Result linking adverbials in learner corpora

O uso de conectores resultativos em *corpora* de aprendizes

*Deise Prina Dutra**

*Barbara Malveira Orfanó***

*Valdênia Carvalho e Almeida****

ABSTRACT: This article aims to present an investigation of the use of the following linking adverbials: *so, therefore,* and *thus* in essay writing of Brazilian university students. Few corpora-based studies have been carried out on the use of cohesive devices in the written production of Brazilian university students in comparison to the writing of American and British university students. It is worth mentioning Matte and Sarmento (2018) and Dutra et al. (2017) who identified differences among the groups concerning the use of additive linking adverbials. The present study compares the occurrences of result linking adverbials in the Corpus of English for Academic Purposes (*CorIFA*), in the Corpus of English Without Borders (*CorIsF*), and in the Louvain Corpus of Native English Essays (*LOCNESS*). The linking adverbials were identified by using *AntConc* 3.4.4 (ANTHONY, 2016). The concordance lines generated by the program were analyzed, and the linking adverbials were classified according to their meaning and syntactic position. Then, the data were submitted to the

RESUMO: Este artigo tem como objetivo apresentar um estudo sobre o uso dos seguintes elementos de coesão textual: *so, therefore*, e *thus* na escrita de redações de estudantes universitários brasileiros. Poucos estudos baseados em *corpora* foram desenvolvidos sobre o uso de elementos de coesão na escrita de alunos universitários brasileiros em relação à escrita de universitários ingleses e americanos. Destacam-se Matte e Sarmento (2018) e Dutra et al. (2017) que observaram diferenças entre os grupos quanto ao uso de elementos que expressam adição. Neste trabalho comparamos as ocorrências dos elementos de coesão resultativos, os quais podem ser chamados de conectores resultativos[1], no Corpus de Inglês para Fins Acadêmicos (CorIFA), no Corpus do Inglês sem Fronteiras (CorIsF) e no *Louvain Corpus of Native English Essays* (LOCNESS). Identificamos esses elementos utilizando o *AntConc* 3.4.4 (ANTHONY, 2016), analisando linhas de concordância e classificando sua posição sintática e seu sentido em cada sentença. Em seguida, os dados foram submetidos

---

* Doutora, Universidade Federal de Minas Gerais. deisepdutra@gmail.com

** Doutora, Universidade Federal de Minas Gerais. barbara.orfano@gmail.com

*** Doutora, Universidade Federal de Viçosa. valdeniaalmeida5@gmail.com

1 Em português uma construção semelhante, a qual utiliza esses elementos de coesão, poderia ser interpretada como uma construção resultativa. O uso do termo 'conector' foi escolhido para diferenciar de 'conjunção' que não é o foco do presente trabalho.

statistical log-likelihood test and the overuse, and underuse of items were studied across the dataset. The results indicate differences mainly concerning syntactic position and meaning of the linking adverbials. Analyzing these differences is instrumental as it may contribute to a better understanding of Brazilian learners' written academic discourse and, in turn, it can assist language educators and material designers who work with to the teaching of English for Academic Purposes.

KEYWORDS: English for Academic Purposes. Learner corpus. Result linking adverbials.

ao teste estatístico log-likelihood e o subuso ou sobreuso dos itens foram investigados. Os resultados indicam diferenças, principalmente, em relação à posição sintática e ao sentido do conector. A análise dessas diferenças é relevante, pois contribuem para uma melhor compreensão do discurso acadêmico escrito de aprendizes universitários brasileiros e podem, futuramente, auxiliar professores e autores de materiais didáticos que tenham como foco o ensino de inglês para fins acadêmicos.

PALAVRAS-CHAVE: Inglês para fins acadêmicos. Corpus de aprendizes. Conectores resultativos.

1. Introduction

Producing a text that expresses what an author has in mind seems to be a difficult path for many people. Members of different communities wish to write clearly, including the ones who belong to the academic community. This community is not homogenous and many researchers have acknowledged that there are several academic communities according to specific disciplines (i.e. HYLAND, 2009), for instance, the community of applied linguists, of chemists and of historians.

> ... the notion of community puts individual's decision-making and engagement at centre-stage and underlines the fact that academic discourse involves language users in constructing and displaying their roles and identities as members of social groups. (HYLAND, 2009, p. 52)

Being a member of an academic community implies that one knows the implicit genre rules (SWALES, 1990). For novice writers, eager to enter a specific discourse community, learning such rules is paramount. At first, they need to be able to

distinguish academic genres (SWALES; FEAK, 2012) that can range from statement of purpose to research articles (SWALES; FEAK, 2011). Second, this awareness of academic textual genres should prepare them to learn genre structure as well as the linguistic preferences of the more experienced community writers. "If genres are symbolic worlds, we textually inhabit in order to participate in consequential actions, then part of invention must involve explicitly examining the conventions that shape these worlds." (JOHNS *et al.*, 2006, p. 244). In order to help novice writers face these challenges to clearly understand genres, researchers, material designers and teachers need to know their necessities. Describing learners' use of the language they are learning is, therefore, the focus of learner corpus research (LCR) (GRANGER, 1998; GRANGER, *et al.*, 2013). LCR is a vast area in corpus linguistics, contributing empirically to the understanding of language usage (RANKIN, 2015) by non-native speakers of different languages.

> This new research strand emerged in the late 1980s as an offshoot of corpus linguistics, a field which had shown great potential in investigating a wide range of native language varieties (diachronic, stylistic, regional) but had neglected the non-native varieties. (GRANGER *et al.*, 2015, p. 1)

Previously to LCR learners' data had been mainly investigated by second language acquisition (SLA) researchers. The two main advantages of the LCR over the SLA approach is: a) the large amount of data that can be analyzed; b) the electronic format of data allows the use of a variety of software tools enhancing the possibilities of extracting and analyzing linguistic features (GRANGER *et al.*, 2015). The LRC studies since the 1980s have focused on learner corpus design and methodology (GILQUIN, 2015; GRIES, 2015), analysis of learner language (COBB; HORST, 2015) and learner corpus as a pedagogical tool (CHAMBERS, 2015) among other investigation topics. In this article we present a descriptive study which concentrates on the analysis

of learner language, following the contrastive interlanguage analysis (CIA) (GRANGER 1996, 2015). The CIA consists of comparing students' oral or written data with native speakers' production, first language (L1) vs second language (L2) or comparing different types of learner corpora (L2 vs L2). One of the main characteristics of CIA is to enable a better understanding of learners' production, taking into consideration different learning environments, the genre produced and their specific characteristics. Granger (2015) explains that LCR in combination with CIA have contributed to extend the range of linguistic phenomena investigated mostly in the areas of lexis and discourse. Another important characteristic of CIA is the identification of overuse and/or underuse of linguistic features. A word or expression is found to be overused or underused when its frequency is statistically higher or lower when compared to a reference corpus. This corpus is one that provides comprehensive information about a language and its data come from successful users of a language, native or highly proficient non-native speakers. Some authors (i.e., SEIDLHOFER, 2001) claim that corpora of English as a Lingua Franca should be the focus of language studies. The CIA does not discard this possibility, but it also considers valid (GRANGER, 2012, p. 18) the comparison of "either two learner varieties (L2 vs. L2) or one learner variety and one native (or expert) variety (L2 vs. L1)."

