ‘her best apron, the white one’), whereas verbal and clausal groups accept *do* and *so* respectively as possible substitutes. In line with substitution is the category termed ellipses, which would be the omission of an item in a clause without compromising the meaning of what has been omitted. An example would be ‘He's leaving but I'm not’, where *leaving* need not be repeated. Due to the close association between these two last categories, some authors prefer treating them as a single resource, including Halliday & Hasan (1989) and Halliday & Matthiessen (2004). Martin (2001), for instance, explains that, from the point of view of the English language, ellipses would be a substitution for zero, whilst if we look across other languages, it is best “to think of substitution as ellipses (signaled) by something” (p. 48).

The fourth category is what we call lexical cohesion. This type of cohesive device is achieved when the repetition of lexical items serve the purpose of unifying distinct messages, for instance through the use of reiteration (including synonymy, repetition of the same word, use of a superordinate word, and use of a general word), as can be seen below:

[3] “...writers must display a competence as disciplinary insiders. **This competence** is, at least in part, achieved through a writer-reader dialogue which . . .” [CRAN, 16a]

[4] “When present, adverbs appear after the verb. **This grammar** is unlike English in many respects . . .” [CRAN, 13a]

Example [3] shows a cohesive relation between one lexical item to another by means of a reiteration of the same word. In this case, the noun phrase *this competence* has the noun *competence* as its referent. When this happens, the second item very frequently carries an anaphoric determiner (e.g. the/*this*) that refers back to the work it is linked to. The use of an anaphoric determiner is not used exclusively when the same word is being repeated, as we can see from example [4]. There, *grammar* is being used as a general word which encompasses the meaning of the whole preceding underlined clause, thus establishing a cohesive relation between the two items.

Finally, the last category of ties, and the one which is of greatest interest to this study, is that of conjunction, realized through what are called conjunctive adjuncts or conjuncts. Our interest in conjunction comes from the fact that linking adverbials can be placed under the umbrella of this type of cohesive relation, albeit for Halliday and Hasan (1976) the notion of conjunction encompasses a much larger set of possible realizations subdivided into four categories: additive, adversative, causal and temporal. In any case, the authors explicitly argue that conjunction is different from the other cohesive relations already discussed, as it is neither grammatical nor lexical. Reference, substitution and ellipses are termed grammatical cohesion, whereas lexical cohesion, as its name implies, is not. Conjunction, on the other hand, falls somewhere between these two ends. It is important to point out, however, that

although these grammatical/lexical labels are used, the cohesive relation is, as mentioned before, fundamentally semantic.

Another feature that distinguishes a conjunction is the fact that it does not simply establish an anaphoric relation, but “a specification of the way in which what is to follow is systematically connected to what has gone before” (p. 227). In addition, they are considered distinct from other ties in that they can link either short stretches of discourse (e.g. phrases), or be used as a means of connecting longer messages conveyed by sentences or even paragraphs.

Interestingly, the ideas initiated by Halliday and Hasan (1976) are further extended by Halliday and Matthiessen (1994, 2004). When discussing the class of conjunction, they subdivide it into three subclasses, namely *linker*, *binder* and *continuative* (the two first being considered ‘conjunction proper’). Linkers include coordinating conjunctions such as *and* and *but*, in addition to linking adverbials (to use the term preferred in Biber et al., 1999) like *however* and *therefore*. Binders include nominal/complement clause subordinators such as *that* and *whether*, along with adverbial clause subordinators like *because* and *before*. As for the continuative, the authors explain it encompasses words which signal a move in discourse, such as *well* and *oh*, typically occurring in dialogical texts.

Although Halliday and Matthiessen propose this new classification for conjunction, it is used mainly when discussing the structure of clauses. When speaking specifically of cohesion, they seem to abandon the original classification and go on to establish logic-semantic relations that distribute conjunctive elements into *elaboration*, *extension* and *enhancement*. In this sense, the first classification of conjunction into additive, adversative, causal and temporal is largely enhanced to encompass a much wider scope of possible semantic meanings, as shown in Fig. 2.

Taking into account that Halliday himself expanded his views on the topic of cohesion, it is only natural that a range of other scholars would do the same. However, while some followed the lines proposed in the hallidayan approach, others deviated and suggested new ways of dealing with textual cohesion. For instance, Martin (1992), cited in Martin (2003), moved away from the notion of

ties which is present in Halliday and Hassan (1976) and into a more abstract level than that of lexicogrammar, namely identification, negotiation, conjunction, and ideation. In either case, there seems to be a general consensus among academics that cohesion is a fundamental device for the construction of meaning relations within the text, making it intelligible and meaningful.

As we have seen, textual cohesion can be realised through various different forms. But our investigation will focus on one of its realizations; that which is achieved when linking adverbials are present in the text. The notion behind linking adverbials is what is discussed in the section that follows.

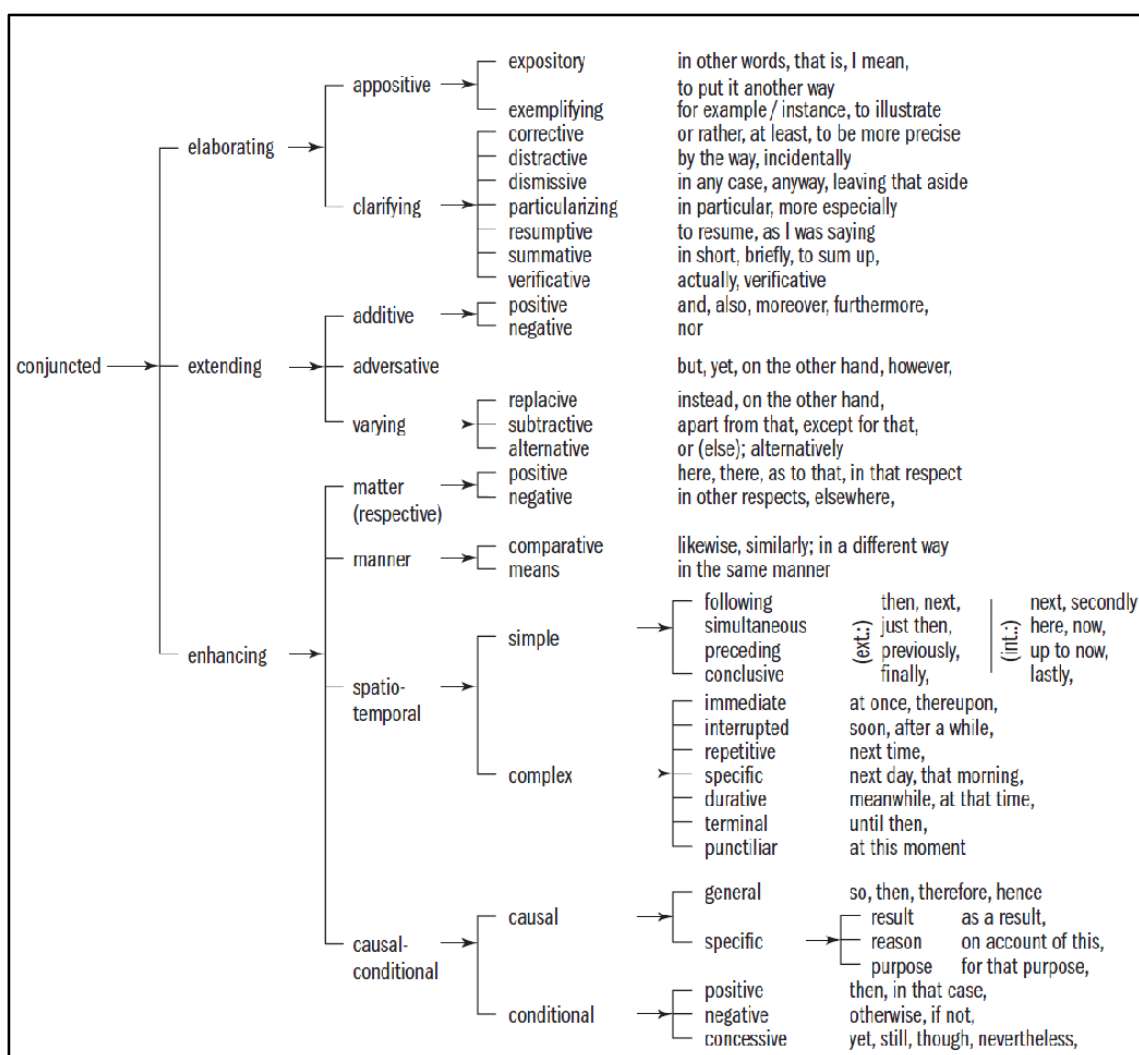


Fig. 3: The system of conjunction  
Source: Halliday & Matthiessen (2004, p. 541).

## 2.2. Understanding Linking Adverbials

This section will introduce the concept of linking adverbials. We begin by discussing the varied terminology in the field of linguistics used to refer to that type of connector. Next, a more structural approach aimed at analysing their meaning and form is taken. Lastly, a brief inventory of relevant studies on the topic is brought to illustrate what has so far been researched concerning these linguistic elements.

### 2.2.1 The conceptualisation of linking adverbials

The term linking adverbials was first introduced in the literature by Biber et al. (1999), and, since then, this term has been widely used in studies in the field of linguistics. However, it has also been given various different names in different grammar books. Quirk et al (1985), for instance, in their *Comprehensive Grammar of the English Language* use the term 'conjuncts'. Greenbaum (1996), in *The Oxford English Grammar*, employs the term 'logical connectives'. As for *The Grammar Book*, by Celce-Murcia & Larsen Freeman (1999), the preferred terminology was 'conjunctive adverbials'. Also, for Huddleston and Pullum (2002), the term chosen in *The Cambridge Grammar of the English Language* was 'connective adjuncts', whereas Carter & McCarthy (2006), in the *Cambridge Grammar of English*, chose 'linking adjuncts'.

Nevertheless, despite the great variety in terminology, all of those authors agree that linking adverbials serve the function of connecting segments of discourse by signalling semantic relations among messages in a text. They also seem to agree on the separation between linking adverbials on one side, and coordinating (*and, but*) and subordinating (*although, when*) conjunctions on the other<sup>2</sup>. That is because the former category has its scope over larger units of discourse (clauses, sentences, or paragraphs) besides exclusively establishing semantic relations, while the latter would only be able to link shorter units at or below clausal level, such as phrases or clauses, in addition to having a

---

<sup>2</sup> Except for Celce-Murcia & Larsen Freeman (1999), who state that, when considering the clause-connecting level, there is no reason to separate subordinating conjunctions and logical connectors.

syntactic function, which is probably its most distinguishing feature. This difference can be seen more clearly in the examples that follow.

- [5] “*Gender orientation as a focus of oppression is present, **but** its manifestations in the literature of applied linguistics need to be greater.*” [CRAN, 39a]
- [6] “**Although** *it can be encouraged by teaching approaches as well as by teachers, autonomy is a complex phenomenon which emerges from the interaction of elements ...*” [CRAB, 04]
- [7] “... *these types of errors appear to constitute the most frequently occurring mistakes that students make in freely produced texts (Schulze, 2001, p. 121). **However**, the parser approach also produces its share of false acceptances...*” [CRAN, 44a]

All of these examples illustrate different items establishing relations of contrast between clauses or sentences. Example [5] illustrates the use of *but* in its most prototypical function: a coordinating conjunction. Its role in the sentence is to link two independent clauses, i.e., clauses that have the same syntactic role and can stand on their own. Example [6] shows the subordinating conjunction *although* also linking two clauses. However, the clause it introduces is linked to the next one by a relation of dependence, i.e., one needs the other so as to make complete sense. In contrast to these two examples, example [7] presents the linking adverbial *however*, which connects a larger span of text (two sentences) and should be analysed as being above the clause level, for its only function is to connect messages without presenting any syntactic function. This absence of a syntactic role in *however* may be noticed by the fact that it can be removed from the sentence without compromising its grammaticality. But, if that were the case, the text would run the risk of losing cohesion and, consequently, texture.

The examples shown above demonstrate why linking adverbials are treated differently when it comes to connecting sentences and clauses. Nevertheless, in some cases, there appears to be an overlap of categories. When this happens,

coordinators behave somewhat like linking adverbials, as can be seen through the use of *yet* in the following examples.

[8] *“This phrase is clearly beyond the limit of 8-10 words, **yet** native speakers can say it without hesitation.”* [CRAN, 42a]

[9] *“According to the suggested answer, the formal variety is the standard; and the informal one, in this activity, comprises all the other language variants, not being helpful for the students’ written production. **Yet**, the organization of the activity takes for granted that students and teachers know the genre specificities...”* [CRAB, 42]

In example [8], *yet* is functioning as a coordinator, linking two independent clauses. In contrast, example [9] shows *yet* linking two sentences: one compound and the other complex, thus establishing a much more elaborate connection between messages. In this case, *yet* could be easily replaced by *however* or *nevertheless*, which makes us believe that it is a linking adverbial, and not simply a coordinator. Nonetheless, a structural analysis of those items reveals that they are indeed coordinators, not linking adverbials. Putting it simply, adverbials have a flexible position in clauses, while coordinators do not. Changing the position of *yet* would result in an ungrammatical sentence, but using *however* would certainly be acceptable, as shown below.

[10a] *\*The organization of the activity, **yet**, takes for granted that students and teachers know the genre specificities...”*

[10b] *The organization of the activity, **however**, takes for granted that students and teachers know the genre specificities...”*

Similarly, other coordinators can be misleadingly understood as linking adverbials in sentences. For instance, *and* may be interpreted as an adverbial such as *in addition* or *moreover*, whereas *but* may be said to acquire the same value as that possessed by *however* or *nevertheless*, as can be seen in the examples that follow.

[11] *“People can also indicate objects by touching them, exhibiting them, turning toward them, or standing in front of or behind them. **And** they can*

*indicate moments in time by vocalizing or tapping at those moments (Clark, 1996; Clark and Fox Tree, 2002).*" [CRAN, 17a]

- [12] *"The visual use of flags, a motto, and the heraldic motif of crossed symbols, coupled with the civic speech asserting the liberal ideals of the sovereign subject, could be seen as a co-optation of the pure notion of civic discourse, its perversion by anti-democratic forces.*

*(Paragraph)*

*But it is not clear that such a civic norm has ever been achieved.*" [CRAN, 11a]

The fact that *and* assumed the unusual function of connecting long stretches of text, and that *but* is being used to link a sentence to an entire preceding paragraph, also do not grant these coordinators the status of a linking adverbial. Again, moving these items in the clause would generate syntactically malformed sentences, whilst for their supposedly structural synonyms, the non-initial position is entirely acceptable.

- [13a] *\*They can, and, indicate moments in time by vocalizing or tapping at those moments...*

- [13b] *They can, moreover, indicate moments in time by vocalizing or tapping at those moments...*

- [14a] *\*It is not clear, but, that such a civic norm has ever been achieved.*

- [14b] *It is not clear, however, that such a civic norm has ever been achieved.*

Despite the fact that coordinators are syntactically different from linking adverbials, it has been shown that, at times, they do present very similar functions in discourse, as illustrated above. In those cases, they seem to be good substitutes for linking adverbials and, therefore, are included in the range of possibilities writers dispose of when connecting messages in a text. As defended Shaw (2009), *but*, for instance, directly competes with *however*, and information may be lost if that is not considered in a study. For this reason, the instances of coordinators identified as playing a directly similar role as that of adverbials will be included in this study for comparability.

### 2.2.2. Usage patterns of linking adverbials

For this discussion, three renowned grammar books have been consulted: A Comprehensive Grammar of the English language (QUIRK et al., 1985); The Longman Grammar of Spoken and Written English (BIBER et al., 1999); and The Cambridge Grammar of the English Language (HUDDLESTON & PULLUM, 2002).

In most studies on linking adverbials, three main features of can be very commonly found: form, meaning, and position of the adverbials in the clause. *Form* is understood as the different possible syntactic realisations of linking adverbials, such as through a single adverb or an adverb phrase. *Meaning* refers to the semantic domains to which those adverbials belong, such as addition or contrast. *Position* has to do with where in the clause the linking adverbial appears, be it in initial, medial (post-subject) or final. This study, however, does not investigate the position of linking adverbials in the clause. The reason for that decision lies in the impossibility of doing such an analysis without necessarily writing an interactive text analysis program that would do the task automatically, as stated by Conrad (1999). Hence, form and meaning will be the features of linking adverbials addressed in the sections that follow. Below, a summary box with the information collected from the aforementioned reference grammar books. In the box, 'yes' means that the information can be found the grammar book, and 'no' means it cannot.

