

CADS innovative Corpus-driven Methodology for the Academic English Instruction of Transitional Linkers.

M. Alesón-Carbonell; A. Ryneiskaya

Departamento de Filología Inglesa
Universidad de Alicante

ABSTRACT

The analysis and description of discourse based on corpora (CADS - Corpus-Assisted Discourse Studies) has experienced an important development during the last decade, essentially due to two main factors: (1) on the one hand the existence of new and more powerful corpus-analysis software and tools for language processing; (2) and, on the other, the realization that corpus methods can extend their application to larger fields of interest. In fact, with regards to L2 teaching, this methodology has promoted the design of specific tools, such as reference corpora, which, in turn, have shaped the ground for the description of, for instance, Academic English discursive conventions. Particularly, the knowledge of discursive and genre practices is paramount for university undergraduates who are, at the same time, acquiring proficiency in L2 and initiating themselves as academic professionals. For that reason, this study attempts to focus on the academic essay and tries to describe the differences in uses of Academic English between Native (corpus BAWE) and non-native speakers (corpus of Spanish speakers-NNS). Likewise, the paper raises issues concerning NNS proficiency assessment and, finally, intends to provide new pedagogical strategies that may facilitate the process of writing skills acquisition in Academic English. Ultimately, the data and results seem to corroborate the adequacy of a CADS approach in order to establish innovative pedagogical procedures in teaching L2 discursive competence.

Palabras clave: Learner Corpus, Genre Analysis, Transitional Linkers, Academic English
1. INTRODUCTION

Academic writing among undergraduate students has been the focus of many recent studies that embrace a wide range of aspects involved in the acquisition of writing proficiency (Hyland, 2007; Swales, 1990a; Swales & Leeder, 2012a; Paltridge et al., 2009). An increasing scholar interest in how writing skills are acquired and how best to teach university-level writing skills has given rise to genre-based studies that embrace the approach of corpus-linguistics methodology. The term Corpus-Assisted Discourse Studies (CADS) has been recently coined to refer to a new discipline in the study of genres that uses the resources of corpus methodology and computers to analyze features of discourse (Partington, 2011; Baker, Gabrielatos, KhoraviNik, McEnery, & Wodak, 2008). Although there has been always an inherent interest on approaching the analysis of discourse features in Corpus Linguistics (Biber, Conrad, & Reppen, 1998: 106-132), this approach differs from the typical analysis of lexical and grammatical associations and patterns. In the field of Academic Writing the association of these two disciplines in the CADS approach has proved very productive as it provides the opportunity of analyzing large quantity of data and providing more accurate descriptions of actual and authentic language use (Biber et al., 1998; Chang & Kuo, 2011; Gabrielatos, 2012).

1.1 Corpus and Genre Studies

Indeed, corpus-studies are nowadays prolific in the field of Academic Writing, especially in those pedagogical approaches aimed at improving non-native speakers (NNS) writing. For NNS, mastering academic genres is essential in order to be able to access the academic community. As most scholars point out, the majority of the research done in Academic Writing has been addressed to the analysis and description of Academic English in order to set up pedagogical tools that would enable native and non-native speakers to improve their writing performance, being Swales’ CARS model a seminal study in the area (Swales, 1990). Obviously, being English the current Lingua Franca of both academic research and publishing, most of the research conducted has been centred either on English academic genres, or in contrasting and comparing native and non-native production in English. Genre Analysis as defined by Swales (1990) provides, through both qualitative and quantitative methodology, a comprehensive description of a specific genre taking into account both detailed macrostructure and microstructure description of discourse features. Also, one of the main developments
proposed by the field is to relate language description with a comprehensive insight of the communicative situations in which the genre is used (Hyland, 2007; Swales & Leeder, 2012; Swales & Feak, 1994; Swales, 2005; Swales, 1990; Eggins & Martin, 1998). In that line, linguistic analysis departs from a mere descriptive approach to evaluate language use and variation in context. As Vijak Bhatia points out Genre Theory tries to offer a ‘grounded description of language” (2004: 22) that not only highlights common patterns and structures, but that leaves back the conception of genres as simply “textual artifacts” to understand discourse as an inherent part of social and professional practices (2008: 62). Therefore, knowledge of discursive and genre practices is essential in order to design appropriate pedagogical materials for students that are acquiring proficiency in Academic English and are trying to become future academics; in the case of NNS, who have, on the one hand, to acquire competence in L2 and, on the other, to be initiated as academic professionals, this demand is paramount.