Comparative analysis of learners′ and native speakers' writings have previously focused on a variety of linguistic aspects, for instance, spatial prepositions (RORVIK; EGAN, 2013), vague language in spoken interlanguage (ORFANÓ, 2013) and lexical bundles of referential meaning (DUTRA; BERBER SARDINHA, 2013). Other studies have also investigated the use of adverbial connectors by Swedish learners of English (ALTENBERG; TAPPER, 1998) and use of connectors[2] in the

---

[2] Connectors are used as a cover term by Rorvik and Egan (2013) to refer to conjunctions and adverbial connectors. In this paper we will use the term linking adverbials (BIBER *et al*., 1999) rather than adverbial connectors. More details on this terminology issue will be given in the following article section.

writing of Norwegian novice writers (RORVIK; EGAN, 2013). Granger and Tyson (1996) influenced closely our study as they investigated the use of linking adverbials by French learners of English and their data come from the French subcorpus of the International Corpus of Learner English (ICLE[3]). The results revealed the French learners' underuse the linking adverbials *however, therefore, thus* and overuse *moreover, for instance, of course, on the contrary*. The qualitative analysis showed the inappropriate use of some linking adverbials by the French learners of English. Linking adverbials characteristic of informal language, such as *anyway* and *so* were commonly used by the French learners in their essays. For the authors, this inappropriate use has its roots in the widespread teaching approach, which does not emphasize style issues and does not make students aware of oral and written discourse differences. They have also identified that the French learners have a preference for using the linking adverbials in initial syntactic position.

There have been few studies focusing on Brazilian learner use (DUTRA; SOUZA, 2012) and learning (MARQUES, 2007) of linking adverbials. Highlight should be given to two other recent studies based on corpus. Matte and Sarmento (2018) investigated the use of linking adverbials in a corpus composed of texts written by Brazilian learners studying in British universities as compared to highly graded texts in the British academic written English corpus (BAWE). Their results showed that the Brazilian learners underuse some cohesive devices and overuse others (e.g. enumeration and addition category). Dutra *et al*. (2017) concentrated on the use of additive linking adverbials in essays through the analysis of two learner corpora - *Corpus de Inglês para Fins Acadêmicos* (*CorIFA*), *Corpus do Inglês sem Fronteiras* (*CorIsF*) and one reference corpus - Louvain Corpus of Native English Essays (LOCNESS). The occurrence of linking adverbials in Brazilian learners' texts was compared to the occurrence in American and British university students' texts. The quantitative

---

[3] ICLE's coordinator is Sylviane Granger from the Université Catholique de Louvain.

analysis revealed an overuse of *besides* and an underuse of *also* by the Brazilian students in contrast to the native speakers' corpus. Above all, their analysis showed that semantic pragmatic distinctions in the use of these cohesive devices need to be addressed in pedagogical contexts, as Brazilian learners use, for instance, *besides* for listing points or as a redundant feature which contrasts with the result and inference meaning commonly observed in LOCNESS.

The fact that there are few studies based on Brazilian learner corpora shows that there is a paucity of research focused on these students. Above all, there is a need for more investigation on university Brazilian English learners' use of linking adverbials. As an LCR study which follows the CIA model, this article compares and contrasts the use of *so*, *thus* and *therefore* in three corpora of both native and non-native speakers of English. The descriptive results of the present article will likely enlighten future pedagogical decisions.

We have chosen to focus on the essay writing of Brazilian university students since this is a popular genre taught in English classes in Brazil and elsewhere. Our hypotheses are that:

- Brazilian learners overuse the result linking adverbial *so* and underuse *therefore* e *thus*;

- The three linking adverbials are more frequent in the initial position in the learner corpora;

- The three linking adverbials are more frequently used with the meaning of logical consequence.

2. Linking adverbials

For more than twenty years (GRANGER, 1998) the corpus linguistics community has systematically put effort into studying second language learners of a

variety of languages, especially, English. Consequently, it has contributed to our understanding of learner's necessities, being one of them the use of linking adverbials (i.e. RORVIK; EGAN, 2013). Before addressing specific results regarding English learners' use of linking adverbials, it is important to briefly present what linking adverbials are and why they seem to be linguistic features worth of close investigation.

Linking adverbials are cohesive devices that contribute to the cohesion of a text. They "are important devices for creating textual cohesion, alongside coordinators and subordinators" (BIBER *et al.*, 1999, p. 875). Zihan (2015) explains that in different studies the term linking adverbials has been addressed as 'conjunctive adjuncts', 'adverbial conjuncts', 'linking adjuncts', 'conjunctive adverbials', 'connectives', 'connectors', 'discourse markers' or 'pragmatic markers'. Linking adverbials is not a commonly accepted term by linguists. Even if one follows the tradition of descriptive English linguists, as we do, several terms come into play. Quirk et al. (1985) refer to this category as 'conjuncts' and Carter and McCarthy (2006), the authors of a widely accepted corpus-based English grammar, use the term 'linking adjuncts'. Zihan (2015) defends that regardless of trying to reach a consensus on the term to use, researchers should focus on investigating the usage patterns of such cohesive devices. Having that in mind, we have decided to use the term linking adverbials, as in Biber *et al.* (1999). This corpus-based grammar shows how different types of grammatical features vary across registers (spoken or written) and investigates the most frequent usage patterns. For our study, we highlight the fact that linking adverbials play a role in connecting ideas in texts, as they "explicitly identify the logical relations among independent clauses" (BIBER; GRAY, 2016, p. 240). Linking adverbials may be classified in six semantic categories: enumeration/addition, summation, apposition, result / inference, contrast / concession and transition (BIBER *et al.*, 1999) and our research interest in on Brazilian learners' use of result linking adverbials in English.

Since frequency and meaning matter in language use, other studies which focus on the meaning created by the use of certain cohesive devices have also called our attention (ZIHAN, 2014, 2016). Zihan (2016), taking in consideration five different registers[4]: academic lectures, conversation, written academic prose, written news and broadcast news, in the Wellington Corpora of Spoken and Written New Zealand English (WSC and WWC), investigated how some linking adverbials may vary in frequency and meaning across such registers. His careful semantic analysis proposed that there are twenty one "(21) different meaning categories … and that the academic lectures have the most meaning categories (18) eighteen" (ZIHAN, 2016, p. 12). One of his striking results was that the use of *so* to express the conclusion meaning occurred fifteen (15) times more in conversation than in written academic prose or written news.

3. Methodology

This article presents a corpus-based study and, therefore, the analysis was done so as to compare and contrast the occurrences of *so*, *thus* and *therefore* as result linking adverbials in three corpora of both native and non-native speakers of English. Is it important to highlight that a preliminary investigation of six semantic category linking adverbials was done. The categories studied were: enumeration, addition, summation, apposition, result / inference, contrast / concession and transition (BIBER *et al.*, 1999). The occurrences of 66 (sixty six) cohesive devices (e.g. *firstly*, *in addition*, *in sum*, *for instance*, *therefore*, *however*, *meanwhile*) were counted. We observed if they played a cohesive role and, consequently, could be omitted without making the sentence ungrammatical. Thus, only these occurrences were considered. We chose to further investigate, for this paper, *so*, *thus*, *therefore* for two reasons: a) there was a great

---

[4] Zihan (2016) chose to use the term 'register' rather than 'genre' (SWALES, 1990). He does that in the same way as Biber *et al.* (1999) and Biber, Gray (2016) do. These terms may be used with different or similar meanings. Yet, due to the lack of space in this paper, we will not be able to discuss them here.

frequency difference between the occurrences of these three linking adverbials in the learner corpora and in the native speaker corpus, and b) we had noticed, as English teachers, that students had difficulties in using these adverbials, but we had not systematically studied how students use them.