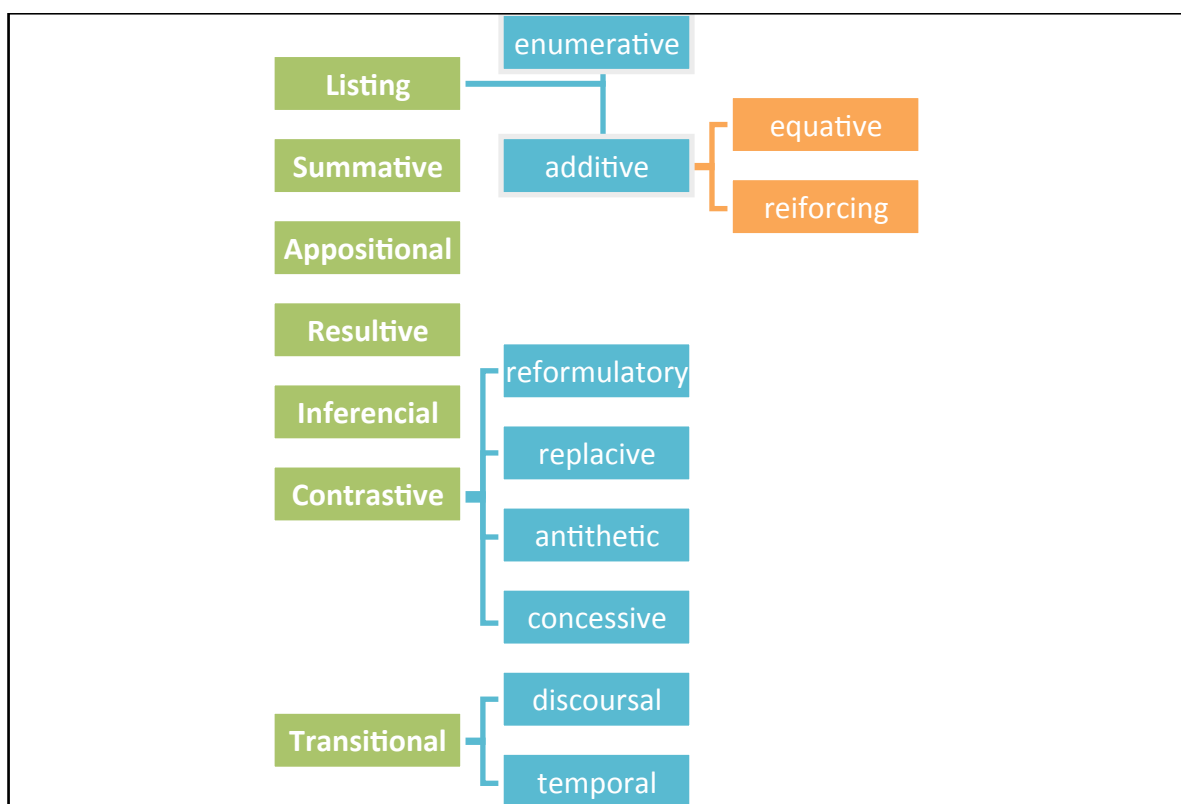
Table 1: Summary of linking adverbials' form, meaning, and position in three grammar books.

	<b>Quirk</b> [conjuncts]	<b>Biber et. al</b> [linking adverbials]	<b>Huddleston &amp; Pullum</b> [connective adjuncts]
<b>Form</b>	No	Yes	No
<b>Meaning</b>	Yes (7 categories + subcategories)	Yes (7 categories)	Yes (5 categories + a subcategory)
<b>Position</b>	No	Yes	No

### 2.2.2.1 Meaning of linking adverbials

The three grammar books use a semantic classification to categorise linking adverbials. Quirk et al. (1985) brings seven categories under which linking adverbials are placed; however, those are further subdivided into ten other subcategories totalling seventeen groups, which makes his classification the most comprehensive of all three books. His classification can be seen from the diagram below.

Box 1 :Quirk's subdivision of semantic categories for linking adverbials



Source: Quirk (1985, p. 634).

Quirk's broad categories have a lot in common with those of Biber et al.'s and Huddleston & Pullum's. He, however, draws very fine lines in his subdivisions. For instance, under *listing*, he considers the *enumerative* class (firstly, next, to conclude, etc.) and the *additive* class, which is then divided into *equative* (likewise, similarly, equally, etc.) and *reinforcing* (also, moreover, in addition, etc.). The four next categories are not subdivided and include the *summative* (therefore, thus, etc.), *appositive* (namely, for example, that is, etc.), *resultive* (therefore, consequently, so, etc.), and *inferencial* (otherwise, then, in other

words, etc). The next category is labelled *contrastive* and is also subdivided into *reformulatory* (rather, more precisely, alternatively, etc.), *replacive* (better, worse, on the other hand, etc.), *antithetic* (contrariwise, conversely, in contrast, etc.), and *concessive* (yet, however, nevertheless, etc.). As for the last category, named *transitional*, the author proposes two subdivisions: *transitional* (incidentally, by the way, etc.) and *temporal* (*meanwhile, in the meantime, etc.*).

It occurs, however, that along with Quirk's vast number of groups and subgroups there comes a range of overlapping possibilities, which, at times, makes it difficult for one to discriminate from. For instance, *on the other hand* can be classified under listing/enumerative, as well as contrastive/replacive and contrastive/antithetic. Similarly, *in other words* may be grouped as appositive, inferential, and contrastive.

Biber et al. (1999), on the other hand, bring a much clearer distinction as regards the semantic categories under which linking adverbials can be found. They do, however, acknowledge that there may be an overlap of categories; not among linking adverbials themselves, but rather an overlap with other adverbial classes, such as stance adverbials and circumstance adverbials. This clear distinction, added to the fact that their grammar book is not only corpus-based but also widely recognized, has made us opt for it as our reference for the data analyses of this study. It appears that Biber et al.'s work was the most successful in bringing together relevant and comprehensive information about linking adverbials, clearly separating them from other classes of adverbials, and condensing information very effectively. The way in which these authors organized linking adverbials has made our study very productive both in terms of time and theoretically. Below is their semantic classification of linking adverbials.

Biber et al. (1999) are not as rigid in establishing their classification. They come up with seven different categories and, rather than creating subclasses, they group together linking adverbials within broad semantic domains, as can be seen in the summary box that follows (table 2).

The first category, *enumeration*, encompasses adverbials used to list pieces of information and organize messages in a text according to the writer's choice.

The order of items being enumerated may follow logical or time sequences, but that is not necessarily always the case. For instance, in giving reasons for developing a certain project, one can place arguments in an order that says *first* reason x, and *second* reason y. But those arguments might be inverted without loss to its logical meaning.

Table 2: Biber et al.'s classification of linking adverbials according to semantic categories.

<b>Enumeration</b>	<ul style="list-style-type: none"> <li>• first, second, firstly, lastly, first of all, to begin with, next, etc.</li> </ul>
<b>Addition</b>	<ul style="list-style-type: none"> <li>• in addition, further, similarly, likewise, moreover, furthermore, etc.</li> </ul>
<b>Summation</b>	<ul style="list-style-type: none"> <li>• in sum, to conclude, in conclusion, to summarize, all in all, etc.</li> </ul>
<b>Apposition</b>	<ul style="list-style-type: none"> <li>• which is to say, in other words, that is, for example, for instance, namely, etc.</li> </ul>
<b>Result/Inference</b>	<ul style="list-style-type: none"> <li>• therefore, consequently, thus, so, as a result, hence, in consequence, etc.</li> </ul>
<b>Contrast/Concession</b>	<ul style="list-style-type: none"> <li>• on the other hand, in contrast, alternatively, however, on the contrary, etc.</li> </ul>
<b>Transition</b>	<ul style="list-style-type: none"> <li>• incidently, by the way, by the by, now, meanwhile.</li> </ul>

Source: Biber et al. (1999, pp. 875 - 879).

As for the *addition* category, adverbials are used to indicate that a new item in discourse is being added to others that have come before. They may also signal that the second item is similar to the first (e.g. *similarly*, *likewise*). Following that come the adverbials grouped under the label of *summation*, which show a piece of message in discourse intends to conclude or sum up the information that comes after it. These adverbials are predominantly found at the end of a paragraph, the concluding paragraph, or in the conclusion section of a research article. Moving on to the next category, we see adverbials of *apposition*.

An appositive linking adverbial shows that the second unit of text is to be treated either as equivalent to or included in the preceding unit, i.e. the second

unit is a restatement of the first, reformulating or stating it in more explicitly. For instance, in the example below, the whole sentence coming after *that is* is a restatement of the quotation that precedes it.

[15] “According to Scholes (2003), ‘no word or expression arises for no reason’, **that is**, they all have some form of motivation and foundation”  
[CRAB, 54]

The two next categories are, in fact, four categories which have been condensed. *Result* was merged with *inference*, as they are semantically very similar. This group of linking adverbials shows that the second unit of discourse brings the result or consequence of the unit that follows it. Similarly, *contrast* and *concession* have also been put together as one group. This category encompasses linking adverbials that signal some kind of contrast or conflict between different messages in a text. When marking a relationship of concession, these adverbials indicate that the next piece of information in discourse expresses something contrary to the expectations raised by the preceding message.

The final semantic category of linking adverbials is the one termed *transition*, which signals a change/transition from one topic in discourse to another. These topics are most often only loosely connected, or even unconnected to one another. This category, however, is rarely found in academic written discourse and, for this reason, it was not included in the analyses of this study.

These categories can also be found in Huddleston & Pullum’s reference grammar book, but these authors have labelled the categories differently. In addition, they distinguish between two kinds of linking adverbials: pure and impure. In their words “pure connectives like *moreover* and *also* have no other function than that of connecting their clause to the surrounding text (or context)” (p. 777). Impure connectives, on the other hand, also have this linking function, but combined with other functions such as that of coordinators. They illustrate it with the adverb *therefore*, which can both *connect* clauses and, at the same time, indicate the *reason* why an event has occurred, as can be seen below.

[16] *His son had been charged with importing illegal drugs; Ed had **therefore** decided to resign from the school board.*

Huddleston & Pullum (2002) consider three categories of impure connectives, to which they give very few examples. These categories are *concession* (nevertheless, nonetheless, still, though, and yet), *condition* (anyway, in that case, otherwise, and then), and *reason/result* (accordingly, as a result, consequently, hence, in consequence, so, therefore, and thus). This type of distinction, however, is not made by either Quirk et al. (1985) or Biber et al. (1999), and it seems that the discussion brought by Huddleston & Pullum (2002) does not affect the true condition of linking adverbials linking spans of text, be them pure or impure. A comparison between Biber et al.'s and Huddleston & Pullum's classification reveals that there is indeed lot in common in their classification, despite the difference in terminology and the absence of the category *summation* in the latter.

Table 3 : Biber et al.'s vs. Huddleston & Pullum's classification of linking adverbials according to semantic categories

Biber et al. (1999)	Huddleston & Pullum (2002)
Enumeration	Ordering (pure)
Addition	Addition and Comparison (pure)
Summation	n/a
Apposition	Elaboration and Exemplification (pure)
Result/ Inference	Reason/Result (impure)
Contrast/Concession	Concession (impure)
	Condition (impure)
Transition	Markers of informational status (pure)

### 2.2.2.1 Form of linking adverbials

As we saw in table 2 above, from the three grammar books consulted, only Biber et al. (1999) present a discussion on the forms of linking adverbials, which can be found in a variety of syntactic realisations, including:

- a. Single adverbs: *anyway, however, therefore, nevertheless*, etc.
- b. Adverb phrases: *even so, first and foremost, more precisely*, etc.
- c. Prepositional phrases: *by the way, for example, in addition, on the other hand*, etc.
- d. Finite clauses: *that is, that is to say, what is more*, etc.
- e. Non-finite clauses: *added to that, to conclude, to sum up*, etc.

According to the authors, every semantic category can be realised by a number of adverbs, but single adverbs are the most frequent, in addition to presenting the most semantically varied structure by which linking adverbials can be realised. Also very common in the academic written register are prepositional phrases. That is because academic prose very frequently presents appositional linking relationships (exemplification and reformulation), and appositive linking adverbials are normally realised through prepositional phrases (e.g. *for example, in other words*). Moreover, enumerative and additive linking adverbials are also very common in academic prose and being commonly realised by prepositional phrases contributes to the prominence of this syntactic form within academic writing.

### **2.2.3 Related corpus-based work on linking adverbials**

Not much work has been dedicated to the study of linking adverbials. Interestingly, a large part of such works were done in China, probably due to Halliday's close connection to that country, as in the beginning of his career he devoted his time to the study of Chinese, and would always draw on that language to make inferences about grammar. In this section we will present studies dating from 1999 onwards, since, as state before, it was in that year that Biber et al. published their grammar book, which gave prominence to the study of linking adverbials.

One of such studies was that published by Conrad (1999). As she is one of the authors of the *Longman Grammar of Spoken and Written English*, her work clearly mirrors what is found in that reference book. Her article, however, is

aimed at introducing language teachers to fundamental characteristics of corpus-based research, and, in order to do that, Conrad uses research on linking adverbials as an illustration. She analyses the registers of conversation, fiction, newspaper reportage, and academic prose so as to establish the differences and similarities in the use of linking adverbials among these registers. The study investigates their frequency, position in the clause, semantic category, and grammatical realizations. Conrad concludes that there are significant differences in the use of linking adverbials depending on the register, being academic prose and conversation the ones which mostly make use of it. In addition, she claims that teachers benefit from corpus-based studies because the answers they give to student questions and the examples they give to illustrate grammar points can more reliably represent real language use.

Liu (2008) also compared the use of linking adverbials in different registers, namely spoken, academic writing, fiction, news, and other writing (nonfiction, miscellaneous - which include writings on commerce, medicine, religion, etc. -, and others). However, as opposed to most studies on the same topic, he did not compile his own corpus, but rather, used the BNC (British National Corpus) as the source for his searches. This choice, he concedes, limits the scope of his conclusions, as there can be some variation in the use of these adverbials among different varieties of English. Still, the study yielded interesting findings, which contradicts some of Biber et al.'s (1999). For instance, whilst Biber et al. state that, in terms of frequency, speaking, fiction, and academic writing have a similar use of contrast/concession adverbs, Liu found that academic writing uses this type of adverbial considerably more. Another contribution of Liu's study is that he suggests a list of linking adverbials divided according to their frequency. For the author, such a list could be used by teachers and materials developers when deciding on the sequence those grammatical devices should be taught. Lastly, Liu acknowledges that there could be some omissions and inaccuracies in his study due to the lack of consensus among grammarians on the classification and terminology of those adverbials.

Shaw (2009) discusses what makes writing mature by comparing the use of linking adverbials in student and professional writing in literary studies. He used four different corpora: one composed of student essays, two others which

comprise academic articles by professional writers, and a last one that encompasses texts of the hard sciences. The reason for the use of two corpora of professional writing is justified by the great variability in the writing of literary disciplines, and the corpus from the hard sciences was included as means of highlighting usages in different disciplinary discourses. As for the results of his study, Shaw found that students use considerably more linking adverbials than professional writers, as they tend to link shorter and simpler ideas. Curiously, he also finds that students make a much lower use of *but* as an adverbial in comparison to professionals, which he accredits to the fact that students tend to follow prescriptive rules of grammar that judge it inappropriate to use coordinating conjunctions as sentence initiators. In addition, Shaw reveals that the use of *yet* is more frequent in both writings in comparison to Biber et al.'s (1999) figures, which suggests that this adverb could be more characteristic of the literary-critical or humanities register.

Finally, Lei Lei (2012) investigated the use of linking adverbials in academic writing on applied linguistics by Chinese doctoral students. He found those students tend to both overuse and underuse linking adverbials in comparison to American professional writers. Interestingly, half the adverbials underused were precisely those expressing contrast. He believes such a tendency in the use of those connectors could be explained by students' lack of register awareness, or inappropriate advice from textbooks, which present lists of connectors as if they were all equivalent substitutes for one another.