With this aim, some contrastive studies of learner language have already been carried out comparing the use of English in academic settings between native speakers and speakers of other languages, i.e. Japanese (Suzuki, Fukushima, Kinjo, & Yoshihara, 2012), Chinese (Desmond, 1995), Cantonese (Hyland & Milton, 1997), French (Carter-Thomas & Chambers, 2012; Freake, Gentil, & Sheyholislami, 2011), German (Lorenz, 1999), or Spanish (Hidalgo Avilés, 2007). Nevertheless, very few studies have been centred on analyzing the Academic English of Spanish learners from a corpus–based approach; one of the exceptions being Moreno Jaén (2007) who uses a corpus approach to testing. Yet, we have observed that in general terms academic genres written by Spanish students remain apparently distanced from native likeness, despite grammar accuracy, so there is a need to further explore the differences in pattern and structure in this type of discourses. As Lorenz (1999) highlights, this is a “sensitive area (ie. difficulty) of foreign language use” that demands special attention.

Therefore, there is a clear need for research in the field, above all of studies that try to advance in the knowledge of typical discursive patterns employed by Spanish speakers in their Academic English writing that are different to typical native English use. This approach is underpinned on the assertion that learners’ language is a “separate linguistic variety” whose ‘variation’ or difference could be studied through the comparative study of the frequency of their linguistic patterns opposed to native-language use. Using learner corpora as Lorenz (1999: 3) states “[…] entails reversing the logic of linguistic description” as it gears research towards the study of production
and variation, rather than towards the traditional vision of teaching grammar rules and vocabulary. Lorenz (1999: 9-12) proposes five principles for learner corpus compilation and research:

1. The strength of learner corpus lies in the detection of patterns, not errors.
2. ‘Naturalness’ and ‘idiomacity’ are probabilistic concepts.
3. Learner corpus studies are best directed at the ‘advanced’ level.
4. The characteristic patterns of advanced learner language are best studied in written production.
5. Naturalness can be improved by making convert patterns explicit.

This five principles constitute the basic rationale for the compilation of a learner corpus prior to a comparative study, and also foresee the practical application of this type of approach. Indeed, corpora of learners’ language are decisive in “discovering learner’s deviations that would otherwise go unnoticed” (1999:8) as they are able to unveil repeated patterns in language that ‘deviate’ from native uses, but which could either correspond to learners’ strategies or to the interference of native language. The objective is then to improve NNS literacy in Academic English by making explicit deviance, so that learners and teachers are aware of these patterns and can work on more natural linguistic behaviour (Lorenz 1999:12):

“It is assumed, here that, once patterns of non-native deviance have been discovered, students can be explicitly made aware of these patterns, and that, given time, motivation and the opportunity to practice, they will eventually be able to modify their linguistic behavior into a more native-like direction.”

1.2 Argumentative Genres and Features of Analysis

From all the genres that could be the target of a CADS analysis in academic settings, Hyland (2009: 40) points out that “argumentative genres play a major role in developing students’ academic knowledge and socializing them into legitimized, and therefore powerful, social practices”. He also adds that “non-native English speaking students appear to find essay demands particularly daunting” (Henry & Roseberry, 1997: 479). An argumentative essay, regardless its extension, is a genre that is especially problematic for its complexity, as its communicative purpose is “putting forward a point of view on a particular issue or topic and defending or explaining it”. In addition, different strategies such as definition, exemplification or persuasion are inherent to the text and there is involved a complex process of reasoning, arguing and persuading. Ching-Fen Chang and Hua Kuo (2011: 223) in their corpus-based study on research articles writing also agree for the mastering of writing competence in essay writing is as important for future professionals, as it is their knowledge on their specific field.
Academic essays and, specifically, research articles have been a major focus of attention of both Genre Analysis and Corpus-based research (see for instance Arrington & Rose, 1987; Crompton 1997; Henry & Roseberry 1997; Salanger-Meyer 1994; Swales 1990; or Swales and Feak 1994); especially in relation to macrostructure and moves, hedging, evaluative language, or intensifiers. Conversely, less attention has been drawn to the analysis of specific discursive features such as connectors or connective structures, when indeed the use of connectors and transitional words in particular, appears to be a problematic point in argumentative writing (Holten, 1998).