The native corpus corresponds to essays compiled for the Louvain Corpus of Native English Essays (LOCNESS), whereas the non-native corpora are represented by two corpora being compiled at *Faculdade de Letras* / Federal University of Minas Gerais (UFMG). For this specific study, our analysis concentrates on argumentative essays resulting in three comparable corpora. The overall description of the three corpora is presented in the next paragraph.

The Louvain Corpus of Native English Essays (*LOCNESS*) comprises argumentative essays written by American and British university students covering different topics. For this study, a subcorpus of LOCNESS, with 228,920 words, was used as our reference corpus. The other two non-native speaker corpora represent our learner data: The Corpus of English for Academic Purposes with 81,586 words and *the Corpus of English Without Borders* with 48,442 words, hereafter, *CorIFA* and *CorIsF* respectively[5]. The former consists of essays written by Brazilian undergraduate and graduate university students from a variety of disciplines taking English for Academic Purposes (EAP)[6] classes at UFMG that range from B1 to C1 level (intermediate to advanced), according to the Common European Framework of Reference for Languages (2001). The latter corresponds to the essays compiled in *The English Without Borders*[7] classes under the *Language Without Borders Program*[8].

---

[5] These numbers correspond to the size of the subcorpora in October 2017.
[6] The Portuguese acronym for EAP is *IFA* (*Inglês para Fins Acadêmicos*). Therefore, the corpus is called *Corpus de Inglês para Fins Acadêmicos* (*CorIFA*).
[7] This program offers English classes to students and employees at federal and state universities in Brazil. The aim is to improve language proficiency and create better opportunities for internationalization - for details on the IsF Program, check at Sarmento *et al.* (2016).
[8] All essays were collected for research upon students´ agreement (see Consent Form Sample in Appendix A). Examples of essay tasks and themes are in Appendices B and C.

First, using the *AntConc* 3.4.4 software (ANTHONY, 2016) concordance lines with the result linking adverbials, *so*, *therefore* and *thus*, were generated in order to analyze all occurrences of these cohesive devices in the three corpora. Then, the lines were analyzed, and phrases in which the item searched was not used as a result linking adverbial, such as *even so*, *so far*, *so on* and the uses of *so* as an adverb were excluded for the purpose of this study. Only the lines that contained *so, therefore* and *thus* used to connect result logical relations among independent clauses were kept as data for this study. In such cases, the omission of the cohesive device does not make the sentence ungrammatical. The next step was to analyze the linking adverbial syntactic position: initial, medial or final. This analysis concentrated on the position of each linking adverbial in the sentence. This position is evidenced by the use of punctuation and capitalization chosen by the students as they wrote their essays. Initial position means that the linking adverbial is the first element in the sentence. If it occurs in the middle or at the end of the sentence, preceded or not by a comma, the position was considered medial or final, respectively. The specific meaning of each result linking adverbial was also a focus of the study so as to allow the research to identify possible inappropriate uses in the corpora under investigation. In the following example, *so* is used in initial position to start a new idea which is not a frequent use in the academic context (BIBER et al. 1999), consequently, it was analyzed as inappropriate.

> [...] I think I am a practical learner, in other words, who gets the point without   wasting too much time. **So**, thinking about my process of learn English, I discovered that I learn better English through movie [...] (CorIFA-UFMG-B1.Ind.Ne.Ess.2013-1.0008.0008)

The assessment of inappropriate uses was conducted following the uses or meaning categories reported in Biber et al (1999). In this study each line containing *so, therefore* and *thus* was evaluated by two researchers in order to check the appropriateness of use of the linking adverbial. If there was a discrepancy in the

evaluation between the two researchers, then, the same line was independently assessed by a third researcher.

After all sentences that carried the targeted linking adverbials were counted, the log-likelihood (LL) test was applied to verify if there was a significant difference in the number of occurrences of the cohesive devices among the corpora. The log-likelihood (LL), a test of significance, was chosen to be performed on our data since it involves inferential statistics. Although some corpus-based studies use basic descriptive statistics, such as frequency count, percentage or normalized frequency (relative frequency)[9], our study uses an inferential statistic test required when comparison of occurrences in different size corpora are done. Such tests are used to test hypotheses, allowing the comparison between two sets of data. The LL results permits us to make inferences about differences in frequency between two corpora. In other words, it enables us to state if the difference is random or if it is statistically significant. According to Rayson (2003), if the log-likelihood result is greater than 3.84 and smaller than 6.63, there is less than 5% probability for the result to be by chance, being expressed by $p<0.05$[10]. If the LL result is greater than 6.63, we can be 99% sure that the result is not random. This means that "[T]he probability of the result - i.e. the difference between the two corpora - happening by chance is less than 1%"[11] and this is represented by $p< 0.01$. The higher the LL result, the more certain one can be that the difference between the two corpora is statistically relevant. The size of the two corpora is taken into consideration when the LL is calculated. The value that comes after the calculations is a positive number. The log-likelihood and effect size calculator that we used for this research is available at http://ucrel.lancs.ac.uk/llwizard.html. This

---

[9] Normalized frequencies are usually given per thousand words or per million words, which allows one to assume that a word might appear a certain number of times, for instance, per a thousand words of a running text.

[10] See more on log-likelihood on http://ucrel.lancs.ac.uk/llwizard.html

[11] Quote taken from https://www.lancaster.ac.uk/fss/courses/ling/corpus/blue/l08_4.htm

calculator, however, provides results with "an indicator for '+' overuse and '-' underuse of corpus 1 relative to corpus 2" (RAYSON, 2018, online). This means that the LL calculation is relative to corpus size and that it is able to detect overuse or underuse of specific linguistic features as it compares the relative frequency in two corpora. To calculate the log-likelihood value of the three linking adverbials, the occurrences of each linking adverbial in each learner corpus and the frequency of the same adverbial in the native speaker (e.g. CorIFA-LOCNESS - *thus* frequencies; CorIsF-LOCNESS - *thus* frequencies) were entered in the calculator as well as each corpus size. The data, then, were analyzed both quantitatively and qualitatively.

The hypotheses put forward in this study refer to the overuse of the linking adverbial, *so*, and the underuse of *therefore* and *thus* by the Brazilian learners of English whose texts are part of the studied corpora. Besides, the other hypothesis is that the items would be used more frequently in the initial position by the Brazilian learners. Thus, the quantitative analysis of the data aimed at the identification of overuse and underuse of *so*, *therefore* and *thus* whereas the qualitative analysis aimed at verifying the syntactic position and the meaning associated with the use of the cohesive devices.

3. Results

The quantitative results confirm, partially, our hypotheses. First, significant differences were detected once the frequency of *so* in the learner corpora was compared to its frequency in the native speaker corpus. Using the statistical test log-likelihood (see Table 1) to compare the occurrences in LOCNESS and in CorIsF, the result was LL= +77.13 (p< 0.0001**[12]), while in LOCNESS and in CorIFA, the outcome was LL= +13.34 (p.<0.001**)[13]. Thus, the difference in the use of *so* is statistically

---

[12] When the LL result is higher than 15.13, the probability of it happening by chance is less than 0.01%. Therefore, we can be 99.99% certain of the result.