None of the above-mentioned studies, however, were done in Brazil. In this country, there seems to be a paucity of research on this topic, especially when it comes to professional writing. Marques (2007), for instance, has done a great work on the use of concessive connectors in adult learners. Nevertheless, her study was not corpus-based, and does not focus on the use of such connectors per se, but rather, on the attention construct with consequence to the improvement of learners' interlanguage.

### 3. METHODOLOGY

What follows is a detailed description of the methods adopted for the development of this research. The procedures are based on the corpus linguistics (CL) methodological approach and, in this respect, our study shares the views of those who, like Gries (2009), see CL as a methodology, rather than a theory. We understand, however, that it would not be reasonable to disregard the fact that CL has made room for the foundation of important principles in linguistics<sup>3</sup>, thus deserving merit as a theoretical field of study (TOGNINI-BONELLI, 2001) cited in (MCENERY & HARDIE, 2012).

The study design falls into the mixed methods paradigm (ANGOURI, 2010) in that it combines both quantitative and qualitative analyses for a better understanding of the phenomena under investigation. We began by compiling two corpora (described below), followed by the application of a statistical test which guided us into discriminating relevant data from the data which would not be of interest. Next, the data were analysed with the help of a concordancer, which allowed for a detailed examination of linking adverbials straight from their original context of use.

#### 3.1 The corpora used

Two corpora were compiled for this study. The first, CRAB (Corpus of Research Articles written by Brazilians), contains 323.909 tokens and consists of applied linguistics research articles published in prestigious journals<sup>4</sup> and written in English by Brazilian professionals. The other, CRAN (Corpus of Research Articles written by Natives) has a total of 363.346 tokens and comprises applied linguistics published research articles written by native speakers of English. By the total number of tokens, both corpora can be said to be middle-sized according to Berber Sardinha's (2004) classification. Yet, in comparison to other renowned large online corpora, such as the BNC (British National Corpus) and COCA (Corpus of American Contemporary English), ours would certainly be

---

<sup>3</sup> Such principles may include those such as colligation, semantic preference, and semantic prosody brought by the so-called Neo-Firthian school of linguistics (MCENERY & HARDIE, 2012)

<sup>4</sup> Listed in Table 4

considered small. Regardless of size, it is known that both large and small corpora allow for interesting findings in linguistic studies. Tribble (1997), cited in Evison (2010) argues that, for very specialized registers, small corpora can better offer insight into the features of that register. His point of view is shared by Koester (2010, p. 66) who, by advocating in favour of small specialised corpora, states that “they allow a much closer link between the corpus and contexts in which the texts in the corpus were produced”.

For they are specialised, the corpora used in this study will be more appropriate for the generalizations we wish to uncover. It is true, however, that the linguistic elements under scrutiny in this research could have been searched for in papers from any other academic area. Nevertheless, our choice for applied linguistics research articles was due to two reasons. First, we wanted to work with texts genuinely written by their authors, i.e. without any sort of mediation involved in the writing process, such as the cooperation of an outside translator. In the applied linguistics field we could search for articles whose authors we knew had most probably written their texts themselves, as they are professors who use English in their workplace. That helped us conform with the construct of research practicality during the data collection phase, as it was unnecessary to contact authors to verify language authenticity. Secondly, it is our goal to better understand academic written language in applied linguistics so as to contribute to the area of which we ourselves are part. Thus far, there is a relatively small number of studies on written language used by applied linguists, especially in comparison to other areas, such as the biological sciences, for which a great deal of research on language use has been done, such as that by Swales (1986, 2014).

### **3.1.1 The corpora compilation process**

Having decided on the kinds of texts that would be comprised in our corpora, it was time to define where those texts would come from. The CAPES<sup>5</sup> web portal seemed to be the best choice as, from it, thousands of journals worldwide can

---

<sup>5</sup> <http://www.periodicos.capes.gov.br/>.

be accessed free of charge. But, being so great the number of possible sources, some criteria should be used as per what journals exactly would be selected. As a starting point, the *CAPES QUALIS*<sup>6</sup> platform, as well as the SLR (Scimago Journal Rank), were used as a reference and only high-ranked journals were accessed. Both Brazilian and international journals were used, and are listed below:

Table 4: Journals used as a source of texts for the compilation of CRAB and CRAN.

Journal	Origin
<i>Bakhtiniana - Revista de Estudos do Discurso</i>	<i>Brazil</i>
<i>Cambridge journals - annual review of applied linguistics</i>	<i>UK</i>
<i>Cambridge journals - language in society</i>	<i>UK</i>
<i>Cambridge journals - language teaching</i>	<i>UK</i>
<i>Canadian journal of applied linguistics</i>	<i>Canada</i>
<i>Delta - Documentação de Estudo em Linguística Teórica e Aplicada</i>	<i>Brazil</i>
<i>Ilha do Desterro - A Journal of English Language, Literatures in English and Cultural Studies</i>	<i>Brazil</i>
<i>Language policy</i>	<i>Netherlands</i>
<i>Linguistics and Education</i>	<i>UK</i>
<i>RBLA - Revista Brasileira de Linguística Aplicada</i>	<i>Brazil</i>

Another key decision concerns the texts which would make up the corpora. For this purpose, we followed Biber's (1993, p.243) concept of representativeness, referred to as "the extent to which a sample includes the full range of variability in a population". Knowing that the field of applied linguistics, though stable and widely recognized, still today faces controversies as for which sub areas are included or excluded from its boundaries (SEIDLHOFER, 2003), the decision concerning which texts would make up our population was based on the definition of areas incorporated in the 2014 AILA Congress (International Association of Applied Linguistics) call for papers. For the 2014 AILA World

<sup>6</sup> *QUALIS* refers to the set of procedures used by CAPES to stratify the quality of the intellectual production in the post-graduate programs. Journals are divided into 8 categories, depending on their quality, and are given a score that ranges from A1 to C.

Congress, applied linguistics falls into five broad categories, namely: language acquisition and language processing; language teaching and learning; language in the professions; language in society; and analysis of spoken and written discourse. For a detailed list of the disciplines these areas encompass, see Appendix A.

Consequently, it would be important to select texts from all of the areas listed above. However, as this study intends to compare two corpora, it would not be appropriate to simply select any text from any category; otherwise, the corpora would not be comparable. This means that, to allow for comparability, the corpora would have to be balanced, i.e., “text categories [should be] sampled proportionally” (McENERY et al, 2006, p.16). Hence, for every subject matter chosen, two texts were collected; one to be part of CRAB, and the other to be included in CRAN. In other words, if, for instance, a discourse analyses research article written by a NNS were selected, a discourse analyses article written by a NS would also be included.

A final step involved converting all text files - originally downloaded in .pdf format - into .txt format. This would enable us to manipulate them in a concordancer (described later in this chapter). In order to do so, the software UniPdf<sup>7</sup> was used. As far as converting files is concerned, this software proved to be a rather time-saving choice as, in the conversion process, all images and unknown characters are automatically removed. It is worth mentioning that all texts were kept in full. This decision was based on the fact that linking adverbials could be found anywhere in the texts and, as pointed by Meyer (2002, p.30) “if the corpus is intended to permit the study of discourse features, then it will have to contain complete texts”.

### **3.1.2 The cleaning-up process**

Even after images and unknown characters have been removed, a corpus is not ready to be manipulated. It is still replete with the so-called text noise, i.e. unwanted information that could both bias and inflate results. There were three

---

<sup>7</sup> Available for download at <http://unipdf.com/>

types of noise that needed to be removed from our corpora. (1) Instances of oral language (i.e. transcribed dialogues) which generally appeared in texts on analyses of oral discourse or classroom research. These features were removed from both corpora. (2) Words and characters that would inflate the actual total number of tokens in the corpora, which, in turn, would lead to a miscalculation of the normalized frequencies. Examples include tables, numbers, and highly repeated words such as author's name or name of the journal (which generally appears on every other page of an article). These features were also removed from both corpora. (3) Quotations that appeared in CRAB. Quotations were deleted because they are perfect copies of authentic text and, therefore, do not depict authentic L2 English usage.

Those steps were carried both automatically and manually, as shown below:

Table 5: Steps involved in the clean-up process of CRAB and CRAN.

Step	Corpus	Mode	Software
1	CRAB and CRAN	manual	-
2	CRAB and CRAN	semi-automatic	Notepad++
3	CRAB	automatic	R

The first phase was done manually and texts files were analysed separately in search for any occurrence of transcribed dialogues. The second phase was done semi-automatically with the help of software Notepad++. One great advantage this program offers is the possibility of working with all text files at the same time. This means that by using the 'search and substitute' command, a single word or term could be deleted from all files at once.

Finally, the last cleaning-up phase was done automatically in the R environment. In order to be able to better deal with texts using R it is necessary that, first of all, the TM (text mining) package be installed. Its installation can be done rather straightforwardly through the command `/ install.packages("tm") /`.

One of the advantages of using this package is that one can deal with all texts of a corpus at once. After having done that, the following script, adapted from Grondona (2014, in press), was used to delete all instances of quotations present in CRAB:

```
texto <- gsub(pattern="<[^\<>]+>",replacement="",x=texto)
return(texto)
}
```

To put it simply, the action performed by this script is the deletion of any string of text which falls in between quotations marks, i.e. it deletes what we know as direct quotations. After that, the corpus was ready to be explored.

### 3.2 Procedures for data analyses

A list containing the linking adverbials to be investigated was obtained from different grammar books: Quirk et al. (1985), Biber et al. (1999), Huddleston & Pullum (2002), and Carter & McCarthy (2006), in addition to Shaw's (2009) study.

The first part of the analyses requires that linking adverbials be searched for in a concordancer. The software used for this purpose was AntConc 3.4.2<sup>8</sup>, developed by Professor Lawrence Anthony. AntConc has been successfully used in other corpus-based studies, such as Shaw (2009) as well as Römer & Wulff (2010), and various researchers advocate in favour of its use, among whom, Mello & Souza (2012) and Gries (2009). Each linking adverbial was, then, searched in this concordancer and their raw frequencies were recorded. It is important to mention that the variation in spelling among different varieties of English words had to be taken into account. For instance, when trying to find the prepositional phrase *to summarise*, its American correspondence *to summarize* was added up to the final count.

---

<sup>8</sup> Available for download at <http://www.antlab.sci.waseda.ac.jp/software.html>

After having found the number of occurrences for each linking adverbial in both corpora (CRAB and CRAN), their frequencies were normalised. Normalising frequencies is a fundamental step for working with corpora of different sizes, as otherwise results could be entirely biased. All frequency counts were, then, normalised to counts per 100.000 words. Despite the fact that Biber, Conrad & Reppen (1998, p.264) state that “frequency counts should be normed to the typical text length in a corpus”, which means we should opt for the 10 thousand/word normalization, normalizing our results to counts per 10.000 words would generate figures that could hinder any prompt visualization of the data, due to the very small numbers that would be acquired. In view of that, the choice for the 100.000 normalization seemed more appropriate for analyses.

Up to this point, the exact linking adverbials to be scrutinized had not been defined. Two different criteria were used for the inclusion of the linking adverbials in this study. First, to be included in the analysis, they would have to be statistically significant. Not being statistically significant means that, as regards frequency, the differences in the use of those linking adverbials between NNS and NS is virtually non-existent, thus suggesting great balance. One of the most widely used statistical tests used by corpus linguist in the chi-square test but, although it can be reliably used to test the significance of values between two datasets, it has its limitations. One of such limitations is the fact that its reliability is reduced if we are dealing with very small numbers (MEYER, 2002). In view of that, the Log-likelihood (LL) test was adopted. This parametric test seems adequate for our study since: i) it can be used with smaller volumes of text; ii) it does not presuppose the data is normally distributed, and iii) it allows the comparison of the significance between both rare and common phenomenon (DUNNING, 1993). The statistical cut-off point used here was  $LL > 3.84$  at the  $p < 0.05$  level, as defined by McEnery and Hardie (2012).

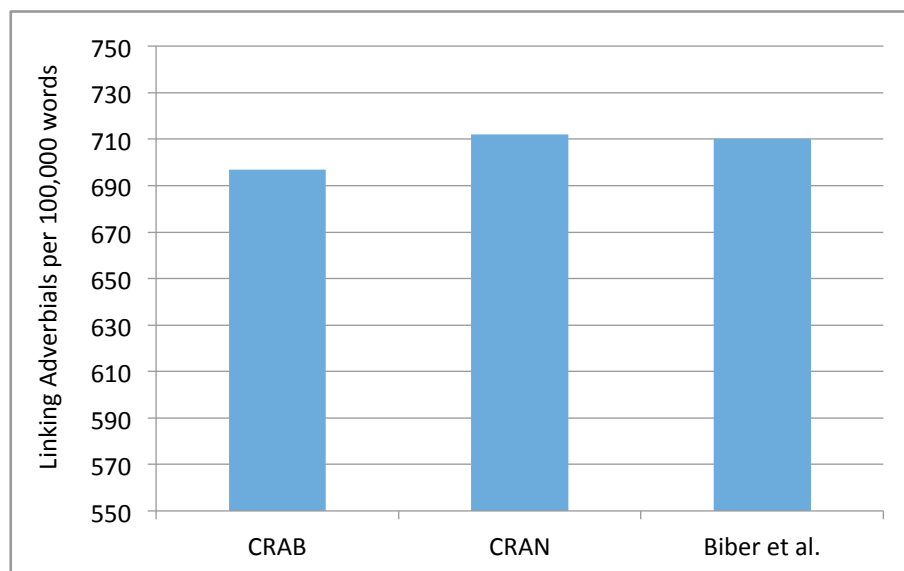
The second criterion used for the inclusion of the linking adverbials in this study was their overall frequency. As the number of individual linking adverbials is very large, it would not be possible to subject all of them to analyses due to time constraints. This step was, therefore, taken as another means of reducing the total number of elements to be examined. Two studies were used as a reference for the determination of a relatively acceptable cut-off point: Biber et

al. (1999), and Biber, Conrad & Cortes (2004). In determining the cut-off point for the inclusion of lexical bundles to be examined in their investigation, the authors set a conservative approach of 40 times per million words in the first study and 10 times per million words in the second. For our study we found it would be best not to be so conservative while, at the same time, not lack rigour. Hence, the cut-off point established here is an average value of the above-mentioned studies, i.e. 2,5 occurrences per 100.000 words (or 25 occurrences per million).

## 4. DATA ANALYSIS

This section addresses two topics regarding the use of linking adverbials: their meaning and form. For such, we begin by presenting the overall data from our corpora and compare them to those of Biber et al.'s (1999), our reference study. Then, we introduce the discussion on the meaning of linking adverbials. More specifically, we deal with the various semantic categories to which linking adverbials can be assigned and compare the frequency and use of these adverbials by professional Brazilian academic writers (henceforth NNS) and English L1 professional academic writers (henceforth NS). It is worth reminding the reader that the writers we refer to belong specifically to the field of applied linguistics and by no means have we intended to generalize to other fields of study. Whenever that is the case, we will make it explicit. Finally, we present a brief discussion of the various syntactic realisations of linking adverbials found in our two corpora.

### 4.1 Data Overview



Graph 1: Overall distribution of linking adverbials in CRAB, CRAN, and Biber et al. (1999).

The overall distribution of linking adverbials in our corpora conforms to the results presented in Biber et al. (1999). For the academic written register<sup>9</sup>, the authors found an average of 710 linking adverbials per one hundred thousand words, whereas CRAB and CRAN have 697 and 712, respectively, as can be seen from Graph 1.

Notably, NNS make a slightly lower use of linking adverbials in their texts and, at first sight, this could indicate that Brazilians and NS use linking adverbials quite similarly. However, a closer look at the data will reveal differences as per exactly which of those adverbials are either more or less consistently used by NNS in comparison to their NS counterparts.

The graph below (graph 2) shows that, concerning the aforementioned semantic categories, the differences between CRAB and CRAN are not strikingly marked. It is notable, however, that when it comes to conveying messages that express the idea of result/inference, NNS overuse linking adverbials in comparison to NS (54% more). In contrast, NNS underuse the category of apposition (40% less). Interestingly, when taking Biber et al.'s data into consideration, a whole new picture arises, showing what seems to be a tendency in NNS' texts.

By looking at graph 2 what seems to be that previously mentioned tendency is the fact that NNS's use of linking adverbials appears to parallel that which is found in Biber et al.'s study. Putting it more clearly, where NNS use of linking adverbials is not so high, Biber et al.'s corpus also shows a not so high use of these adverbials. And the same is true where NNS's use of linking adverbials is more noticeable.