Alfred Rosa and Paul Escholz (1995) define transition words as the “words and phrases that are used to signal the relationships between ideas in an essay and to join the various parts of the essay together”. One of the main reasons why connectors and transitional linkers are crucial in argumentative writing is that they present a cohesive function as “bridges” to link ideas (Holten, 1998). As Jan Renkema (2009) points out in his Theory of Connectivity, word connectors are one of the constituents of the texture of discourse in the creation and maintenance of what he calls relational cohesion. Also, Hyland (2007:199) highlights their importance as elements of cohesion in the text. Therefore, we consider that more research on transitional linkers is necessary, in terms of frequency and colligation in order to improve NNS students’ argumentative competence.

1.3 Research questions.

In sum, taking into account the advances done in the area of academic writing the present paper aims to offer a contrastive study of native and non-native use of transitional linking adverbials in academic essays that answers the following research questions:

1. Which transitional linkers are specifically typical of academic argumentative writing?
2. Are academic transitional linkers more frequent in NS or in NNS writing?
3. Are there differences in the use of specific transitional linkers as regards frequency in NS between NNS writing?
4. Are specific academic linkers over/under used in the NNS corpus?

2. METHODOLOGY

2.1. Description of the corpora and selection principles

As the main purpose of the study is to analyze students’ performance and use of the academic transitional adverbs to determine over or underuse of these structures, the
main source of data will be from learner corpora. However, in order to be able to evaluate the use of the structures, the study is also adding insights from control corpora from professional academic writers. Indeed, following Gabrielatos (2005), there are usually three types of corpora that can be used for learning and pedagogical purposes, namely *native speaker corpora, learner corpora* and corpora extracted from *text and course-books* (see Figure 1: Corpora and ELT (Gabrielatos, 2005). Figure 1).

![Figure 1: Corpora and ELT (Gabrielatos, 2005).](image)

As regards the native and non-native speakers’ data sources, this study consisted of two learner corpora, namely the BAWE corpus and an ad-hoc Spanish students’ corpus (SpL1-EW). The native speakers’ corpus BAWE (The British Academic Written English) was obtained through the free downloadable files available from the Oxford Text Archive (http://ota.ahds.ac.uk/). It was developed at the Universities of Oxford Brookes, Reading and Warwick under the directorship of Hilary Nesi, Sheena Gardner, Paul Thompson and Paul Wickens with funding from the ESRC (Nesi et al., 2010). It consists of 2761 assignments covering 35 disciplines. However, for our project, only fifty-one papers were selected due to their representativeness, which was based on the criteria of appropriateness of the essay-genre employed and their comparability as regards the academic disciplines involved, namely *Linguistics, English, Sociology, Comparative American Studies, and Classics*. The number of native-speakers papers chosen for our study is lower than the Spanish for the sake of maintaining a balanced number of running words in both corpora, as Spanish essays tend to be shorter in length. In fact, as showed in Table 1, there is a mean of 1,311.58 words per essay in BAWE, whereas SpL1–EW has a mean of 592.82 words per paper. Additionally, author-related information provided along with the corpus was taken into consideration while
delimiting the most appropriate papers; for instance, having attended entirely the UK high school system seemed an important factor to us as it “guaranteed” the “Englishness” of participants.

<table>
<thead>
<tr>
<th></th>
<th>BAWE</th>
<th>SpL1-EW</th>
</tr>
</thead>
<tbody>
<tr>
<td># Tokens</td>
<td>66,891</td>
<td>58,689</td>
</tr>
<tr>
<td># Essays</td>
<td>51</td>
<td>99</td>
</tr>
<tr>
<td>Mean</td>
<td>1,311.58</td>
<td>592.81</td>
</tr>
</tbody>
</table>

Table 1. Means of Tokens per Essay.