[13] Since the LL result is higher than 10.83, the probability of it happening by chance is less than 0.1%. This allows us to be 99.9% sure the difference is not random.

significant at the level of 0.01% between LOCNESS and CorIsF and at the level of 0.1% between LOCNESS and CorIFA. The linking adverbial *so* is overused in both learner corpora. Second, significant differences were also observed through the log-likelihood statistical test between the occurrences of *therefore* in LOCNESS and in CorIsF (LL= -14.98, p < 0.001**) and between LOCNESS and CorIFA (LL= -5.63, p<0.05*). The difference in the use of *therefore* is statistically significant at the level of 0.1% between LOCNESS and CorIsF and at the level of 0.5% between LOCNESS and CorIFA. Thus, we can be 99.9 % certain that the linking adverbial *therefore* is underused in the essays that are part of CorIsF and 95% sure that the same adverbial is underused in CorIFA. Third, no significant results were found when the occurrences of *thus* were compared, considering CorIFA and CorIsF, separately, compared to LOCNESS. Our third hypothesis of underuse by the Brazilian students was not confirmed.

Table 1 – Raw frequency in each corpus - *log-likelihood* results between LOCNESS and each of the learner corpora (CorISF and CorIFA).

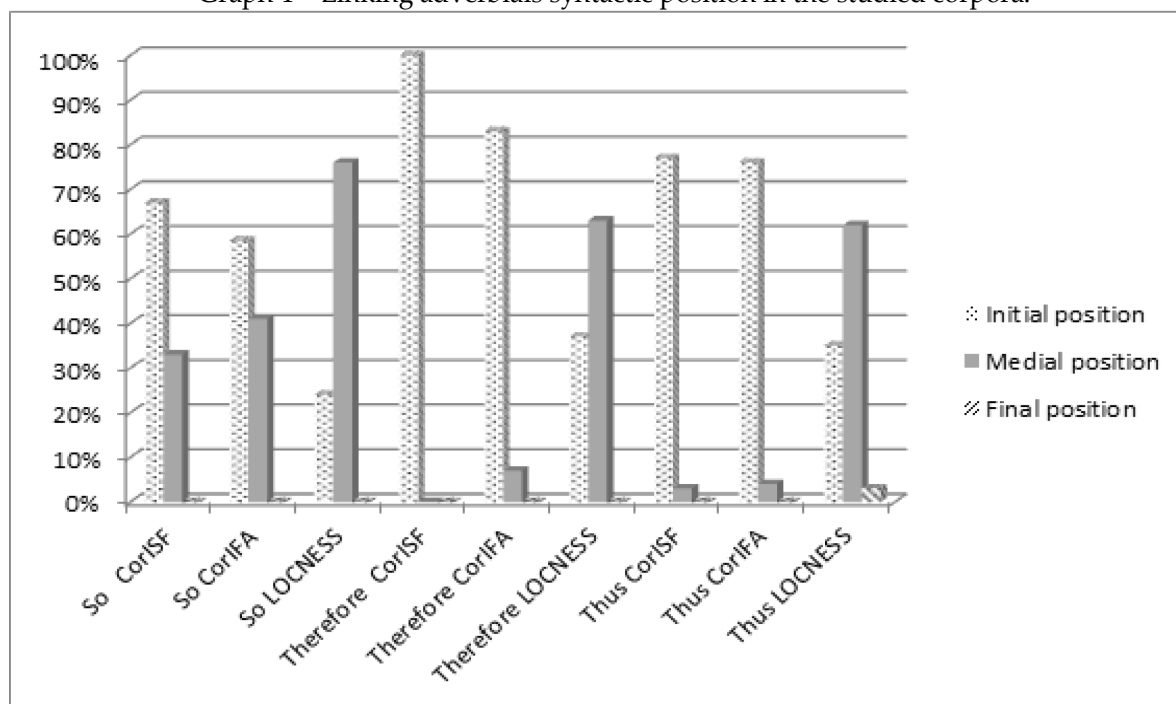| Linking adverbials	LOCNESS	CorIsF		CorIFA	
	occurrences	occurrences	LL	Occurrences	LL
So	243	136	**+77.13****	130	**+13.34****
therefore	157	11	**-14.98****	34	**-5.63****
thus	75	13	-0.46	25	+0.08

Source: the authors.
** = significant statistical results[14]

[14] Plus sign (+) before the LL result = overuse of corpus 2 (CorIsF) relative to corpus 1 (LOCNESS) or overuse of corpus 3 (CorIFA) relative to corpus 1 (LOCNESS)
Minus sign (-) before the LL result = underuse of corpus 2 (CorIsF) relative to corpus 1 (LOCNESS) or underuse of corpus 3 (CorIFA) relative to corpus 1 (LOCNESS)

It was observed that the preferred syntactic position of the result linking adverbials *so, therefore* and *thus* also varies if the learner corpora (CorIsF and CorIFA) are compared to LOCNESS (see Graph 1 and Table 2). Brazilian university English learners prefer the initial position for *so*, *therefore*, and *thus*, while university native speakers use the same linking adverbials more often in medial position, either connecting clauses or in the middle of a clause after the subject. These differences called our attention since the syntactic positions of linking adverbials "have slightly different effects" (CONRAD; BIBER, 2009, p. 88). In the following sections, such effects, which impact on the linking adverbial meaning in the specific linguistic context it is in, are discussed as samples of students' writings are presented.

Graph 1 – Linking adverbials syntactic position in the studied corpora.

Source: the authors.

Table 2 – Percentage of linking adverbial occurrences in syntactic position.

Linking Adverbials	Corpora	Initial	Medial	Final
So	CorISF	**67%**	33%	0
	CorIFA	**59%**	41%	0
	LOCNESS	24%	**76%**	0
Therefore	CorISF	**100%**	0	0
	CorIFA	**83%**	7%	0
	LOCNESS	37%	**63%**	0
Thus	CorIsF	**77%**	23%	0
	CorIFA	**76%**	24%	0
	LOCNESS	35%	**62%**	3%

Source: the authors.

3.1 *So*

The quantitative analysis, in terms of percentage, shows that *so* in medial position is more frequent than in the other positions in LOCNESS. Seventy six percent (76% - see Graph 1 and Table 2) of all occurrences of *so* in LOCNESS were in the middle of the sentence, connecting two clauses. In samples 1 and 2 the linking adverbial *so* has the function of showing the result of the idea presented in the first clause. The logical consequence in the second clause is announce by *so*.

> 1. The whole episode may send the price of beef down **so** people will buy it and take the risk (which must be tiny) rather than forking out for more expensive meat products. (LOCNESS - ICLE-ALEV-0002.7[15])

---

[15] ICLE-ALEV-0002.7 This code contains the following information:
ICLE: International Corpus of Learner English

2. Even if we did, then shorly **so** does the rest of the world. (LOCNESS - ICLE-ALEV-0001.16)

Both in CorIsF and CorIFA the linking adverbial *so* comes more often in sentence initial position (67% in CorIsF and 59% in CorIFA). In quantitative terms, the numbers are much higher than in LOCNESS (24% in initial position). The overuse of *so*, detected as the log likelihood test was run, as previously mentioned, is clearly shown. Once we focus on the syntactic position of this linking adverbial, the differences are also striking. In initial position, where it occurs more often, learners do not express the result meaning. The linking adverbial *so* tends to convey other meanings, such as getting to a conclusion (Samples 3 and 4), restating an idea or restatement (Sample 5) and beginning a topic (Sample 6). Zihan (2016) noticed that *so* was the only result linking adverbial that occurred across the five registers studied, being so much more frequent in conversation. Yet, the three meanings we found as common in both CorIsF and CorIFA, getting to a conclusion (also detected in CHEN, 2006), restating an idea and beginning a topic, were clearly more characteristic of the conversation register in the Wellington Corpora of Spoken and Written New Zealand English (WSC and WWC) (ZIHAN, 2016).