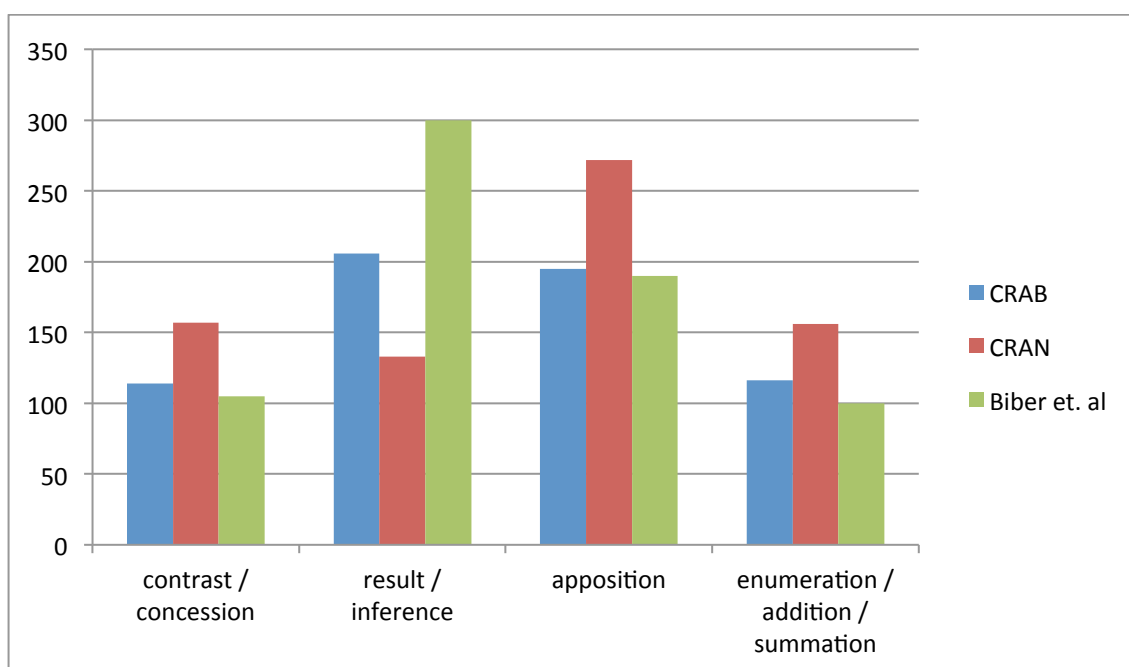
It is important to restate that the corpus used in Biber et al.'s study greatly differs from CRAB in that the first encompasses texts taken from different registers (research articles and book extracts) as well as from a variety of disciplines<sup>10</sup>, whilst the latter is highly specialized, including only research

---

<sup>9</sup> In Biber et al.(1999), the written academic register includes both research articles and book extracts, all of which taken from a wide range of different academic disciplines.

<sup>10</sup> These disciplines include: agriculture, anthropology/archaeology, biology/ecology, chemistry/physics, computing, education, engineering/technology, geology/geography, law/history/politics, mathematics, linguistics/literature, medicine, nursing, psychology, sociology.

articles from the applied linguistics field. Having said that, the fact that there is some similarity in the use of linking adverbials between those two corpora is quite remarkable. It is well known that different academic disciplines present different lexico-grammatical, rhetorical, and argumentative textual patterns, as stated by Hyland (2009). He also argues that “we can see disciplines as language using communities and the term helps us join writers, texts and readers together” (p. 179). Thus, it would be fair to expect that CRAB and CRAN, both corpora with texts from the same discipline, could contain a similar use as regards linking adverbials.

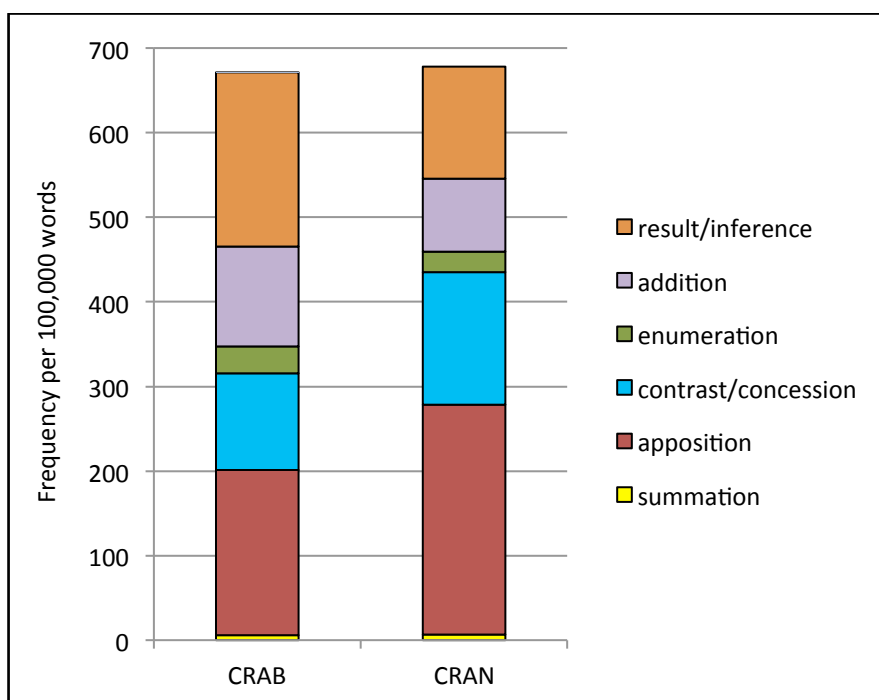


Graph 2: Overall distribution of linking adverbials by semantic category in three corpora: CRAB, CRAN, and Biber et al (1999).

Nonetheless, the fact that it was CRAB and Biber et al.'s corpora that portrayed the greatest similarity might be an indicator that, as far as linking adverbials are concerned, Brazilian and NS applied linguistics do not follow the same lines in writing. It also appears that, regarding these adverbials, Brazilians follow a more general style, which compares not to the specific field of applied linguistics, but rather, to an average use of all disciplines. This might be explained by the fact that NNS could perhaps not abandon the linguistic behaviour acquired when

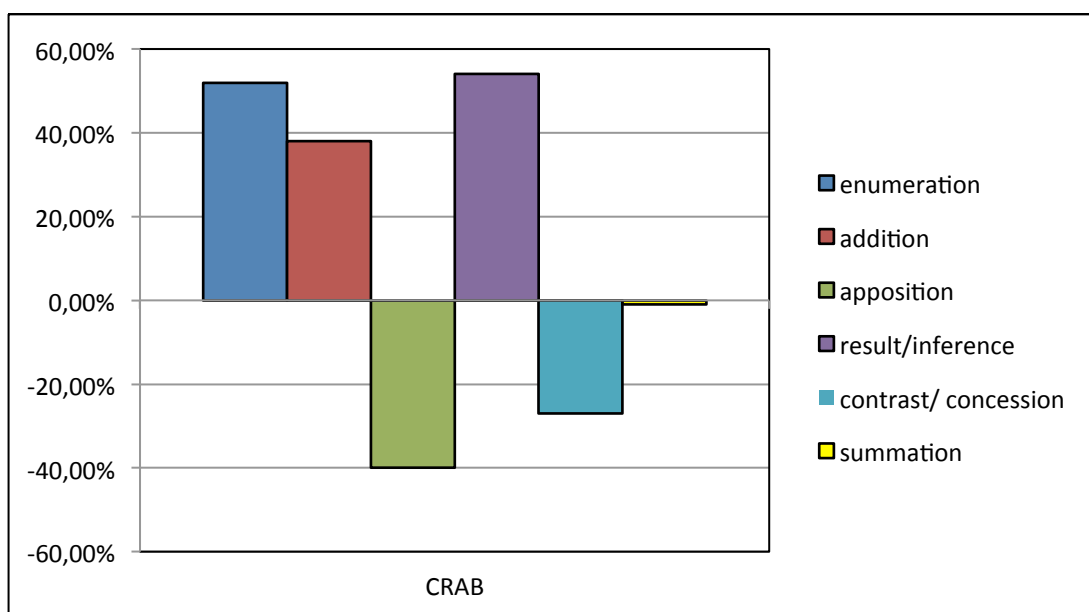
they learned to write in English as an L2, i.e. some of the characteristics found in their writing could be attributed to the fact that when one learns a second language, the input for the acquisition of that language comes not from a specific subject, but from various subjects. Another reason could be the corpus size. The academic prose section alone in Biber et al.'s corpus (5,331.800 tokens) is much larger than ours, and that causes discrepancies in the results, even when using normalized data.

Going further, we move on to group linking adverbials according to each semantic category separately, namely enumeration, addition, apposition, result/inference, contrast/ concession, and summation. Having done that, what can be seen is that there is some variation between NNS's and NS's use of linking adverbials. Remarkably, the categories of apposition, followed by result/inference, are those in which the largest number of linking adverbials concentrates. Below is a graph that illustrates this assertion.



Graph 3: Distribution of linking adverbials among different semantic domains in CRAB and CRAN.

By looking at the quantitative data in terms of broad semantic categories, the category of summation, which includes adverbials such as *in short* and *to conclude*, is the least used, and the one that presents the least variation between NNS and NS, 1.7% only. That is because such adverbials have a very limited scope of use, since, in most cases, they appear almost exclusively in the conclusion section of research articles. On the other hand, linking adverbials that make up the remaining categories may be found anywhere in academic texts, from introduction through conclusion, which, in turn, allows for greater variability. Nevertheless, it is important to point out that, as per the range of variation found between CRAB and CRAN, Brazilians use more linking adverbials than NS concerning some categories, but the reverse is also true. The range of variation is as follows.



Graph 4: Overuse and underuse of linking adverbials in CRAB in comparison to CRAN.

As can be seen from the graph, the variation in the use of linking adverbials in the two corpora is considerable, though not extremely high. The categories in which NNS overuse such items are enumeration (52%), addition (38%), and result/inference (54%). Those which are underused include apposition (-40%), contrast/concession (-27%), and summation (-1%). Nevertheless, more

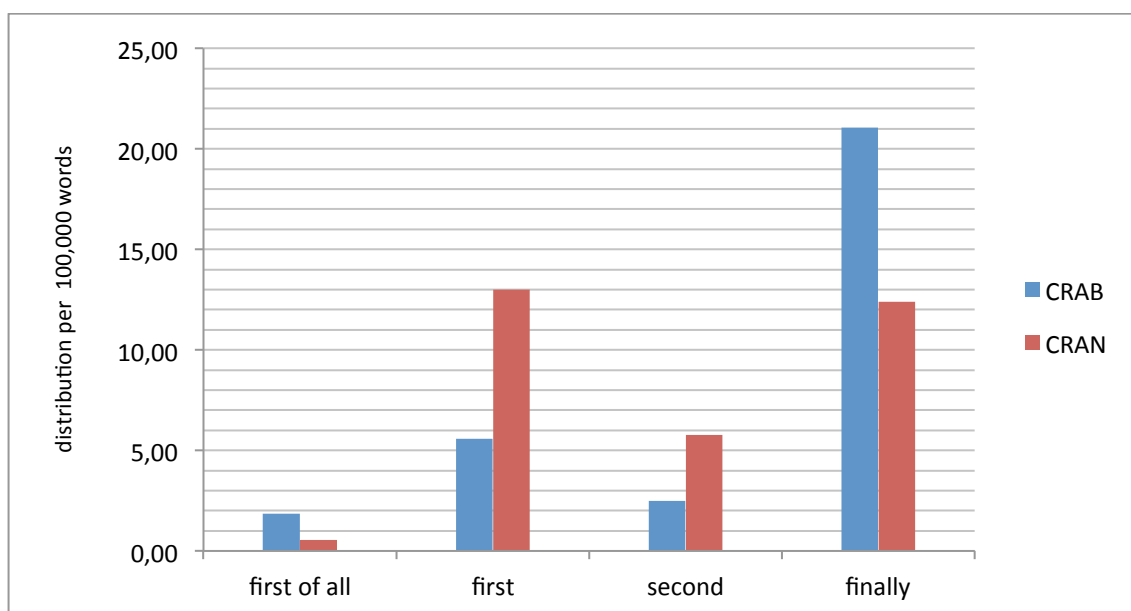
interesting differences are revealed when we look at each semantic category separately from a qualitative point of view. And that is what is done in the section that follows.

## **4.2 Distribution according to separate semantic categories**

In this section we will introduce our discussion on linking adverbials in relation to each semantic category. It is worth mentioning that all graphs in this section display the linking adverbials which have proved to be statistically significant according to the log-likelihood test. The remaining linking adverbials, along with the statistically significant ones, can be found at the Appendix C for reader's reference.

### **4.2.1 Enumeration**

We begin by analysing the enumeration category. The linking adverbials left out for not being statistically significant were: firstly, to start with, to begin with, lastly, secondly, third, and thirdly. Hence, four linking adverbials are shown in the graph, namely *first of all*, *first*, *second*, and *finally*. The greatest range of variation is found in *first*, with NS using it 132% more times, and *finally*, used 70% less times.



Graph 5: Distribution of linking adverbials belonging to the sematic category of enumeration in CRAB and CRAN

As regards the use of *first*, which is much less used by NNS, it appears that when signalling the start of a series of events or arguments, Brazilians make use of a wider range of forms, which include *to begin with*, *to start with*, *firstly*, and *first of all*. This seems to contribute to a lower use of the adverb *first* in their texts. NS, on the other hand, predominantly make use of *first* to convey the same type of message, which can be attested by the graph above.

Another interesting fact comes when we analyse the use of *finally*, which is overused by NNS. It seems that Brazilians feel the need to mark the end of a series of points, even when other textual resources could perform the same function, as can be seen in the example below.

[17] ***Finally***, *the third and final part* of the paper is oriented towards the analysis of future EFL teachers' report excerpts so as to draw conclusions about the processes of knowledge construction involved in the activity of becoming a teacher." [CRAB, 30]

In this example, what we see is that *finally* could have been left out, as the phrase *the third and final part* already signals that the series of events has come to an end. Another point to consider is that, in NNS's writing, the use of

*finally* is most often preceded in discourse by the adverbs *first/first of all*, and *second/secondly*, thus forming the sequence first - secondly - finally. In contrast, NS seem make a larger use of possibilities, including the very frequent sequence: first - second - third. In addition, NS do not seem to be so explicit when marking the end of a series of events. They, at times, finish their argument with the insertion of *second* or *secondly*, or simply do not mark it, or even use some additive adverb, thus making a much lower use of the form *finally*. Below is an example.

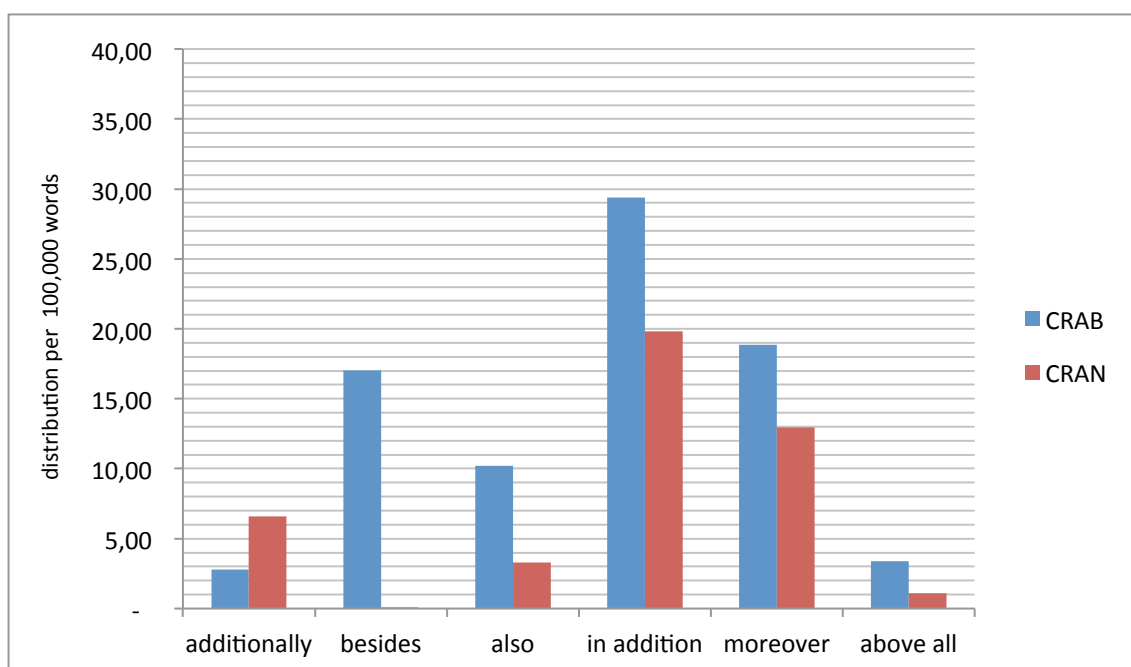
[16] “**First**, I will address the diverse historical inheritances and lines of development associated with critical pedagogies of second or other languages. I then spend time on advocacy. I draw attention to the diverse institutional contexts that could be explored for critical language pedagogy. I point to the broader than appreciated range of languages within which critical approaches have been explored, **then** discuss aspects of EFL (English as a foreign language) critical pedagogy, as this is one of the more challenging, but basically positive, developments.” [CRAN, 39a]

[17] “**First**, I will briefly discuss the main features of the ESP approach and then I will focus on reading as an active process. The notions about concept maps and the characteristics of the CMap Tools software will also be addressed. The design of the research project will then be discussed, followed by a description of the context, including the subjects who participated in the study. **Finally**, I will discuss the results obtained.” [CRAB, 27]

What we see in the first example is that the NS text does not explicitly signal the end of a series of events, but rather, shows that there is one more point to be discussed by using the adverb *then*. The text written by a NNS, on the other hand, does mark the end of events, explicitly using *finally*, as previously discussed.