The non-native learner corpus was compiled only among third and fourth-year students of the English Studies degree at the University of Alicante, contrary to our BAWE selection that comprises students from all undergraduate courses, this decision was taken under the premise, stated by Lorenz (1999:10) that learner corpora compilation with the aim of eliciting differences between NNS and NS language patterns, must be focused on advanced learners. Thus, the SpL1-EW corpus brings together 99 essays covering two disciplines: English Language and Textual Analysis of English Literary Works (both being core subjects in the current curriculum of the English Studies degree). Table 2 provides a basic description of the two corpora.

<table>
<thead>
<tr>
<th>Academic Disciplines</th>
<th>Tokens</th>
<th># Files</th>
<th>Academic Disciplines</th>
<th># Tokens</th>
<th># Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>25,994</td>
<td>22</td>
<td>English Language</td>
<td>18,883</td>
<td>55</td>
</tr>
<tr>
<td>Classics</td>
<td>18,794</td>
<td>13</td>
<td>Textual Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociology</td>
<td>13,802</td>
<td>10</td>
<td>of Literary works</td>
<td>39,805</td>
<td>44</td>
</tr>
<tr>
<td>Linguistics</td>
<td>4127</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp. American Studies</td>
<td>3068</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>65,785</td>
<td>51</td>
<td>Overall</td>
<td>58,688</td>
<td>99</td>
</tr>
</tbody>
</table>

Table 2. Detailed Description of the Corpora Compilation.

2.2. Delimiting transitional linkers

In relation with the inclusion of reference corpora of course-books, as proposed by Gabrielatos (2005), it was important to use a corpus that could be used to evaluate the academic language used by the learners in the above mentioned corpora. The problem was that the description of connectors in course and grammar books is not always consistent; as A. Downing & P. Locke (2002) posit the typical description of these groups in grammar books is a simple classification under semantic labels (cf.
Holten, 1998; Quirk & Greenbaum, 1979). Besides, the focus is, obviously, not restricted to Academic Language only. Hence, along the lines of the academic scope of our research, it was important to limit the list of connectors that was going to be the object of study to those that were just typically academic. To address this problem, the Academic Wordlist (AWL) (Coxhead, 2010) was considered an appropriate reference. The AWL was developed using a written academic corpus of 3.5 million running words. It contained textbooks, articles and laboratory manuals subsequently divided in four major areas: Arts, Commerce, Law and Science. The three criteria for the selection of headwords in AWL were frequency, range and uniformity. As Coxhead (2011) explains “the word families had to occur 100 times or more in each of the four disciplines (frequency), in 15 or more of the subject areas (range), and over 10 times in the four disciplines (uniformity)”. A total of 570 families met the criteria. The 60 most frequent families, each containing a headword and the sublist members, form the first sublist. The second sublist comprises the next 60 more frequent word families, and so on. Overall, as Coxhead reports, the AWL covers 10% of the vocabulary in the written academic corpus it was based on.

It is therefore possible to examine which connectors appear among the 10 sublists and whether their usage by Native speakers in BAWE is as frequent as in SpL1-EW. As a matter of fact, numerous studies, although not focusing exclusively on transitions, had referred to the AWL and profiled the presence of the AWL items in differently-oriented corpora (Belcher, Johns, & Paltridge, 2011; Chen & Ge, 2007; Li & Qian, 2010; Suzuki, Fukushima, Kinjo, & Yoshihara, 2012). In the first step of the study all the AWL ten sublists were examined. Overall, 13 transitional linkers were identified and sampled according to the sublist in which they occur:

Sublist 4: *Despite, Hence, Overall, Prior.*

Sublist 5: *Whereas.*

Sublist 6: *Furthermore, Nevertheless.*

Sublist 8: *Plus, Thereby.*

Sublist 10: *Albeit, Likewise, Nonetheless, Notwithstanding.*

One of the most recurrent pieces of criticism that Coxhead’s AWL has received is to have left out of the list West’s list of most frequent words. This decision was taken because the purpose of the AWL was not General English but specifically academic vocabulary. Li & Qian (2010), for instance, in their revision of previous literature points out that the exclusion of general vocabulary is in a way unnatural and can mislead the
learner as general language words are also employed in Academic Writing. Thus, attending to this reported deficiency and assuming that either native or non-native speakers will have a better command of common linkers, i.e. not strictly academic, the list of most frequent connectors (MICUSP) listed in J. Swales & K. Lockman (2010) was also included:

*In addition, However, Therefore, As a result, i.e., In contrast, On the other hand, For example, For Instance and In other words.*

2.3. Tools and procedure

Once the use of the reference corpora had delimited the list of headwords that had to be analysed and compared, the analysis was carried out using both quantitative and qualitative methods. Firstly, statistical and frequency counts were made in both corpora, followed by a qualitative analysis of the concordance, using WordSmith Tools (v.5).