3. When we go abroad we develop a different point of view about the world and people. **So** if a lot of people go abroad our country will be better because these people will have new experience of life and improve their comunication[16]skills. (CorIFA-UFMG-B1.Ind.NE.Ess.2013-1.0026.0026[17])

ALEV: A level

0002.7: number given to each student's text and to each student

[16] Misspellings were not corrected in neither CorIsF nor CorIFA.

[17] CorIFA-UFMG-B1.Ind.NE.Ess.2013-1.0026.0026. This code contains the following information:

CorIFA: acronym for *Corpus de Inglês para Fins Acadêmicos*.

UFMG: acronym for *Universidade Federal de Minas Gerais*.

B1: language proficiency level following the Common European Framework of Reference for Language.

Ind.NE.Ess.2013-1: Independent non-edited essay collected in the first semester of 2013.

0026.0026: number given to each student's text and to each student.

4. This happens in all the world, independent of the religion. And the religion has a big influence in behavior and thoughts of religion people. **So**, my conclusion is that religion is not dying, but its level of importance has changed, in relation of the places and the people. (CorIsF-UFMG-B1.Ind.AEss.NE.out2014.0352.0142)

5. Our opinions about others can change our attitudes toward others as well. After all, we are never the same. **So** the first impression is not ALWAYS that is and will be what will determine one character and morality of a person. (CorIsF-UFMG-B1.Ind.AEss.NE.abr2015.0821.0402)

6. This option is not definitive. In fact, a lot of people change their choice. Some students change the course in the middle and sometimes a few people graduate and back to college to learn other profession. These are the dentists that intent to be engineers, lawyers that become musicians and doctos that leave their sthetoscopes to teach philosophy. **So** what we have to do to do not choose a wrong career? (CorIFA-UFMG-B1.Ind.NE.Ess.2013-1.0007.0007)

The overuse of *so* by the university Brazilian students, represented in this study by the participants who agreed to allow their essays to be part of CorIsF and CorIFA, shows that pedagogical intervention is at stake. The essays analyzed are part of the non-edited writings. Although there is no room in this article for presenting tasks created by our research group to address such matters, we need to acknowledge that this has been a concern. We believe that awareness of what is appropriate in different genres / registers is crucial to be developed in English classes, especially when the audience faces situations where they are evaluated and judged by their academic community (writing articles, presenting conference papers, etc.)

3.2 Therefore

The linking adverbial *therefore* is the least used in *CorIsF* and the second most used in *CorIFA*. According to Zihan (2014, p. 61), *therefore* can be assigned to three main category meanings in academic prose: logical consequence, result, and conclusion. If

we compare the meanings associated with the two other linking adverbials studied in this paper, (*so* and *thus*) we can infer that *therefore* is the one displaying fewer meanings regarding its use (ZIHAN, 2016). This may cause less problems to learners as it will be shown in this subsection.

Regarding syntactic position, the learner corpora demonstrated that the sentence-initial position is the preferred structure among learners, while native speakers seem to structure their arguments using *therefore* in medial position. Data from *CorIsF* show that 100% (Graph 1 and Table 2) of the occurrences of *therefore* in this corpus is in initial position. The number is slightly lower in *CorIFA:* 82% of the uses in the initial position with a timid number of occurrences in a medial position corresponding to 8% of the examples. The connector *therefore* is used in initial position when it links two independent clauses, and the writer wants to highlight the meanings of result or conclusion before the reader gets the second sentence. Sample 7 brings an example of *therefore* in initial position.

> 7. From these side, it is possible to infer that religion is as important and strong as it represented in the past. **Therefore**, there is a change of point of view. (CorIsF-UFMG-B1.Ind.AEss.NE.out2014.0348.0133)

When the native corpus is investigated a different trend emerges from that observed in the learner corpora. Results from LOCNESS indicate that 42% (Graph 1 and Table 2) of the occurrences are in the initial position, while 58% are used in the medial position. The use of connectors in a medial position is usually associated with complex discourse when the connector is placed in the middle of the verb phrase as illustrated in example eight.

> 8. Papers can be printed out neatly with whatever margins, font, and size of print the student wants to use. Computers have **therefore** made paper writing a much easier task. (LOCNESS - ICLE-US-MICH-0039.1)

Concentrating on the meanings that *therefore* can carry, logical consequence, result and conclusion were expressed by learners in both *CorIFA* and *CorIsF* and also by native speakers in LOCNESS. Writers draw on this cohesive device to establish a logical relationship between the ideas expressed in their argumentative essays. The following example illustrates the conclusion meaning expressed in CorIFA.

> 9. According to the program's official website, 21.418 scholarship has already implement, only 654 were intended for full doctorate, doctoral-3141 and 1940 to the sandwich postdoctoral (which is not and should not be a record of academic training). **Therefore**, only 26% of scholarship contribute to the formation of academic researchers accredited. (CorIFA-UFMG-B1.Ind.Ne.Ess.2013-1.0020.0020)

As mentioned at the beginning of this section, *therefore* does not offer a variety of meanings not imposing difficulties for students in its use. However, the predominance of *therefore* in the initial position of the sentence might be an indication that students need specific instruction in writing texts that belong to the academic domain. It seems that learners choose to organize their writing using *therefore* in a sentence-initial position as a strategy to make clear the meaning between two independent clauses.

3.3 Thus

As presented earlier, our quantitative analysis of the linking adverbial *thus* in Brazilian university students' essay writing demonstrated that there was no statistical significant difference between the corpora of study (CorIFA and CorIsF) and the reference corpus (LOCNESS). In this section we describe, in some details, the results from the qualitative analysis of *thus* to shed some light on the understanding of our students' necessities. Our overall findings revealed that there are certain similarities

and differences among the two groups in the use of *thus*, concerning the syntactic position and pragmatic functions.

Brazilian learners most commonly used *thus* in sentence-initial position, while native speakers had a preference for medial position. As shown by the data, 62% of the occurrences (Graph 1 and Table 2) of *thus* in LOCNESS is in medial position. These occurrences follow three main patterns: 1. *thus* + -ing form (*thus reducing, thus making, thus replacing*); 2. *thus* + phrase (*thus it can be seen, thus scientists cannot*); e 3. *thus* + verb (*and thus allows us, and thus bring*). It can be noticed that in construction type 3 (*thus + verb*) the connector is usually followed by *and*. In one of our study corpora, CorIsF, all occurrences with *thus* are of type 2, and there are no occurrences of *thus* followed by – *ing* form. The same analysis in CorIFA showed that 87.5% of the occurrences of *thus* are of type 2 (*thus* + phrase), and there were only two occurrences of *thus* + -ing form and one occurrence of *thus* + verb. These structural preferences will be more carefully analyzed in the following paragraphs.

Regarding *thus*, Zihan (2014, p.99-105) points out that there are three commonly used meanings in academic prose. (1) Logical consequence: a fact leads to a logical result as shown in the sentence below.

> The purpose of the analysis is to make comparisons over time and between groups and **thus** consistency of classification is more important than marginal classification decisions. (ZIHAN, 2014, p.104).