#### 4.2.2. Addition

A number of linking adverbials within the category of addition were not statistically significant and thus left out of the analyses: *further*, *likewise*, *in any case*, *in effect*, *and* (sentence initial), and *again/ once again*. Hence, the ones represented in the graph include *additionally*, *besides*, *also*, *in addition*, *moreover*, and *above all*. It is interesting to observe that, except for *additionally*, all of the analysed linking adverbials are overused by NNS to a greater or lesser extent.



Graph 6: Distribution of linking adverbials belonging to the semantic category of addition in CRAB and CRAN

From the data shown in the graph, two linking adverbials particularly draw our attention: *also* and *besides*. They are of interest here for the same reason: they are used mainly in informal contexts, as we will see below. Since we are dealing with texts belonging to the academic prose register, i.e. scientific articles, such adverbials were not expected to be found so prominently in our corpora. Indeed, if we look at *besides* we see that this linking adverbial is almost exclusively used by Brazilians. In fact, NS do use it, but, as its frequency is extremely low (only 1.1 times per 100,000 words), it does not appear in the graph above.

Notably, when NS make use of *besides* in their texts, it comes as a preposition in virtually all of its uses, not as a sentence adverb, as can be seen below.

[18] “**Besides** suggesting that social networking sites can serve as an authentic way to be exposed to language and culture outside of the classroom, it is also argued that . . .” [CRAN, 14a]

[19] “Maybe not enough is known about the interpersonal and representational processes involved in distance learning in online contexts. **Besides**, the fact that isolation is still associated with distance learning means that not all distance learning educators are tuned up with the role of digital communication...” [CRAB, 08]

The next linking adverbial of interest here is, as mentioned above, the adverb *also*. In comparison to NS, Brazilians use it 209% more times. This fact alone already shows a great discrepancy in its use between the two corpora. But, by looking at the data more closely, an interesting point is revealed. In NS’s texts, *also* is used to connect simple or short spans of text. It is found, for instance, in a table that presents a summary of findings which is intended to compress information for fast reading. It is also predominantly found in the introduction and results sections of articles, in which authors seem to adopt a less rigid text format. In addition, where *also* is present, other more dialogical textual features can be found, such as the use of first person singular.

Conversely, NNS seem to treat *also* as any other linking adverbial used for conveying the message of addition. They, for instance, use it in essentially any section of the article, and have it connect complex messages or longer spans of text, as can be seen from the examples below.

[20] “**Also**, decide whether you think your use of English was necessary, and in any case, how much you think English helped you complete the task in German, or not.” [CRAN, 35a]

[21] “**Also**, if knowledge is understood as a social construct, learning as a social process and teaching is not detached from the environment, then it is possible to admit that the educational context in which we live nowadays enables the development of different actions in which

*computers, and their different uses, may represent much more than mere tools to be used for gathering data or for communicating.* [CRAB, 28]

Hence, as regards the use of linking adverbials that have the additive value, it is possible to claim that NNS seem to be less conservative in their writing, since they make use of those less formal adverbs in their texts. However, it is also true that NNS overuse virtually all linking adverbials that belong to this semantic domain. This could be explained by the fact that NNS add messages to their arguments more explicitly than NS do. It seems to be the case that NS have more ways of adding information, such as through the use of stance adverbials, which can take the place of the additive linking adverbials. Below are two examples.

[22] *“The formal, or grammatical, syllabus is arguably the easiest way to package language as a subject-language for a variety of reasons. It is, at least in traditional grammars, a self-contained system governed by a set of rules. It can thus be packaged, taught and tested. **Importantly**, teachers can master it as a subject matter and thereby fulfil their roles as experts in the classroom.”* [CRAN, 40a]

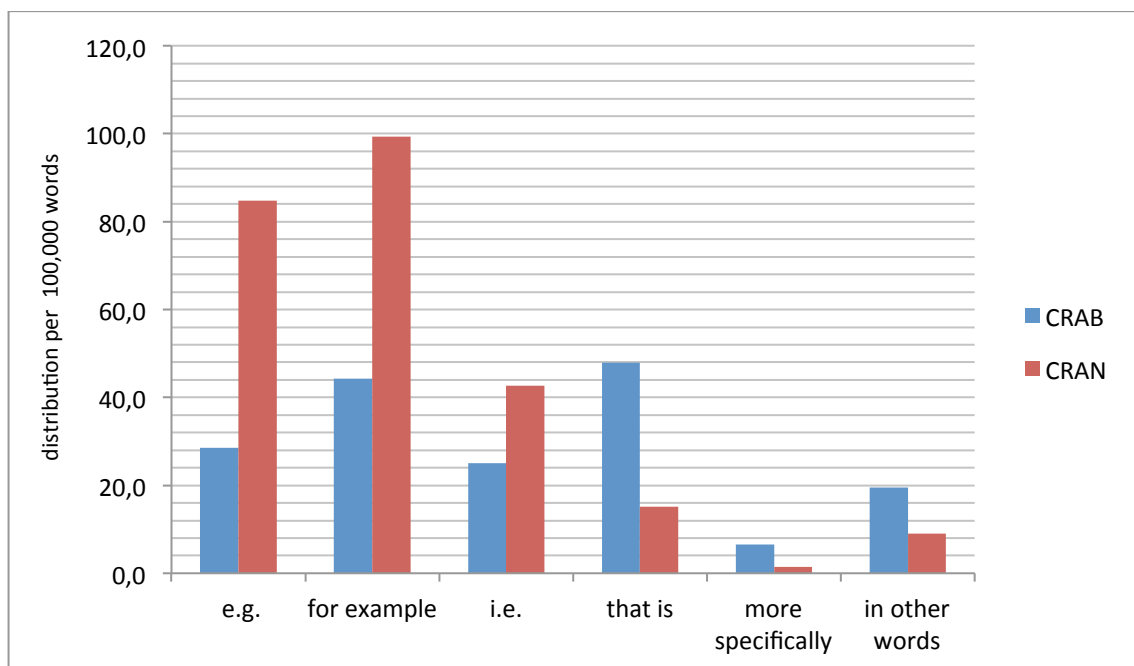
[23] *“Although these resources are generally framed within a communicative approach which emphasises interaction, collaboration and meaning, they often involve a more traditional focus on form, accuracy and individual study. **Interestingly**, they differ in the extent to which they accommodate monolingual teachers...”* [CRAN, 31a]

In the first example, it seems that the adverb *importantly* is not only drawing readers' attention to the fact that the argument just posed is an important one; it is also adding an argument in favour of the so-called formal syllabus. Thus, the excerpt *“Importantly, teachers can master it as...”* could be rephrased as *“In addition, teachers can master it as...”* The same seems to be true for the second example. The author wants not only to emphasise the information given about the pedagogical resources in question, but also add information to what has been stated. In this sense, *interestingly* could be replaced by, for instance,

*moreover*. The use of these stance adverbials seem to be almost not used in NNS' texts that make up our corpora. A parallel search in CRAB and CRAN revealed that *importantly* is used almost four times more by NS, and *interestingly* almost three times more.

### 4.2.3 Apposition

Three linking adverbials within the category of apposition were not statistically significant and thus left out of the analyses: *another example*, *namely*, and *for instance*. Consequently, the following adverbs were used for the analysis: *more specifically*, *in other words*, *for example*, *that is*, *e.g.*, and *i.e.*



Graph 7: Distribution of linking adverbials belonging to the semantic category of apposition in CRAB and CRAN.

A first glance at the graph above shows that the most common form within the apposition category is realised by the phrase *for example*. Remarkably, if we add the totals of the two forms (*for example* and *e.g.*), we see that NS use them 153% more times than NNS do. The data seems to suggest that NS might be

more careful when explaining a certain point by overtly signalling that what is being said refers to an illustration of the topic, not simply an ordinary explanation. Thus, their texts could be said to sound more pedagogical to a certain extent, in the sense that they connect to the reader as though in a classroom communication. A search on the BNC<sup>11</sup> shows that such an assertion seems reasonable since, within the spoken register, the use of *for example* is particularly common in university lectures, while being very infrequent in more monological registers such as sermons or speeches.

In applied linguistics scientific articles, therefore, NS appear to make their exemplifications more explicit. Other searches were done looking for different forms of exemplifying that are not realised through linking adverbials. Such forms included *illustrate*, *illustrates*, *illustrated*, *an example*, *examples*, *exemplify*, *exemplifies*, *exemplified*, *exemplification*, and *such as*. Again, native speakers' use of such appositive forms was higher, thus corroborating this distinguishing characteristic of NS's writing.

The linking adverbial *for instance*, as mentioned at the beginning of this section, was not found to be statistically significant. Nonetheless, an interesting contrast of this adverbial with *for example* has been established in previous research by Lee & Swales (2006). The authors have found that "*for instance* is used a lot more frequently in the social sciences and humanities, [...] whereas in the natural sciences *for example* is clearly favoured (being used to illuminate and clarify a difficult or complex point through the exemplification)" (p.64).

Noticeably, that appears to be the case in applied linguistics. In this field of study, most NS prefer to use *for instance* so as to exemplify more "casual, non-essential information" (LEE & SWALES, 2006), while some NNS use the same linking adverbial to explain complex ideas, where, ideally as it seems, the use of *for example* would have been better appreciated. Below are two examples that illustrate this point.

[24] "*The teachers asked, **for instance**, how do genres and modes vary across disciplines, subjects and fields?*" [CRAN, 8a]

---

<sup>11</sup> The British National Corpus, available at <http://corpus.byu.edu/bnc/>

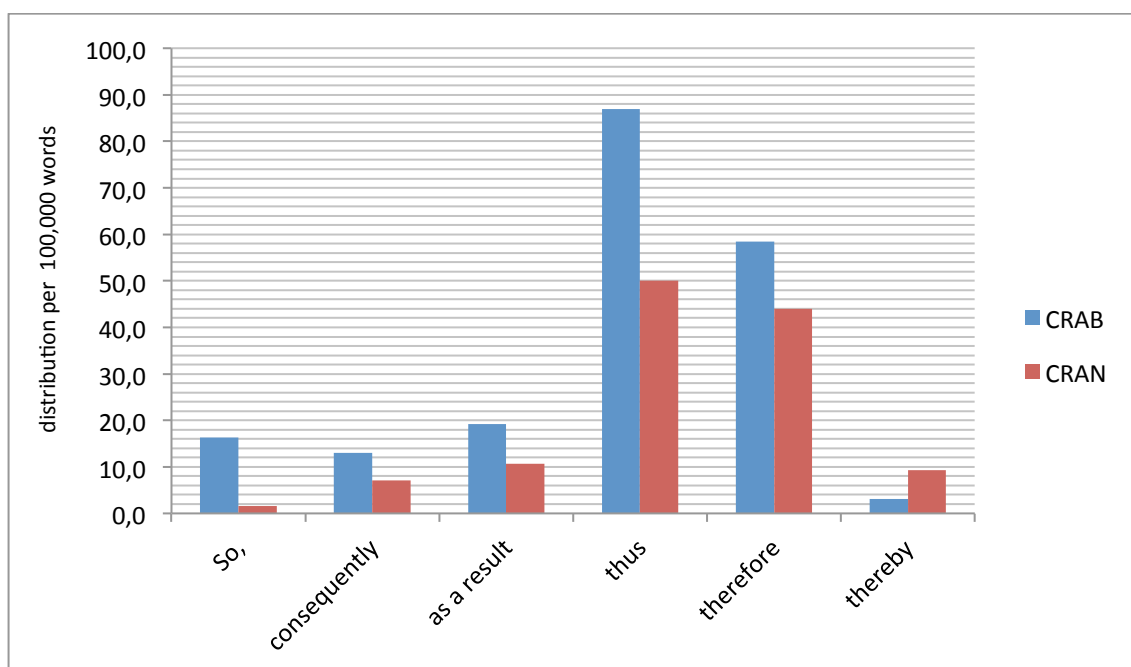
[25] “*Inherent in this discussion is a need to make explicit how language, culture, identity, relations, ideology, and power are relevant to all language teaching and learning situations. For instance, as Kubota (2004, 2006) has argued, the study of culture should resist a color-blind, superficial liberal view of cultural difference that essentializes cultures and creates a dichotomy between. . .*” [CRAB, 40]

As for the form *that is*, what can be seen is that it is preferred among NNS. However, when the abbreviated form *i.e.* is taken into consideration, we notice that the two are somewhat balanced, for NNS prefer the latter. Both *more specifically* and *in other words* are more used by NNS; yet the two are not very frequent in both corpora.

#### 4.2.4 Result / Inference

In the result/inference category, three linking adverbials were left out of the analysis for not being statistically significant: *hence*, *for that reason*, and *alternatively*. Thus, the ones analysed include: *so*, *consequently*, *as a result*, *thus*, *therefore*, and *thereby*.

The most frequent linking adverbial within result/inference group is *thus*, which is the most used in both corpora. Brazilians, however, seem to prefer this adverbial in their writing, displaying an overuse of 74% in comparison to their NS counterparts. These NS, in turn, also have *thus* as the most productive adverbial among those in this category, but they make great use of *therefore*, which nearly equals the total number of *thus*. Despite the fact that *thus* and *therefore* are very frequent linking adverbials in both corpora, NS' use of those two forms is more balanced than that of NNS, that is, the choice between one form or the other is much less discrepant for NS, which seems to grant these linking adverbials a more equivalent value in discourse.



Graph 8: Distribution of linking adverbials belonging to the semantic category of result/inference in CRAB and CRAN.

Curiously, however, as regards the use of *thus*, what seems to be a rather striking difference between the two corpora is not only its frequency, but rather the syntactic forms that collocate with it. The combination *thus* + *present participle* is three times more common in NNS' writing than it is in NS's. Below is an example of such an instance.

- [26] “*Collocation lists were proposed as an instrument for viewing words and their co-texts, **thus enabling** researchers to sample words that seem to be used metaphorically in light of linguistic evidence...*”  
[CRAB, 5a]

A common resource used by NS - and also by NNS, but to a lesser extent - so as to avoid that type of syntactic combination is to use a relative clause in lieu of *thus* + participle, as can be seen below.

- [27] “. . . *my appeal is for a better appreciation of the complexities of metaphorical meaning and understanding, and the relations between meaning and understanding, **which will enable** us to see the true beauty of metaphor in language, thought, and culture.*” [CRAN, 7a]

A last point in the discussion on this adverbial regards the punctuation that follows it. Where *thus* is found as a sentence initiator, it is most commonly followed by a comma in NNS' texts (96% of Brazilians use a comma after it). In contrast, NS seem to be less rigid as per the use of a comma following *thus* (65%).