3. RESULTS

With respect to the first and second research questions about specifically academic transitions and its frequency in our corpora, the analyses yielded instances of academic transitional linkers in both corpora:

<table>
<thead>
<tr>
<th>AWL Adverbs</th>
<th>BAWE instances</th>
<th>in SpL1 - EW instances</th>
<th>Awl sublist Category*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furthermore</td>
<td>7</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>Nevertheless</td>
<td>3</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Despite</td>
<td>11</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Whereas</td>
<td>16</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Hence</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Likewise</td>
<td>1</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Plus</td>
<td>3 (only 2 as connectors)</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Nonetheless</td>
<td>4</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Overall</td>
<td>7 (only 4 as connectors)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Albeit</td>
<td>1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Notwithstanding</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Thereby</td>
<td>2</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Prior</td>
<td>4 (only 2 as connectors)</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3. Instances of occurrence in BAWE and SpL1 - EW of 13 adverbs detected in AWL.

In addition, general language connectors were also common and their frequency higher than the strict academic connectors (see Table 3).
In general terms, when comparing the means of frequency of connectors, the analyses of both corpora revealed that from a global perspective there is not a significative difference (p= 0.449) between the BAWE and the SpL1-EW corpora.

<table>
<thead>
<tr>
<th>Source</th>
<th>Transition</th>
<th>Standardized figures (SF)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BAWE</td>
<td>SpL1-EW</td>
</tr>
<tr>
<td>MICUSP</td>
<td>Furthermore</td>
<td>9,23</td>
</tr>
<tr>
<td></td>
<td>Nevertheless</td>
<td>3,95</td>
</tr>
<tr>
<td></td>
<td>In addition</td>
<td>13,19</td>
</tr>
<tr>
<td></td>
<td>On the other hand</td>
<td>15,83</td>
</tr>
<tr>
<td></td>
<td>In other words</td>
<td>0</td>
</tr>
<tr>
<td>AWL</td>
<td>Overall</td>
<td>5,27</td>
</tr>
<tr>
<td></td>
<td>Thereby</td>
<td>2,63</td>
</tr>
<tr>
<td></td>
<td>Albeit</td>
<td>1,31</td>
</tr>
</tbody>
</table>

Table 6. Comparative analysis results of transitions frequency over 100.000 words

When considering the standardized figures, five connectors in particular, all of them listed by Swales (2010), emerged with a remarkable overuse by Spanish users. On
the contrary, only three AWL connectors came forward with a noteworthy result of underuse with a total absence in SpL1–EW papers. Due to the low frequency of these connectors also in the BAWE corpus, no further examination, at least in the present study, was conducted, although it would be desirable to focus on these type of connectors in wider scope research.

These results confirm previous results in literature that show that students, both NS and NNS, tend to be more conservative in their use of language as are still in their process of argumentative skill acquisition. This is also confirmed if we take a close look to the most frequently used connectors in both corpora that are *And* and *But*. Results tend to confirm the students’ mastering of general non-academic connectors over more specific academic linkers.

<table>
<thead>
<tr>
<th>Connector</th>
<th>BAWE instances</th>
<th>SpL1 - EW instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>and</td>
<td>2041</td>
<td>1586</td>
</tr>
<tr>
<td>but</td>
<td>190</td>
<td>286</td>
</tr>
</tbody>
</table>

*Table 7. And and But frequency results.*

**Furthermore**

This formal transitional linker or discourse marker is “used to add information or arguments to what has already been said” (Swan, 2000). In comparison with the other rather overused transitions, the *furthermore* case turned out to be the most relevant. In fact, as the raw figures and the standardized calculus show, there is almost a six-time difference in its usage between BAWE and SpL1–EW frequencies (see Table 8).