(2) Result: when *thus* is used to introduce a practical result.

> The third proposes changes in the Ranfurly Shield rules consequent upon the creation of the third division of the national provincial championship; and the fourth, which is to cause no excitement whatever, urges that North Otago, Otago and Southland be created as a fourth district within the jurisdiction of the Maori Advisory Board,

thus splitting Te Waipounamu (the South Island) into two. (ZIHAN, 2014, p.104).

(3) Signal of evidence: The connector is used to provide evidence in support of a previous claim, as it can be seen in another example from Zihan.

> The author recorded under a headword every occurrence of a word or construction which seemed to be quantifying in a particular context. **Thus**, 'He won a fistful of prizes in 1977' has the word 'fistful' recorded as the headword or 'type'. (ZIHAN, 2014, p.104).

In our reference corpus and study corpora *thus* was used similarly to mark result, logical consequence, and signal of evidence. The examples below show these three different meanings of *thus*. In sample 10, from our reference *corpus*, *thus* is used in medial position to introduce a practical result and a logical consequence (Sample 11 - *and thus*) respectively.

> 10. With people not eating beef it may have its side effects. People will decide to eat more lamb or pork **thus** increasing sales and demand in these areas. (LOCNESS - ICLE-ALEV-0012.9)
> 11. Apart from being a necessity to some, it makes the human being lazy to a certain extent **and thus**, allows us to use our brains even less and less. (LOCNESS - ICLE-ALEV-0013.9)

In example 12, from CorIsF, *thus* is used in medial position, similarly as in example 11, connecting clauses, a less frequent position in this corpus, to mark a logical consequence.

> 12. They believed that the God would give them the promised land and **thus** get the happiness (CorIsF - CorIsF-UFMG-B1.Ind.AEss.NE.ago2015.1098.0672)

In example 13, from CorIsF, **thus** is used in initial position to provide evidence in support of a previous argument.

> 13. These set of facts have contributed to people become more independent and count less on the divine blesses. **Thus**, it can explain reasonably why religion is loosing its power to answer and remedy the problems of people in contemporary world. (CorIsF - CorIsF-UFMG-B1.Ind.AEss.NE.out2014.0351.0141)

Zihan (2014) found that *thus* is the only linking adverbial used to mark a signal of evidence. It seems that there is a relation between the structures and the meanings expressed by *thus*. The structure *thus + -ing* form usually introduces a practical result. The structure *thus* + verb, commonly preceded by *and*, seems to mark a logical consequence. The structure *thus* + phrase seems to be used to provide evidence in support of a previous argument. A more detailed analysis of larger amount of data is necessary in order to confirm these impressions concerning the relation between structures and meanings. Further analysis could investigate, for example, the combination *and thus* + verb, and the combination *and thus* + phrase, as it seems that both can often mark logical consequence and this meaning could be related to the presence of *and*.

Even though there was no significant difference between the learner corpora and the reference corpus, we believe that a more detailed qualitative analysis can help us better understand the tendencies of usage assigned to *thus* by Brazilian university students. The small variety of structures with *thus*, few occurrences of *thus + -ing form* and of *thus* + verb, in the learner corpora may indicate the need for specific pedagogical interventions.

3.4 Focus on individual essays

The quantitative and qualitative analysis of the concordance lines demonstrated a high concentration of connectors in a single essay. In some cases *so* was used in initial position to introduce a new idea or as a marker of conclusion in several lines of the same essay (excerpt 14), as can be seen in the example below.

> 14. **So**, people started to express their religion in different ways. People began creating new religions, or started to modify the rituals and habits in old religions. The catolicism and protestantism are examples of religions that changed with the years. **So**, it appears that the people are losing their religiosity, but they are only changing their way to express it, to adequate a new world and a new society. But the faith is not lost. (CorIsF-UFMG-B1.Ind.AEss.NE.ago2015.1131.0705)

Many essays in our study corpora carry similar characteristics to the example presented above. Despite the fact that the linking adverbials are commonly used in the academic register, a large number of these cohesion devices within the same text can impair the reader´s ability to understand the text (CONRAD; BIBER, 2009). According to Field and Yip (1992) learners preference for using connectors in initial position in sentences and paragraphs may characterize the overuse of such connectors. In the authors' words "The impression of too many devices in the L2 scripts may be compounded by a strong use of the initial sentence and paragraph position" (FIELD; YIP, 1992, p. 25).

4. Conclusion

This study investigated the use of the linking adverbials *so*, *therefore* and *thus* in essays written by Brazilian university learners of English. We hypothesized that there would be an overuse of the connector *so* and an underuse of *therefore* and *thus* in the learner corpora (CorIsF and CorIFA) when compared to LOCNESS. The quantitative data derived from the log likelihood test confirmed the overuse of *so* and an underuse

of *therefore*. However, our initial underuse hypothesis of linking adverbial *thus* was invalid. Our study also explored and compared the meanings and syntactic position of the three linking adverbials in native and non-native corpora. The results revealed some differences between learners´ and native speakers´ texts.

The comparative analysis on the basis of syntactic position showed that Brazilian learners frequently used these linking adverbials in initial position, while native speakers tended to place them in medial position, which in turn can indicate a more complex discourse.

As for the meanings marked by the connectors under study, six different meanings or functions were identified altogether: result, logical consequence, signal of evidence, initiating a topic, conclusion, and restating an idea. In the native corpus the linking adverbial *so* carried the meanings of result and logical consequence. Differently, in the learner corpora *so* was used to initiate a topic, restate an idea and to mark a conclusion. According to Biber *et al.* (1999), *so* starting a topic or restating an idea is most frequent in spoken rather than in academic written registers. This use of *so* that was detected in CorIFA and CorIsF may indicate that learners seem to be unaware of the register restrictions of this linking adverbial, which is most frequently used in conversation to initiate a topic and it is not so appropriate to be used in the same position and with the same meaning in academic texts. Zihan (2014, p. 108) also states that written academic prose has far less use of *so*.

Linking adverbial *therefore* most commonly marked the meaning of result and logical consequence in the native speaker corpus, while in the learner corpora it marked the meaning of result and conclusion. *Thus* was used similarly in the learner and native corpora to provide evidence in support of a previous argument and present a practical outcome.

It is important to highlight that the qualitative analysis revealed that the three connectors were used to introduce a new idea or a conclusion. This could be noticed

in our corpora of study and reference corpus, although the incidence was clearly bigger in CorIFA and CorIsF.

Our study confirms the results from many other studies (ZIHAN, 2016; FIELD; YIP, 1992; GRANGER; TYSON, 1996) on the use of linking adverbials by learners, by indicating that students tend to overuse, underuse and misuse these linking adverbials in their written production. As stated by Biber *et al.* (1999), the use of linking adverbials is specific to genre and register. More studies need to be carried out in specific language learning contexts in order to better inform classroom priorities, syllabus, and teaching material design.

Future studies may choose to address pedagogical implications based on the description provided in this article. They could involve raising learners' awareness of the use of linking adverbials. This awareness can be important to help learners understand the nuances of use, regarding sentence position and meaning of the linking adverbial in academic texts. Studies such as the present one can serve as a guidance for second language learners on how/when/where each identified linking adverbial is used. As Zihan (2015) argued, studies on the use of linking adverbials should have more detailed discussions of their three aspects: form, meaning and position.