Interestingly, by looking at NS' texts one can establish a fair parallel between *thus* and *thereby*, which seem quite synonymous with each other. In CRAN, concordance lines containing *thereby* reveal similar syntactic combinations to those of *thus*. Both linking adverbials accept the collocation with the present participle, in addition to the coordinator *and* as a common collocater, which generates the phrases *and thus* or *and thereby*. Conversely, NNS use of *thereby* does not always parallel that of *thus*. At times, Brazilians seem to misuse that adverbial by choosing it as a synonym to *therefore* (25% of the times). For instance, they place *thereby* between commas, which is not at all a common use for that linking adverbial, as shown in the next example.

[28] “ . . . *considering that knowledge, skills and attitudes modification takes place through performing tasks of content production, promoting, **thereby**, the teaching and learning process .*” [CRAB, 35]

Indeed, *therefore* is commonly found between commas, not *thereby*. Such an assertion can be attested through simple searches in large online corpora such as the BNC and COCA<sup>12</sup>, as shown below, where it can be seen that the use of *thereby* between commas is rather uncommon in academic prose.

Table 6: Total count per million words of *therefore* and *thereby* in the BNC and COCA.

	BNC	COCA
therefore (total tokens)	551.02	275.15
therefore (between commas)	126.80	52.31
thereby (total tokens)	59.75	60.00

<sup>12</sup> Corpus of Contemporary American English, available free of charge at <http://corpus.byu.edu/coca>

thereby (between commas)	0.85	0.96
--------------------------	------	------

Finally, a linking adverbial that shows great discrepancy in terms of frequency of use between NNS and NS is *so*. This adverbial displays an 800% overuse by NNS; it seems to be the case that NNS writing does portray instances of oral language. It is true, however, that NS also make use of *so* in their texts (though rarely), but when this connector is used by NS, it comes in sections of the article that allow for such informality, such as in instances in which the author chooses a more dialogical tone and uses, for instance, the first person pronoun *I*. Many NNS, contrariwise, make use of *so* in parts of their texts where no other feature of oral language is found. In such cases other more formal linking adverbials - e.g. *therefore*, *consequently*, or *hence* - could have been preferred. Below are two examples that illustrate these points.

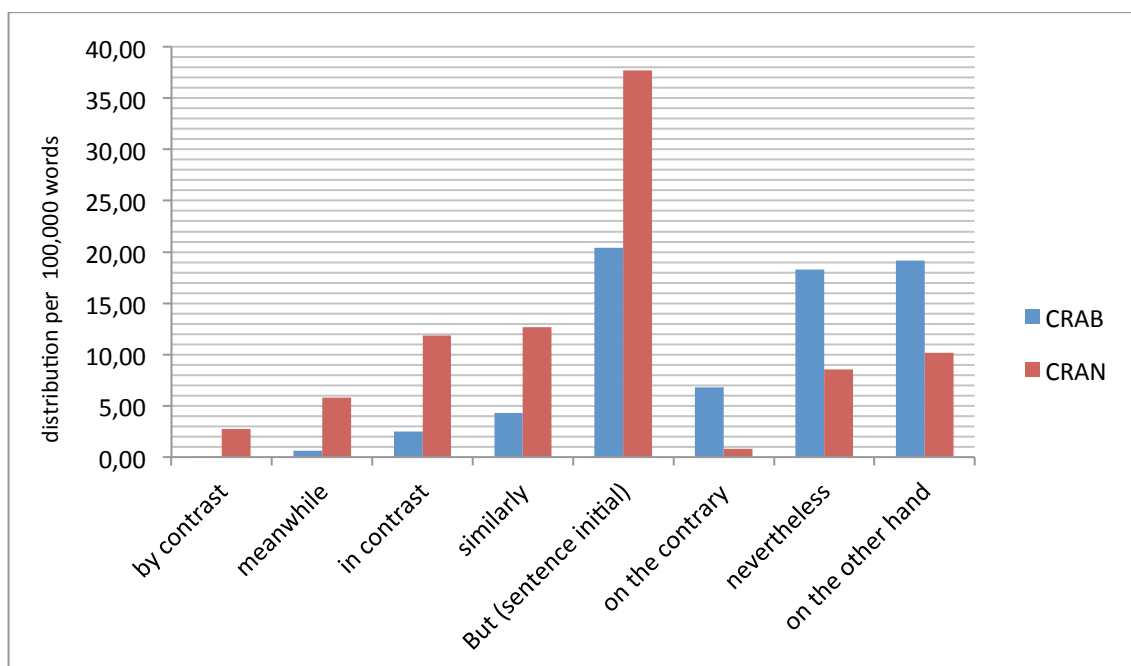
[29] “*So, in the course of this essay, I want to use my reflections on Phillip Goode in order to explore how people in San Carlos address the questions “What is a language? and “What is a language expert? I want to consider the ways in which evangelical missionary practices have hindered language revitalization. . . .”* [CRAN, 46a]

[30] “*However, the subject-producer can use paraphrase for the argumentative exercise, in order to counter-argue the paraphrased passage. So, this dialogical property refers to consensus and polemics indifferently. Fundamentally, rhetoric exigencies order the paraphrastic enunciate status. . . .”* [CRAB, 38]

#### 4.2.4 Contrast / Concession

In the contrast/concession category, ten linking adverbials were left out of the analysis for not being statistically significant: *alternatively*, *in turn*, *notwithstanding*, *otherwise*, *or* (sentence initial), *yet*, *conversely*, *at the same time*, *at any rate*, and *nonetheless*.

From the reading of the graph, one clearly sees that *but* is the most used form to establish a relation of contrast amongst messages in our corpora. We know, however, that *but* is not a linking adverbial, but rather a coordinating conjunction and is presented here for comparability, as when it is found in initial position in a sentence, it may perform a very similar function as that of linking adverbials.



Graph 9: Distribution of linking adverbials belonging to the semantic category of contrast/concession in CRAB and CRAN.

The fact that *but* is much more frequent in comparison to all other linking adverbials used to perform the same semantic relation in our corpora comes as a great surprise. Throughout our school years, many have been told that coordinating conjunctions such as *but* should be avoided as sentence initiators. Nevertheless, what corpus findings show us is quite the contrary. Proficient writers in the field of applied linguistics (our corpora) display a great preference for *but* exactly in sentence initial position. This comes to corroborate what the Oxford Fowler's Style Guide has long maintained. In the author's words:

"The widespread public belief that *but* [emphasis added] should not be used at the beginning of a sentence seems to be unshakeable. Yet it has no foundation. In certain kinds of compound sentences, *but* is used to introduce a balancing statement 'of the nature of an exception, objection, limitation, or contrast to what has gone before; sometimes, in its weakest form, merely expressing disconnection, or emphasizing the introduction of a distinct or independent fact' (OED). In such

circumstances, but is most commonly placed after a semi-colon, but it can legitimately be placed at the beginning of a following sentence, and frequently is". (FOWLER, 1983, p.121).

In line with that, Biber et al (1999, p. 83) argue that "there is a well-known prescription reaction against beginning an orthographic sentence with a coordinator. Nevertheless, in actual texts we quite frequently find coordinators in this position". Interestingly, it seems that NNS have adhered to that type of prescription to a larger extent than NS; the use of *but* at the beginning of a sentence is almost two times more common in the latter group. Still, the use of such a coordinator among NNS is quite high, nearly equalling the use of *nevertheless* and *on the other hand*, the two second most used linking adverbials found in CRAB.

Also very evident in the data presented in the graph is the fact that, NNS have a strong preference for the three aforementioned linking devices - *but*, *nevertheless*, and *on the other hand*. Notably, NNS underuse all other linking adverbials in the contrast/concession category in comparison to NS. In contrast, in CRAN, the various linking adverbials are more equally distributed, without much variation among them, which suggests that they have a more or less equal value semantic wise. The only exceptions are found in the frequency use of *but* and *on the contrary*.

Remarkably, *on the contrary*, which is seven times more used by NNS, seems to have been inappropriately used in many of its instances in CRAB. In some cases, where NNS used *on the contrary*, the most appropriate linking adverbial would have been *on the other hand* or *in contrast*. As stated by Burchfield (2000, p. 180) the phrase *on the contrary* is only accurately used "in a statement intensifying a denial of what has just been stated or implied", or, as asserted by Swan (2005, p. 144) this linking adverbial is used "to contradict - to say that what has been said or suggested is not true". In this sense, it would be closest in meaning to 'quite the opposite'. Below is an example of the misuse of that adverbial in text about metaphors.

[31] “*The foot and the inch, initially representing the size of the king's foot and thumb respectively, have today taken on fixed values within a duodecimal system. The HEAD, on the contrary, has no fixed value, but only an approximate spatial notion. . .*” [CRAB, 17]

It is clear that the use of *on the contrary* in the example above is not the most appropriate. The metaphors using *head* do not contradict those that refer to the *foot* and the *inch*. Thus, linking adverbials such as *on the other hand* or *in contrast* could have been used for the establishment of a contrastive semantic relation without any conflict in meaning. An example of the adequate use of *on the contrary* is shown in the example below, in which readers understand that the first stretch of text is contradicted by the second.

[32] “. . . *I found no evidence that interrupting the ongoing reading process to search out the meaning of an unfamiliar word has a negative impact on L2 readers' comprehension, either word comprehension or global text comprehension. On the contrary, there is growing evidence that when students have access to the meaning of unfamiliar words either through a dictionary or marginal gloss, both local word comprehension (...) and global text comprehension are enhanced...*” [CRAN, 27a]

#### 4.2.4 Summation

All of the linking adverbials within the category of summation were not statistically significant and were thus not scrutinised. They include: *to sum up*, *all in all*, *in short*, *to conclude*, *to summarize*, *in sum*, and *in conclusion*. The reason behind that lies in the fact that the semantic category of summation includes adverbials whose uses are limited to specific parts of a research article, most often the end of a section and the conclusion section.

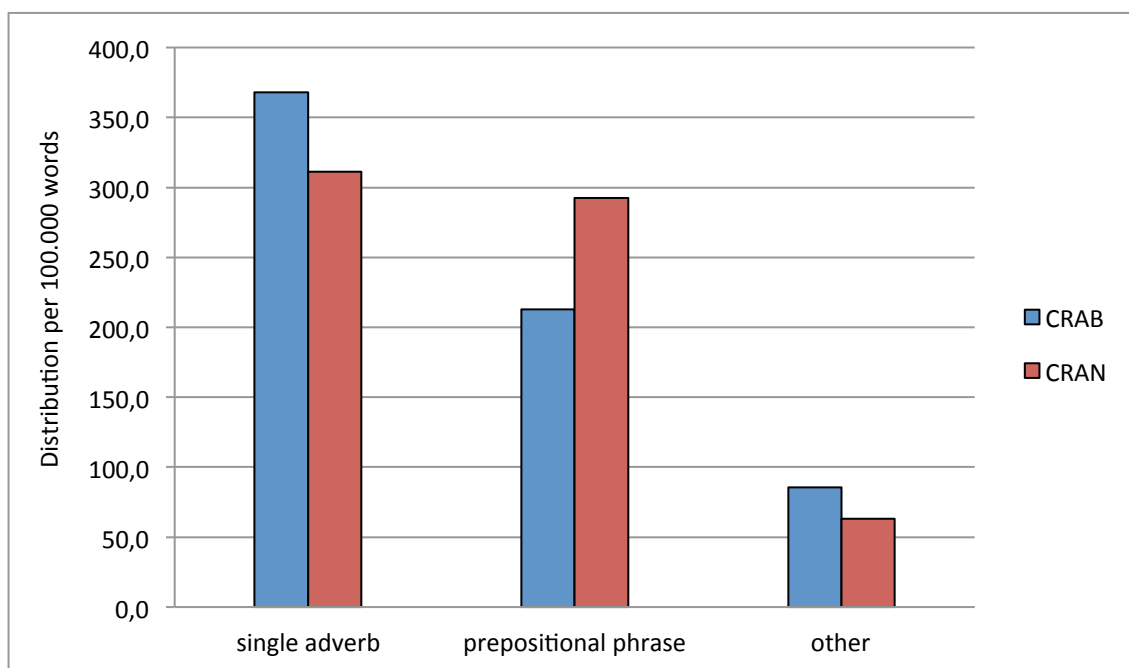
Nevertheless, despite the relative low frequency of linking adverbials of this category, it is worth mentioning that NS show a strong preference for the use of *in short* (167% more times than NNS). Brazilians, on the other hand, strongly

prefer *in sum*, which is slightly more used than their second and third most common linking adverbials: *to conclude* and *to summarise*, which, in turn, do not seem to be very productive in NS's writing.

### 4.3 Syntactic realisation of linking adverbials

As mentioned earlier in this study, linking adverbials may be realised by a variety of syntactic structures, which include items such as *single adverbs* (however, finally), *adverb phrases* (more precisely, even so), *prepositional phrases* (for example, in addition), *finite clauses* (that is, that is to say), and *non-finite clauses* (to conclude, added to that).

In our corpora, however, two of those forms account for the majority of the occurrences: single adverbs and prepositional phrases, as can be seen below.



Graph 10: Syntactic realisation of linking adverbials in CRAB and CRAN.

The data from CRAB show that Brazilians have great preference for single adverbs to connect messages in a text. CRAN also reveals that NS favour single adverbs in the array of options to achieve the same purpose. Also, both

corpora have prepositional phrases as their second most used forms of linking adverbials. However, the distribution in CRAN is more even than in CRAB. In texts written by Brazilians, single adverbs are used 73% more times than prepositional phrases, whereas in texts written by NS this difference accounts for only 6%.

As regards the syntactic forms of linking adverbials, Biber et al. (1999) have compared two registers, conversation and academic prose, and they have also found that single adverbs are the most common syntactic form realised by linking adverbials. As the conversation register is not the focus of this study, we will focus on the right side of the graph shown in the figure below, which presents the distribution of linking adverbials in the academic prose register.

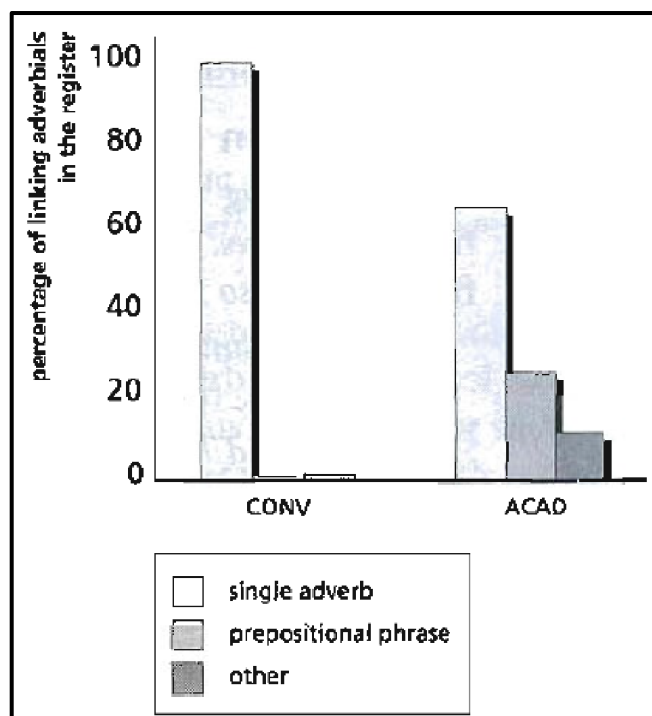


Fig. 4: Syntactic realisations of linking adverbials in conversation and academic prose  
Source: Biber et al. (1999, p. 884).

It is worth mentioning, one more time, that the data presented in the graph includes a wide range of different disciplines in the academic domain. This means that it serves as a good picture of how linking adverbials are used in the

academic prose register as a whole. In contrast, our findings come from data in a specific discipline, applied linguistics, and therefore it is not surprising that there are differences between CRAB/CRAN and Biber et al.'s figures.

Having said that, it is interesting to note that in academic prose as a whole, there is a great discrepancy between the number of single adverbs and prepositional phrases. This great discrepancy, however, does not happen in the academic research articles written by NS, as just mentioned above (the difference is only 6%). Nevertheless, there seems to be some correspondence between NNS applied linguists' writing and the general academic prose register. Again, this could indicate that, as far as the use of linking adverbials is concerned, professional Brazilian applied linguistics seem to follow a more general trend when they write, and that they do not use the preferred style used specifically in the applied linguistics field.

## 5. CONCLUSION

With this study, it was possible to draw some very interesting conclusions regarding the use of linking adverbials by NNS professional writers in comparison to NS'. In the lines that follow, I intend to elucidate the facts that can possibly contribute to a better understanding of professional academic writing by highlighting the most significant differences found in the two corpora analysed, CRAB and CRAN.