<table>
<thead>
<tr>
<th>Type</th>
<th>BAWE</th>
<th>SpL1 - EW</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>9.23</td>
<td>56.45</td>
</tr>
<tr>
<td>RAW figures</td>
<td>7</td>
<td>33</td>
</tr>
</tbody>
</table>

*Table 8. Furthermore ratio*

**Examples:**

The intervention of the Olympian deities in Greek myth is a recurring motif, and in many myths they play pivotal roles. **Furthermore**, the idea of virginity is a major theme in a fair number of ancient Greek myth ([BAWE: 6179a](#)).

The fact that the child is looking down and not up at the camera, apparently concentrating on using the prosthetic arms, **furthermore** creates a sense of intrusion and voyeurism ([BAWE: 0318d](#)).

To sum up, it is important to put emphasis on the idea that it is an example of narrative discourse that combines action, speech and thoughts. **Furthermore**, it shows a clear progressivity marked not only by the […] ([SpL1-EW: OT-student02](#)).

As we can see, concerning its position within a sentence, the majority of Spanish students followed the same pattern of placing this connector right after a full stop, with
an exception of only one instance. However, almost a third of the examples provided by Native speakers, reveals its usage in the middle position of a sentence.

**Nevertheless**

It is an ambivalent marker since it can be used both: to accentuate a contrast, and to introduce a concession. When performing the first function, it “emphasizes the fact that the second argument contrasts with the first”, but, unlike *however*, “*nevertheless* is very formal”. Following Swan (2000), *nevertheless* can also perform a concessive function in which case it is used in a “three-part argumentative structure: (1) the writer mentions facts pointing at a certain direction; (2) it is agreed that particular contradictory facts point the other way; (3) but the writer dismisses this and returns to the original direction of argument.”

<table>
<thead>
<tr>
<th></th>
<th>BAWE</th>
<th>SpL1 - EW</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>3.95</td>
<td>22.23</td>
</tr>
<tr>
<td>RAW figures</td>
<td>3</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 9. *Nevertheless* ratio

As data shows, NNS use of *Nevertheless* is frequent 5 times more than the NS use.

[…](SpL1–EW: tiz.txt) Spanish new government promises. **Nevertheless**, the credibility of national governments has been called into question since the beginning of this crisis. (SpL1–EW: tiz.txt) Braverman envisaged being consumed by deskilling, but *nevertheless* skill was taken from the ‘shop floor’ labourers into the ‘planning department’ (BAWE: 6195a.txt) Silver mining in Mexico *nevertheless* stimulated the commercial growth of the colony, establishing itself as the central part in which all economic activity came to be related. (BAWE: 0130a.txt)

Moreover, all Spanish instances are similar in their use of *nevertheless* after a full stop, whereas two of the Native samples attest a different, middle, position. In addition, if we analyse carefully the concordances, we find that although the frequency of use is higher in the SpL1-EW corpus, the argumentative structure highlighted by Swan is nearly never followed. In the majority of instances, Spanish students use *Nevertheless* as a synonym of *However* or just as a resource for adding information, thus showing their insufficient mastering of complex argumentative structures.

*Nevertheless*, there exist a huge amount of factors involved in the process of preparing an exam which are often overlooked, but they are as significant as the test itself. Those factors are even more relevant […](SpL1–EW: student85.txt)

**In addition**

This connector has the same additive function as *furthermore* (Swan, 2000).
Table 10. *In addition* ratio

<table>
<thead>
<tr>
<th></th>
<th>BAWE</th>
<th>SpL1- EW</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>13.19</td>
<td>70.14</td>
</tr>
<tr>
<td>Raw figures</td>
<td>10</td>
<td>41</td>
</tr>
</tbody>
</table>

Once again, the ratio of NNS overuse is remarkable.

[...] the gulf between himself and his mistress reveals a lack of commitment; *in addition*, he seems to make [...] (BAWE:0179).txt

[...] student risks everything on one throw and, *in addition*, the result obtained could be an unfair [...] (SpL1-EW:10student09).txt

[...] something related to the English language. *In addition*, this kind of studies is a good opportunity to travel [...] (SpL1–EW:beg).txt

As regards positioning, in this particular case there are no striking differences between the corpora. All Spanish students except one and all Native students except two used this transition in its initial position.