It is equally important to address the issue of register adequacy and the main differences between spoken and written register. Further corpus-based studies on the use of linking adverbials by Brazilian learners, such as Dutra et al. (2017) and Matte and Sarmento (2018), are needed as they can provide more evidence that may help the preparation of well-informed pedagogical material.

References

ANTHONY, L. **AntConc** (Version 3.4.4) [Computer Software]. Tokyo, Japan: Waseda University. Available from: http://www.laurenceanthony.net/ 2016.

ALTENBERG, B.; TAPPER, M. The use of adverbial connectors in advanced Swedish learners' written English. *In*: GRANGER, S., **Learner English on Computer**. Harlow: Addison Wesley Longman Limited., 1998, p. 80-93.

BIBER, D.; GRAY, B. **Grammatical complexity in academic English**: linguistic change in writing. Cambridge: Cambridge. 2016. DOI https://doi.org/10.1017/CBO9780511920776

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

CARTER, R.; McCARTHY, M. **Cambridge grammar of English**: a comprehensive guide: spoken and written English grammar and usage. New York: Cambridge University Press, 2006.

CHAMBERS, A. The learner corpus as a pedagogic corpus. *In*: GRANGER, S.; GILQUIN, G.; MEUNIER, F. **The Cambridge Handbook of Learner Corpus Research**. Cambridge: Cambridge, 2015, p. 445- 464. DOI https://doi.org/10.1017/CBO9781139649414.020

CHEN, C. W. The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. **International Journal of Corpus Linguistics**, v. 11, n.1, p.113-130, 2006. DOI https://doi.org/10.1075/ijcl.11.1.05che

CONRAD, S.; BIBER, D. **Real grammar**: a corpus-based approach to English. White Plains: Pearson/Longman. 2009.

COBB, T.; HORST, M. Learner corpora and lexis. *In*: GRANGER, S.; GILQUIN, G.; MEUNIER, F. **The Cambridge Handbook of Learner Corpus Research**. Cambridge: Cambridge, 2015. p. 185-206. DOI https://doi.org/10.1017/CBO9781139649414.009

COUNCIL OF EUROPE. COUNCIL FOR CULTURAL CO-OPERATION. EDUCATION COMMITTEE. MODERN LANGUAGES DIVISION. **Common European Framework of Reference for Languages**: learning, teaching, assessment. Cambridge University Press. 2001.

DUTRA, D. P.; BERBER SARDINHA, T. Referential expressions in English learner argumentative writing. *In*: GRANGER, S; GILQUIN, G; MEUNIER, F. **Twenty Years of Learner Corpus Research**: Looking back, Moving ahead. Corpora and Language in Use – Proceedings 1. Louvain-la-Neuve: Presses universitaires de Louvain, 2013, p. 117–127.

DUTRA, D. P.; QUEIROZ, J. M.; ALVES, J. C. Adding information in argumentative texts: a learner corpus-based study of additive linking adverbials. **Estudos Anglo Americanos**, v. 46, n. 1**,** p. 9-32, 2017.

DUTRA, D. P.; SOUZA, M. Learner corpora and classroom activities: discourse organizing markers on focus. *In*: CONGRESSO INTERNACIONAL DA ABRAPUI: LANGUAGE AND LITERATURE IN THE AGE OF TECHNOLOGY. **Comunicação**. Florianópolis: UFSC. 2012.

FIELD, Y.; YIP, L. M. O. A comparison of internal cohesive conjunction in the English essay writing of Cantonese speakers and native speakers of English. **RELC Journal**, v. 23, n. 1, p. 15-28, 1992. DOI https://doi.org/10.1177/003368829202300102

GILQUIN, G. From design to collection of learner corpora. *In*: GRANGER, S.; GILQUIN, G.; MEUNIER, F. (ed.), **The Cambridge Handbook of Learner Corpus Research** Cambridge: Cambridge University Press, 2015. p. 9-34. DOI https://doi.org/10.1017/CBO9781139649414.002

GRANGER, S. Contrastive interlanguage analysis A reappraisal. **International Journal of Learner Corpus Research**, v.1, n. 1, p. 7–24, 2015. DOI https://doi.org/10.1075/ijlcr.1.1.01gra

GRANGER, S.; GILQUIN, G.; MEUNIER, F. (ed.) **Twenty years of learner corpus research**. Louvain: UCL. 2013.

GRANGER, S. How to use foreign and second language learner corpora. *In*: MACKEY, A.; GASS, S. (ed.). **A Guide to Research Methods in Second Language Acquisition.** Malden: Basil Blackwell, 2012. p. 7–29. DOI https://doi.org/10.1002/9781444347340.ch2

GRANGER, S. (ed.) **Learner English on Computer**. London: Addison Wesley Longman, 1998.

GRANGER, S. From CA to CIA and back: an integrated approach to computerized bilingual and learner corpora. *In*: AIJMER, K.; ALTENBERG, B.; JOHNSSON, M. (ed.). **Languages in contrast. Text-based cross-linguistic studies.** Lund: Lund University Press, 1996. p. 37-51.

GRANGER, S., GILQUIN, G.; MEUNIER, F. **The Cambridge Handbook of Learner Corpus Research**. Cambridge: Cambridge University Press. 2015. DOI https://doi.org/10.1017/CBO9781139649414

GRANGER, S.; TYSON, S. Connector Usage in the English Essay Writing of Native and Non-Native EFL Speakers of English. **World Englishes**, v. 15, n. 1, p. 17–27, 1996. DOI https://doi.org/10.1111/j.1467-971X.1996.tb00089.x

GRIES, S. Statistics for learner corpus research. *In*: GRANGER, S.; GILQUIN, G.; MEUNIER, F. **The Cambridge Handbook of Learner Corpus Research**. Cambridge: Cambridge University Press, 2015. p. 159 - 181. DOI https://doi.org/10.1017/CBO9781139649414.008

HYLAND, K. **Academic Discourse English in a Global Context**. London: Continuum. 2009.

JOHNS, A.; BAWARSHI, A.; COE, R.; HYLAND, K. PALTRIDGE, B.; REIFF, M.; TARDY, C. Crossing the boundaries of genre studies: Commentaries by experts. **Journal of Second Language Writing**, v. 15, p. 234–249, 2006. DOI https://doi.org/10.1016/j.jslw.2006.09.001

MARQUES, A. L. S. P. **O papel da percepção consciente de conectivos concessivos na produção escrita e na reformulação da interlíngua**. 2007. Dissertação (Mestrado em Linguística Aplicada) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

MATTE, M. L.; SARMENTO, S. A corpus-based study of connectors in student academic writing. **English for Specific Purposes World**, v. 20, n.55, p. 1-21, 2018.

ORFANÓ, B. Analysing the use of vague language in spoken interlanguage: A corpus-based study of a group of Brazilian university students learning English as a second language. *In*: GRANGER, S; GILQUIN, G; MEUNIER, F. **Twenty Years of Learner Corpus Research**: Looking back, Moving ahead. Corpora and Language in Use – Proceedings 1. Louvain-la-Neuve: Presses universitaires de Louvain, 2013. p. 367-376.

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

RAYSON, P. **Testing for significance**: Log-likelihood. Available at: https://www.lancaster.ac.uk/fss/courses/ling/corpus/blue/l08_4.htm. Access on: July 29, 2018.

RAYSON, P. **Matrix**: A statistical method and software tool for linguistic analysis through corpus comparison. 2003. 207 f. Thesis (Ph.D. in Computer Science), Computing Department, Lancaster University, Lancaster. 2003.