First of all, NNS only slightly underuse linking adverbials in comparison to NS. This assertion, however, has to be taken with caution, as the differences in the use of these adverbials only surface when the various semantic categories are taken into consideration. With respect to that, an initial interesting finding was that NNS in the field of applied linguistics use linking adverbials in a way that parallels not NS within the same field of study, but rather NS from various disciplines put altogether. As texts written by authors of the same discipline tend to portray similarities which are specific to that discipline (HYLAND, 2009), this finding may indicate that, as far as linking adverbials are concerned, Brazilian applied linguists portray a more general style in writing, thus probably indicating that their writing style was unaffected by the readings they did on applied linguistics when they were already professionals of that field.

Differences between NNS and NS are made more explicit when, as stated above, different semantic categories are taken into consideration. In this case what we see is that NNS overuse linking adverbials belonging to the semantic domains of enumeration, addition, and result/inference. In contrast, they underuse the ones that belong to the categories of apposition and contrast/concession. The category of summation, however, is the only one which presents virtually no difference in frequency between the two corpora.

Notably, when considering the analyses of each semantic category separately, some differences are even more apparent. For the category of enumeration, it appears that when signalling the start of a series of events or arguments, Brazilians make use of a wider range of linking adverbials. NS, on the other hand, predominantly make use of *first* to convey the same type of message. Also, NNS tend mark the end of events explicitly using *finally*, whereas that linking adverbial is not as frequent for NS.

For the category of addition, it could be said that NNS seem to be less conservative in their writing, since they make use of less formal adverbs such as *besides* and *also* in their texts. However, it has also been found that NNS overuse virtually all linking adverbials that belong to this semantic domain. This could be explained by the fact that NNS add messages to their arguments more explicitly than NS do, or that NS have more ways of adding information, such as through the use of stance adverbials with an additive value, such as *importantly* and *interestingly*, *for example*.

Linking adverbials in the category of apposition, the most common of all in research articles, are much more frequent in NS' texts than in NNS'. Since this type of linking adverbial is used to exemplify or reformulate an idea, it appears that NS' texts could be said to sound more instructive to a certain extent, implying that authors seem to connect to the reader as though in a classroom communication. It has also been found that the use of *for example* and *for instance* differ in terms of meaning in our two corpora.

The category of result/inference also reveals some interesting facts. Brazilians prefer the use of *thus* to signal the results or consequences of previously stated messages in their texts. NS, on the other hand, use *thus* and *therefore* more equally in terms of frequency. Also, the combination *thus* + *present participle* is three times more common in NNS' writing than it is in NS's. As for the meanings associated with these linking adverbials, Brazilians seem to take *thus* and *therefore* as synonymous, whereas NS treat *thereby* as a synonym to *thus*, not *therefore*.

The last category analysed, contrast/concession, is the one that presents the greatest number of different linking adverbials. Interestingly, the most common

connector within this category is not actually a linking adverbial, but rather the coordinating conjunction *but*. With respect to its use in research articles, NS tend not to conform to prescriptive rules that claim that *but* should be avoided as a sentence initiator. In contrast, NNS use *but*, *nevertheless*, and *on the other hand* with virtually the same frequency. An even more striking difference, however, is that Brazilian tend to underuse almost all linking adverbials in this semantic category.

Other remarkable findings in this research have to do with misuses or inappropriateness in NNS writing. The first ones are the use of *besides* and *so*. In a comparison with NS writing, Brazilians highly overused these two linking adverbials which, as it is known, are considered too informal to be found in academic research articles. Moreover, Brazilian's use of *thereby* also seems inappropriate when this linking adverbial is written in between commas, thus resembling the use of *therefore*. But perhaps the most utterly incongruous misuse of a linking adverbial is that of *on the contrary*. NNS many times mistakenly used this linking adverbial to mean *on the other hand* or *in contrast*.

Moving on we arrive at the discussion on the *syntactic* realisations of linking adverbials. Two of the five possible forms assumed by linking adverbials are the most prominent in CRAB and CRAN: single adverbs and prepositional phrases. Interestingly, however, once again professional Brazilian applied linguistics seem to follow a more general trend when they write, as revealed by comparing our data to that found in Biber et al.(1999).

Having said that, we would like to point out that this research could be further enhanced by adding to its analyses a discussion on the position of linking adverbials in the clause. That, however, would require that the researcher write a program that would be able to automatically retrieve this type of information. In addition, further research using corpora built with texts from different scientific fields would be necessary to better understand how Brazilian professional writers produce their texts. This, we believe, could shed more light to interventions which would aid authors to better write for publication. Finally, we expect that undergraduate and graduate researches can profit from corpus-

based writing workshops which could be designed from the findings of this and other similar research in order that they, too, can improve their writing skills.

As stated in the introductory section of this study, the findings we intended uncover with this research are only a tiny piece of a much more complex jigsaw that composes the hard work of writing. We, however, hope that we have been able to contribute to the enhancement of knowledge in this field of study, though we are aware that there is still a great deal of work to be done regarding empirical research in writing.

## REFERENCE

ANGOURI, J. Quantitative, qualitative or both? Combining methods in linguistic research. In: LITOSSELITI, L. (Ed). *Research Methods in Linguistics*. London: Continuum, 2010. p. 29-49.

BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, SP: Manole, 2004.

BIBER, D. et al. *Longman Grammar of Spoken and Written English*. Essex: Pearson Education, 1999.

BIBER, D. Representativeness in corpus design. *Literary and Linguistic Computing*, v.8, n.4, p 243–57, 1993.

BIBER, D.; CONRAD, S.; CORTES, V. If you look at...:Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics*. v.25, n.3, p. 371-405, 2004.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 1998, 300 p.

CARTER, R.; MCCARTHY, M. *Cambridge Grammar of English*. Cambridge; New York: Cambridge University Press, 2006.

CELCE-MURCIA, M., LARSEN-FREEMAN, D. 1999. *The Grammar Book: an ESL/EFL Teacher's Course*. 2<sup>nd</sup> ed. Boston: Heinle & Heinle, 1999.

CONRAD, S. M. The importance of corpus-based research for language teachers. *System*, v. 27, n1, p. 1-18. 1999.

COOLIDGE, H. J. ed. *Archibald Cary Coolidge: Life and Letters*. Boston: Houghton Mifflin Company, 1932.

CRANE, P. A. Texture in Text: A Discourse Analysis of a News Article Using Halliday and Hasan's Model of Cohesion. *Journal of School of Foreign Languages*. n. 30, p.131-156, 2006.

CURRY, M. J.; LILLIS, T. M. Academic research networks: Accessing resources for English-medium publishing. *English for Specific Purposes*. v. 29, n. 4, p. 281–295, 2010.

DISNEY, H. Evaluating Scientific Papers and Research Proposals. *Cortex*, v. 37, n. 4, p. 583–584, 2001.]

DUNNING, T. Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, v.19, n. 1, p. 61-74, 1993

EMIG, J. Writing as a Mode of Learning. *College Composition and Communication*, v. 28, n. 2, p. 122-128, 1977. Disponível em: <<http://www.jstor.org/stable/356095>>. Acesso em: 28 de outubro, 2014.

EVISON, J. What are the basics of analysing a corpus? IN: O'KEEFFE, A.; MCCARTHY, M. *Routledge Handbook of Corpus Linguistics*. Oxford: Routledge, p. 122-135, 2010.

FILHO, J. E. C. A responsabilidade ética das Revistas Científicas frente à pressão para publicar. *Rev. Bras. Saude Mater. Infant*, v. 12, n. 1, 2012.

FLOWERDEW, J. Problems in writing for scholarly publication in English: The case of Hong Kong. *Journal of Second Language Writing*, v. 8, n. 3, p. 243–264, 1999.

\_\_\_\_\_. Scholarly writers who use English as an Additional Language: What can Goffman's "Stigma" tell us? *Journal of English for Academic Purposes*, v. 7, n. 2, p. 77–86, 2008.

FOWLER, H. W. *Fowler's Modern English Usage*, 2<sup>nd</sup> edition. Oxford/New York: Oxford University Press, 1983

GIBBS, W. W. *Scientific American*. Lost Science in the Thrid World, 1995.

GREENBAUM, S. *Oxford English Grammar*. Oxford: Oxford University Press, 1996.

GRIES, S. *What is Corpus Linguistics?* Language and Linguistics Compass 3. University of California, 2009, pp. 1225–1241.

GRONDONA, C.B. *Eliminação de pacotes lexicais relacionados ao tópico e de pacotes lexicais em contexto de sobreposição: uma proposta metodológica para os estudos da linguística de corpus*. 2014.

Dissertação (Mestrado em Linguística Aplicada) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, Mar, 2014.

HALLIDAY, M. A. K.; HASAN, R. *Cohesion in English*. New York: Longman, 1976.

HALLIDAY, M. A. K.; MATTHIESSEN, C. M. I. M. (2004). *An Introduction to Functional Grammar* (3rd ed.). London: Edward Arnold, 2004.

HEWINGS, M. English language standards in academic articles: Attitudes of peer reviewers. *Revista Canaria de Estudios Ingleses*, n.53, p.47-62, 2006.

HUDDLESTON, R. D.; PULLUM, G. K. *The Cambridge grammar of the English language*. New York: Cambridge University Press, 2002.

HYLAND, K. Writing in the disciplines: Research evidence for specificity. *Taiwan International ESP Journal*, v. 1, n.1, p 5-22, 2009.

KACHRU, B. B. Standards, codification and sociolinguistic realism: The English language in the Outer Circle. In R. Quirk; H. Widdowson (eds.), *English in the World: Teaching and Learning the Language and Literatures*, 1985.

LEE, D.; SWALES, J. A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*. v. 25, p. 56–75, 2006.

LEI, L. Linking adverbials in academic writing on applied linguistics by Chinese doctoral students. *Journal of English for Academic Purposes*, v.11, n.3, p. 267-275, 2012.

LIU, D. L. Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, v. 13, n.4, p. 491-518, 2008.

MARQUES, A. L. S. P. O papel da percepção consciente de conectivos concessivos na produção escrita e na reformulação da interlíngua. 2007, 174f. Dissertação. (Mestrado) Universidade Federal de Minas Gerais. Faculdade de Letras da Universidade Federal de Minas Gerais. 2007.

MARTIN, J. R. Cohesion and Texture. IN: Schiffrin, D. ;Tannen, D.; Hamilton, H. E. (eds.). *The Handbook of Discourse Analysis*. Oxford: Blackwell, 2001.

MCENERY, A.; XIAO, Z.; TONO, Y. *Corpus-based Language Studies: An advanced resource book*. London / New York: Routledge, 2006.

MCENERY, T.; HARDIE, A. *Corpus linguistics: method, theory and practice*. Cambridge, New York: Cambridge University Press. 2012.

MELLO, H; SOUZA, R.R. A Linguagem da Ciência: *Prospecção de Dados Baseados em Corpora*. Anais de Seminários Teóricos Interdisciplinares do SEMIOTEC – I STIS. vol 1, n.1, Nov, 2012.

MEYER, C. *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press, 2002

QUIRK, R.; GREENBAUM, S.; LEECH, G.; SVARTVIK, J. A *Comprehensive grammar of the English language*. London ; New York: Longman, 1985.

RICHARDS, J. C. *The language teaching matrix*. Cambridge: Cambridge University Press, 1990.

RÖMER, U.; WULFF, S. Applying corpus methods to writing research: Explorations of MICUSP. *Journal of Writing Research*, v.2, n.2 pp. 99-127, 2010.

SAMPSON, G. *Empirical Linguistics*. London: Continuum, 2002.

SEIDLHOFER, B. (Ed.). *Controversies in Applied Linguistics*. Oxford, New York: Oxford University Press, 2003, p. 269-323

SHAW, P. Linking adverbials in student and professional writing in literary studies: what makes writing mature. Nn: Charles, M.; Pecorari, D.; Susan, H. (eds). *Academic Writing: At the Interface of Corpus and Discourse*. London: Continuum. 2009, p. 215-235.

SWALES, J.M. What has happened to academic medicine? *The Lancet*, V. 327, pp.1194-1196, 1986.

\_\_\_\_\_. *Genre Analysis: Ensligh in Academic and Research Settings*. Cambridge & New York: Cambridge University Press, 1990.

\_\_\_\_\_. Variation in Citational Practice in a Corpus of Student Biology Papers. *Written Communication*. v. 31, n.1, pp.118-141, 2014.

SWAN, Michael. *Practical English Usage*. 3. ed. New York: Oxford University Press. 2005.

## **APPENDIX A**

### **DISCIPLINES CONSIDERED FOR THE AILA WORLD CONGRESS 2014**

#### **A: Language Acquisition and Language Processing**

1. First Language Acquisition
2. Second Language Acquisition
3. Reading, writing and visual Literacy
4. Psycholinguistics

#### **B: Language Teaching and Learning**

1. Mother Tongue Education
2. Standard Language Education
3. Second/Foreign Language Teaching and Teacher Development
4. Language and Education in Multilingual Settings
5. Educational Technology and Language Learning
6. Language Evaluation, Assessment and Testing

#### **C: Language in the Professions**

1. Business and Professional Communication
2. Translating, Interpreting and Mediation
3. Language and the Law/Forensic Linguistics
4. Language and the Workplace
5. Language in the Media and Public Discourse
6. Language, Health and Aging

#### **D: Language in Society**

1. Sociolinguistics
2. Language Policy and Planning
3. Bilingualism and Multilingualism
4. Intercultural Communication
5. Language and Ideology
6. Language, Culture and Socialization
7. English as a Lingua Franca and World Englishes
8. Language and Technology

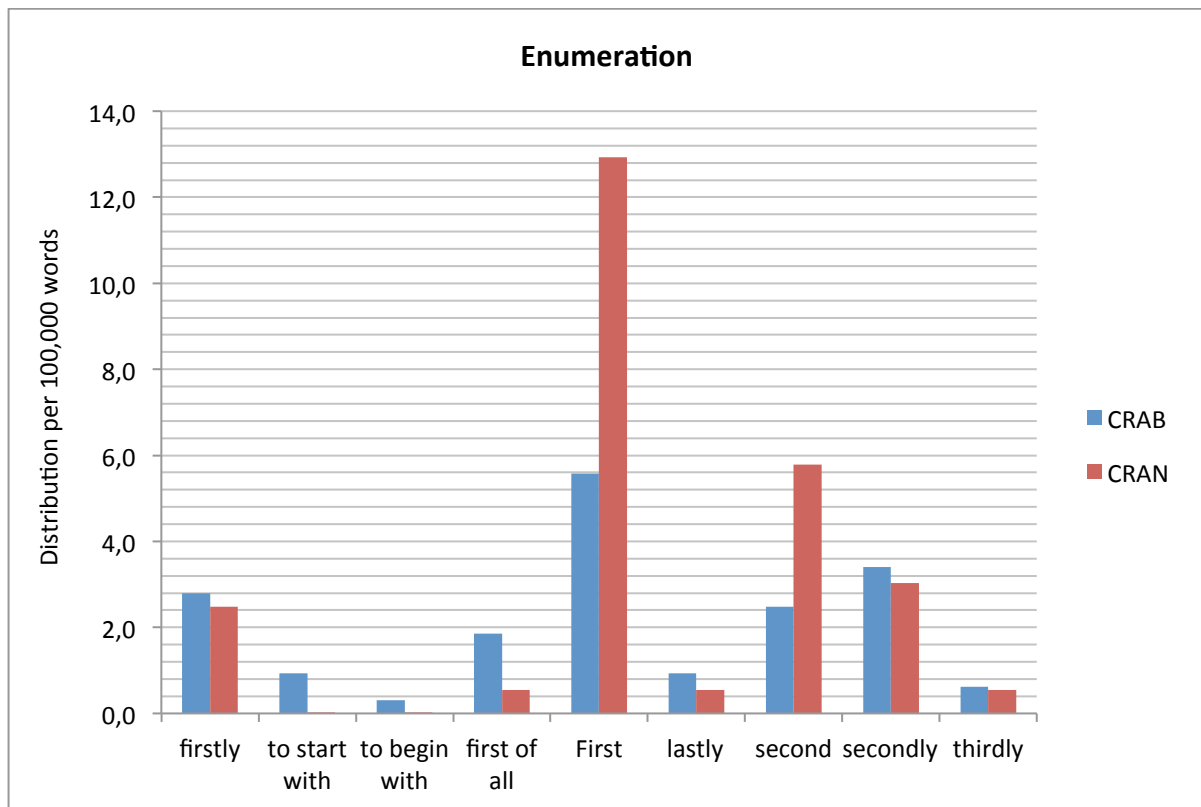
**E: Analysis of spoken and written discourse**

1. Language and Social Interaction
2. Pragmatics
3. Multimodality
4. Corpus Linguistics
5. Discourse Analysis
6. Rhetoric and Stylistics
7. Lexicography and Lexicology

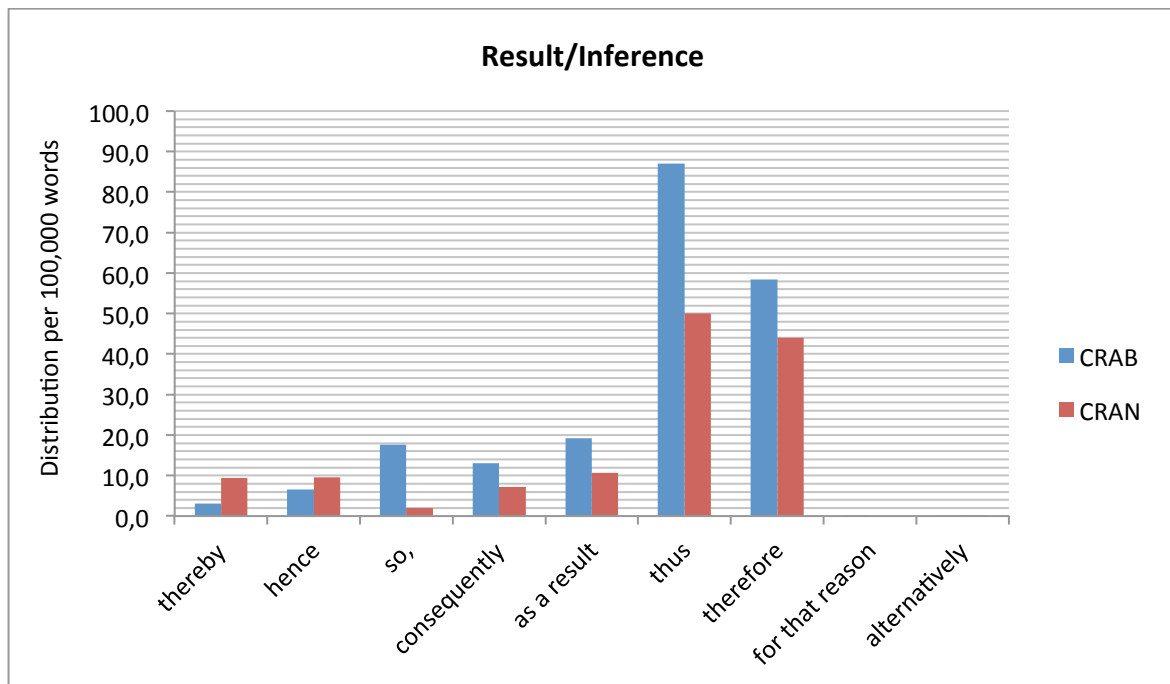
**APPENDIX B - LIST OF LINKING ADVERBIALS SEARCHED FOR IN THIS STUDY.**

- above all
- additionally
- again / once again
- all in all
- also
- alternatively
- alternatively
- and (sentence initial)
- another example
- as a result
- at any rate
- at the same time
- besides
- besides
- but
- by contrast
- conversely
- consequently
- eg / e.g
- finally
- first
- first of all
- firstly
- for example
- for instance
- for that reason
- further
- furthermore
- hence
- i.e.
- in addition
- in any case
- in conclusion
- in contrast
- in effect
- in other words
- in short
- in sum
- in turn
- lastly
- likewise
- meanwhile
- more specifically
- moreover
- namely
- nevertheless
- nonetheless
- notwithstanding
- on the contrary
- on the other hand
- or (sentence initial)
- otherwise
- second
- secondly
- similarly
- so,
- still,
- that is,
- thereby
- therefore
- thirdly
- thus
- to begin with
- to conclude
- to start with
- to sum up
- to summarize
- yet,

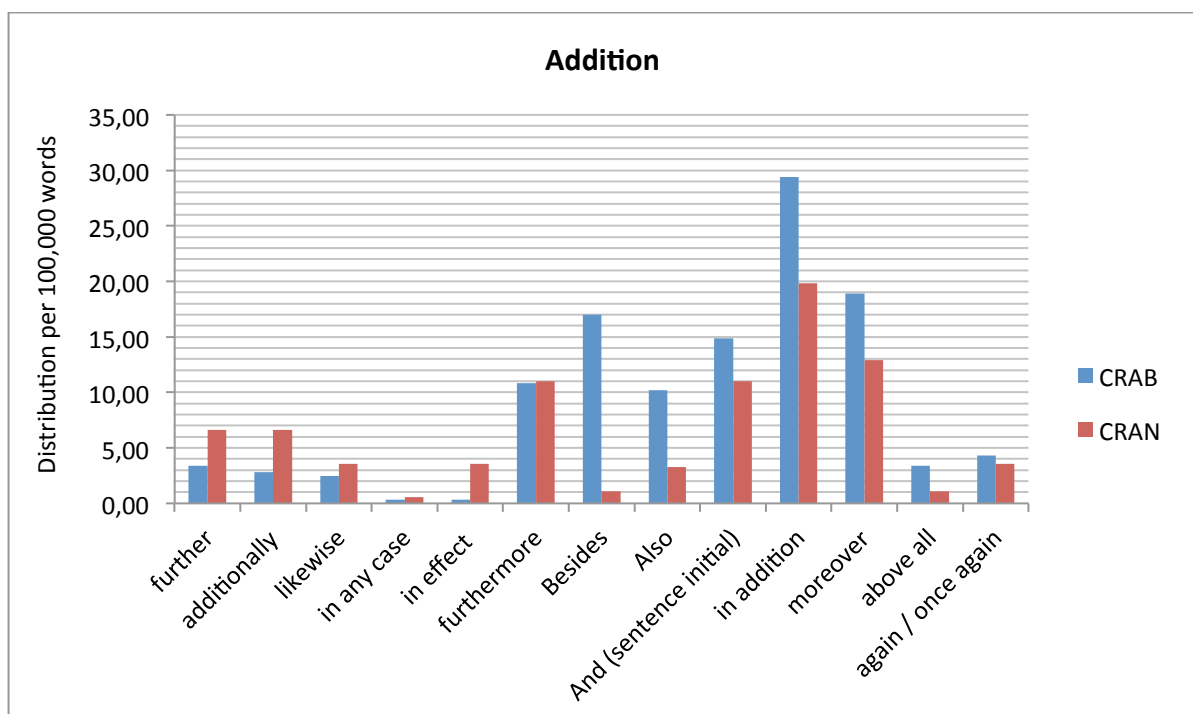
**APPENDIX C - GRAPHS CONTAINING LINKING ADVERBIALS ACCORDING TO SEMANTIC CATEGORIES: STATISTICALLY SIGNIFICANT AND NON-SIGNIFICANT ADVERBIALS**



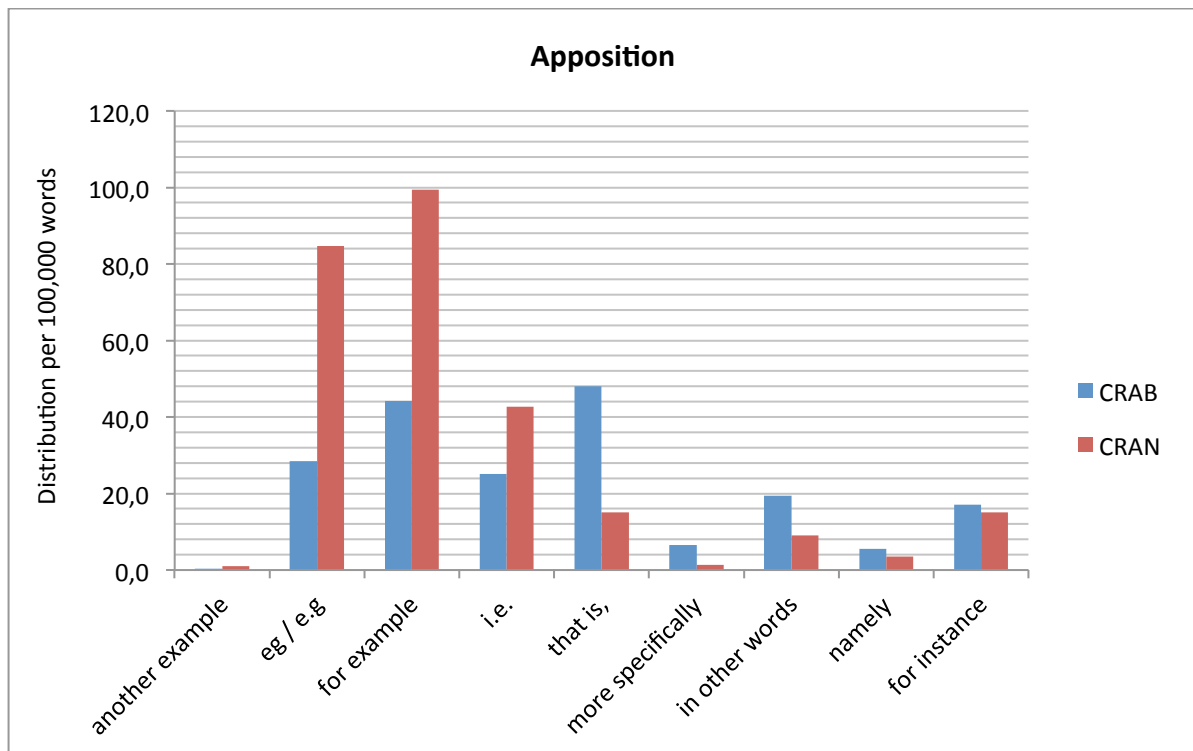
Graph 1a : Enumeration, excluding finally (finally distorted the results for its use is very high)



Graph 2a: result/ inference

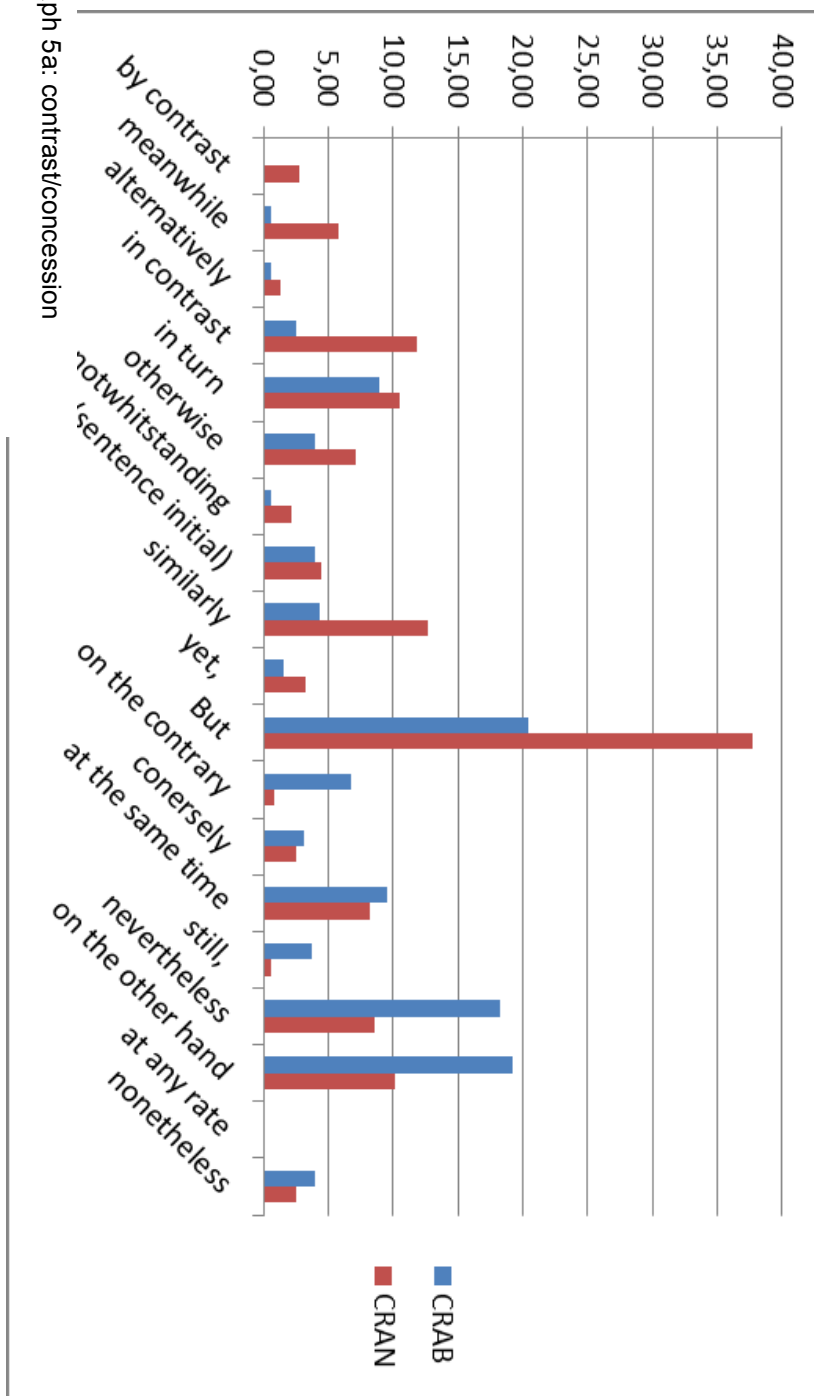


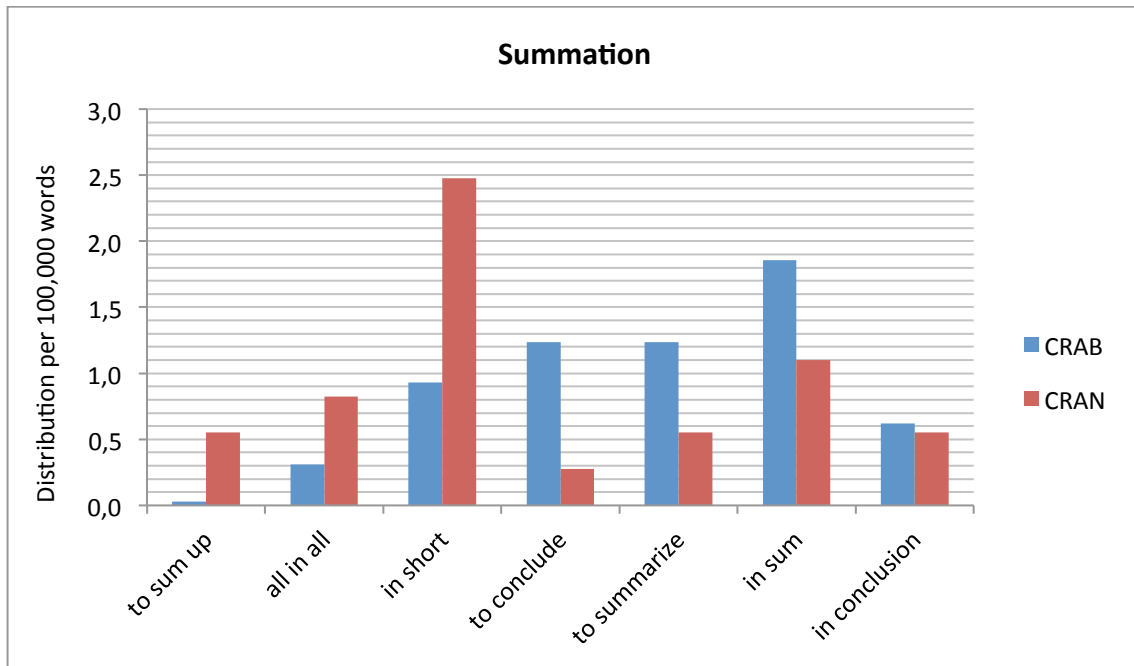
Graph 3a: addition



Graph 4a: apposition

Graph 5a: contrast/concession





Graph 6a: summation