*On the other hand*

The main function of this connector is “to balance two factors or ideas that contrast, but do not contradict each other” (Swan, 2000).

Table 11. *On the other hand* ratio

<table>
<thead>
<tr>
<th></th>
<th>BAWE</th>
<th>SpL1- EW</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>15.83</td>
<td>68.42</td>
</tr>
<tr>
<td>Raw figures</td>
<td>12</td>
<td>40</td>
</tr>
</tbody>
</table>

For written language *on the other hand*, the answer is the opposite. (BAWE: 3080b).txt

[...] he is about to commit the irreparable; and *on the other hand* she becomes a prey for a predator pushed by revenge. (SpL1 – EW10student07).txt

Such predicts are generally optimistic in timescale. *On the other hand*, the development of powerful and highly intelligent robots can take over or destroy the human race. (SpL1–EWire).txt

Here, the NNS use exceeds in four times the native results. The Spanish tendency to place connectors at the beginning of a sentence is newly verified, since only one NNS student used it in a middle position as opposed to a half of Native students.

*In other words*

This linker marks the intention of the speaker or writer “to say something again in another way and, accordingly, to make things clear or give details” (Swan, 2000).
Table 12. In other words data

<table>
<thead>
<tr>
<th></th>
<th>BAWE</th>
<th>SpL1- EW</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>0</td>
<td>11.98</td>
</tr>
<tr>
<td>Raw figures</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

[...only associated to other notions like space, matter, motion, events, in other words with aspects closely related to the existence of human being, centred in the middle of the World as each of us is able to understand it. (SpL1–EW: 10student16.txt)
We notice that she is religious and attaches herself to God's help, in other words she still has faith.(SpL1–EW: germ.txt)
First of all, we must highlight the importance of isosemy, especially the semantic fields of force and violence, and how it helps, together with syntax, to create an increasing tension and violence throughout the passage; in other words, its progressivity.(SpL1–EW: OT_student17.txt)

In this particular case, as figures prove, there were no instances of this transitional linker in the BAWE corpus. As for SpL1–W corpus, half of instances were located in a middle position (some of them provided above).

4. CONCLUSION

This study has suggested that undergraduate L2 students have special difficulty in using specific academic transitional markers in argumentative discourse. Our study is an attempt to focus on real texts and explore the use of particularly outstanding connectors by the Spanish learners in terms of their incidence and situation in a sentence. Although not many contributions in this field to date have focused on the use of connectors in learner corpora, we believe that by identifying writing differences and deviations between NNS and NS use, precise pedagogical methods can be developed. However, wider-scope studies and more extensive corpora could shed more light on these issues.

As regards to specific results, similarly to the other corpora–based studies (Lorenz, 1999; Suzuki et al. 2012), frequency findings have proved to be an adequate method to elicit differences between NNS and NS speakers. In the case of specific academic connectors, data has showed that there are no overall differences on the subject of occurrence and distribution of transitions between NNS and NS. However, certain transitional linkers proved to be deviant in the way they are used by SpL1–EW. As a matter of fact, and as hypothesized above, Spanish learners tend to resort to a relatively narrower number of transitional linkers, which leads to their prominent overuse. Additionally, as for the intra-sentential location of connectors, there is a prevailing bias towards initial positions. Presumably, NNS students are more
conservative and try to stick to basic syntactic order without venturing into more creative structures; in this respect further investigation would be advisable.

Although the results found in this study are not totally conclusive, some of the issues raised have highlighted certain needs in the Spanish speakers’ mastering of English connectors, and also have established ground for pedagogical actions to be designed and applied, which, as these exploratory results show, should focus on highlighting and reinforcing various key points; namely the knowledge of specific academic connectors, the instruction about their positional variation, and, finally, the development of argumentative structure.

5. REFERENCES


Notes:
1The AWL list classifies academic words in categories that run from 1 the most common words, to 10 the least common. Transitional Words start to appear from category 4 onwards.
2For instance, the GSL (General Service List of English Words) created by West in 1935, on the contrary, was not considered as it did not fulfill the objective of focusing on specifically academic vocabulary.
3Calculation of frequencies on an estimate of occurrence over 100.000 words (see Lorenz 1999:19-20).