RANKIN, T. Learner corpora and grammar. *In*: GRANGER, S.; GILQUIN, G.; MEUNIER, F. **The Cambridge Handbook of Learner Corpus Research.** Cambridge: Cambridge University Press, 2015, p. 231-254. DOI https://doi.org/10.1017/CBO9781139649414.011

RØRVIK, S; EGAN, T. Connectors in the argumentative writing of Norwegian novice writers. *In*: GRANGER, S; GILQUIN, G; MEUNIER, F. **Twenty years of learner corpus research**: Looking back, moving ahead. Louvain**:** Presses Universitaires de Louvain, 2013, p. 401-410.

SARMENTO, S. ABREU-E-LIMA, D. M.; MORAES FILHO, W. B. (Org.) **Do Inglês sem Fronteiras ao Idioma sem Fronteiras**: A construção de uma política linguística para a internacionalização. Belo Horizonte: Editora UFMG, 2016.

SEIDLHOFER, B. Closing a conceptual gap: The case for a description of English as a lingua franca. **International journal of applied linguistics**, v. 11, n. 2, p. 133-158, 2001. DOI https://doi.org/10.1111/1473-4192.00011

SWALES, J.; FEAK, C. **Academic writing for graduate students**. 3rd Edition. Ann Arbor: MUP, 2012. DOI https://doi.org/10.3998/mpub.2173936

SWALES, J.; FEAK, C. **Navigating academia**: writing supporting genres. Ann Arbor: MUP, 2011. DOI https://doi.org/10.3998/mpub.3073304

SWALES, J. M. **Genre Analysis**: English in academic and research settings. Cambridge: University Press, 1990.

ZIHAN, Y. Register-specific meaning categorization of linking adverbials in English. **Journal of English for Academic Purposes**, v. 22, p. 1-18, 2016. DOI https://doi.org/10.1016/j.jeap.2016.01.004

ZIHAN, Y. The Use of Cohesive Devices in News Language: Overuse, Underuse or Misuse? **RELC Journal**, v. 46, n. 3, p. 309–326, 2015. DOI https://doi.org/10.1177/0033688215597578

ZIHAN, Y. **Linking Adverbials in English**. 2014. 190 f. Thesis (Ph.D. in Applied Linguistics) – Victoria University of Wellington, New Zealand, 2014.

APPENDIX A

Consent Form

CARTA DE CONSENTIMENTO LIVRE E ESCLARECIDO:
Para os participantes (alunos da graduação)

<u>Caro(a) Senhor(a)</u>

A coordenação do Programa Inglês sem Fronteiras/UFMG conduz pesquisas que visam estudar o desenvolvimento das habilidades de leitura, de escrita, de audição e de fala de aprendizes de língua inglesa para fins acadêmicos. Cada projeto de pesquisa está devidamente autorizado pela Câmara de Pesquisa da Faculdade de Letras da UFMG.

A fim de que os projetos possam ser desenvolvidos, é necessária a sua autorização, vez que as pesquisas constarão da coleta das suas redações produzidas enquanto aluno do curso. A sua participação nesta pesquisa é voluntária e não determinará qualquer risco nem trará desconfortos. Além disso, sua participação é importante para o aumento do conhecimento a respeito dos processos de aquisição e desenvolvimento das quatro habilidades supracitadas por alunos universitários brasileiros, podendo beneficiar outros alunos futuramente na melhoria do ensino de língua inglesa no nível superior.

Informamos que o/a Sr(a). tem a garantia de acesso, em qualquer etapa dos estudos, sobre qualquer esclarecimento de eventuais dúvidas. Se tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com a coordenação do programa (3409-3839) ou com o Comitê de Ética em Pesquisa (CoEP) da Universidade Federal de Minas Gerais, situado na Av. Antônio Carlos, 6627. Unidade Administrativa II - 2° andar - Campus Pampulha, telefone 3409-4592 / 3409-4027.

Também é garantida a liberdade da retirada de **consentimento** a qualquer momento e deixar de participar do estudo.

Fica também garantido que as informações obtidas serão analisadas em conjunto com as de outras pessoas, não sendo divulgada a identificação de nenhum dos participantes.

O/A Sr(a). tem o direito de ser mantido atualizado sobre os resultados parciais das pesquisas e, caso seja solicitado, todas as informações que solicitar lhe serão fornecidas.

Não existirão despesas ou compensações pessoais para o participante em qualquer fase dos estudos. Também não há compensação financeira relacionada à sua participação.

Os participantes das pesquisas comprometem-se a utilizar os dados coletados somente para pesquisa, e os resultados serão veiculados através de artigos científicos, em revistas especializadas e/ou em encontros científicos e congressos, sem nunca tornar possível a sua identificação.

Abaixo se encontra o Termo de **Consentimento Livre e Esclarecido**, para concordância caso não tenha ficado qualquer dúvida.

Deise Prina Dutra – Coordenadora Geral do IsF/UFMG
Ana Larissa Adorno Marciotto Oliveira – Coordenadora Pedagógica do IsF/UFMG

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Acredito ter sido suficientemente informado a respeito dos estudos conduzidos pela coordenação do Programa Inglês sem Fronteiras/XXXX. Ficaram claros para mim quais são os propósitos dos estudos, os procedimentos a serem realizados, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro, também, que a minha participação é isenta de despesas e que tenho garantia do acesso aos resultados e de esclarecer minhas dúvidas a qualquer tempo. Concordo voluntariamente em participar e estou ciente de que poderei retirar o meu consentimento a qualquer momento sem penalidade ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido.

○ Concordo
○ Discordo

APPENDIX B

Task example

INDEPENDENT WRITING TASK

Thought and mind (B1)

Read the question below. Give yourself 30 minutes to plan, write, and revise your essay. Typically, an effective response will contain a minimum of 300 words.
- Do you agree or disagree with the following statement? The first impression is the most important one. Use specific reasons and examples to support your opinion.

Religion (B1)

Read the question below. Give yourself 30 minutes to plan, write, and revise your essay. Typically, an effective response will contain a minimum of 300 words.
- It appears that religion has been around in one form or another for most of human history. Do you think it is getting stronger, dying out, or staying about the same level of importance. Why?

Facebook

Read the question below. Give yourself 30 to 40 minutes to plan, write, and revise your essay. The essay should be about 200 or 300 words.
- Should what you say on Facebook be grounds for getting fired?

APPENDIX C

Themes from Br-ICLE also used in CorIFA

(1) Crime does not pay
(2) The prison system is outdated. No civilised society should punish its criminals: it should rehabilitate them
(3) Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value
(4) A man/woman's financial reward should be commensurate with their contribution to the society they live in.
(5) The role of censorship in Western society

(6) Marx once said that religion was the opium of the masses. If he was alive at the end of the 20th century, he would replace religion with television

(7) All armies should consist entirely of professional soldiers: there is no value in a system of military service

(8) The Gulf War has shown us that it is still a great thing to fight for one's country

(9) Feminists have done more harm to the cause of women than good.

(10) In his novel Animal Farm, George Orwell wrote "All men are equal: but some are more equal than others" How true is this today?

(11) In the words of the old song "Money is the root of all evil"

(12) In the 19th century, Victor Hugo said : "How sad it is to think that nature is calling out but humanity refuses to pay heed. "Do you think it is still true nowadays ?

(13) Some people say that in our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion?