

Master in Meertalige Communicatie

Achieving cohesion through connectors

Connector usage in argumentative essays written
by Flemish EFL undergraduate students

Masterproef aangeboden door
Denver DE CLEER

Tot het behalen van de graad van master in Meertalige Communicatie

Promotor: prof. dr. Lieven BUYSSE

Academiejaar: 2018-2019

Master in Meertalige Communicatie

Achieving cohesion through connectors

Connector usage in argumentative essays written
by Flemish EFL undergraduate students

Masterproef aangeboden door
Denver DE CLEER

Tot het behalen van de graad van master in Meertalige Communicatie

Promotor: prof. dr. Lieven BUYSSE

Academiejaar: 2018-2019

Code of conduct

- **Code of conduct voor geloofwaardig auteurschap**
 - **Correct verwijzen naar bronnen houdt in:**
 - dat ik als auteur transparant ben over mijn bronnen, zodat de lezer op elk moment correct kan inschatten welke de bron is van wat beweerd wordt;
 - dat ik als auteur geen relevante informatie waarover ik beschik en die een ander licht zou kunnen werpen op de geboden interpretatie of ontwikkelde redenering bewust weglaat;
 - dat ik als auteur duidelijk aangeef welke relevante interpretaties ik niet in rekening heb gebracht en waarom dat is gebeurd;
 - dat ik als auteur waarheidsgetrouw weergeef wat ik heb geobserveerd;
 - dat ik als auteur op de hoogte blijf van evoluties in manieren van verwijzen en het weergeven van referenties en een bibliografie.

Ik bevestig dat ik deze *code of conduct* heb nageleefd bij het schrijven van deze masterproef.

Naam: Denver De Cleer

Datum: 24 mei 2019

Handtekening:

Preface

They say that a master's thesis is a one-man job and yet, as this work is nearing completion, I realise that such an enterprise could hardly have been done without the precious help of several few.

To my supervisor, professor Lieven Buysse, thank you. Your help, feedback and inexhaustible patience have helped me malleate my many thoughts into ideas and further shape them into a coherent text.

My thanks also extend to the KU Leuven professors of the Brussels and Antwerp campuses, who have kindly provided me with participants for this study: professors Van Steendam, Van Dessel, Dierickx, Roelans, Van den Eynden, Pauwels and Schramme. To the students, too, I owe my thanks. I can only imagine how tedious it must have been to carry out an assignment for seemingly no discernible reason. The students have also shown me just how diverse education has become: it is a wondrous evolution to behold.

For their suggestions and advice, I would like to thank professors Jaspers and Granger.

To the late Joris De Roy, my lecturer when I decided to embark upon my teacher training, six years ago now. Today still, I owe him my warm thanks for instilling into me a passionate love for the English tongue, culture, literature and quirkiness of the language. Without his sound advice, I would probably not have set out on the journey through the highs and lows of this MA in the first place. O Captain! My Captain! I thank you.

Acknowledgements are seldom written without a kind word to one's parents and rightly so. Their continuous all-encompassing support throughout my scholastic career—and beyond—have shaped me into the person I am today.

I have reserved the last and best spot on this page for my equally superlative Joni. If anyone ever showed patience, it has to be her. At last, after much ploughing through my thesis, here I am. Thank you for your unstinting moral support and the proverbial kicks up my backside. I truly needed them. In his memoir *On Writing*, Stephen King succinctly wrote, "Writing is a lonely job. Having someone who believes in you makes a lot of difference." I count my blessings every day.

Abstract

In the past few decades, cohesion theory through corpus linguistics has attracted a great deal of interest. All over the world, researchers have looked at EFL learners' use of cohesive devices. Looking at the way foreign-language learners use such domains as reference, ellipsis, conjunction or lexical cohesion enables teachers, linguists, educationalists, and policymakers to see where learners struggle to communicate clearly and coherently. In Flanders, no such research has yet seen the light of day. This study seeks to provide insight into the connector usage of Flemish university-level EFL learners in their written argumentative essays.

The research focuses on the use of connectors of Flemish EFL learners in written essays. A corpus called CoFLE was compiled from Applied Language Studies students at the KU Leuven and put side by side with another learner corpus, i.e. the Flemish subcorpus of ICLE. These two are compared to LOCNESS, an English native-speaker corpus. First, the connectors frequencies per semantic categories are analysed. Then, the study looks at the most-used connectors per corpus, identifies overused and underused connectors in the learner corpora and examines which syntactic position contains most connectors.

Flemish learners of English use twice as many connectors than native English speakers. When looking at the different semantic classifications of connectors, Flemish EFL learners consistently use more connectors than native speakers. Moreover, Flemish EFL learners overwhelmingly choose contrastive, listing and additive, and resultive connectors. This trend is reflected across all three corpora in their fifteen most-used connectors, where the three aforementioned categories ubiquitous. What's more, Flemish EFL learners from both non-native-speaker corpora overuse connectors in their writing.

This study has also found that Flemish learners of English mostly place connectors in initial position, thus reinforcing the notion that syntactic positioning may be a universal or interlanguage characteristic. The results also show that native-speaker learners, too, mainly use connectors in IP. Positioning connectors at the start of the sentence, as a consequence, is an inherent feature of English L1 and EFL learners alike.

The results of this study have several pedagogical implications. Flemish EFL learners should use other methods of cohesion besides connectors and, when using the latter, ought to diversify the position in which they are placed. Lastly, a holistic approach to language teaching, i.e. language immersion, will help understand the implicit ways texts become cohesive and coherent.

Table of contents

Code of conduct	i
Preface	ii
Abstract	iv
Table of contents	vi
List of tables	viii
Introduction	1
Literature review	3
1.1. Cohesion and coherence	3
1.2. Cohesive domains	4
1.2.1. Reference	4
1.2.2. Substitution	5
1.2.3. Ellipsis	6
1.2.4. Conjunction	7
1.2.5. Lexical cohesion	8
1.3. The foreign language learner perspective	8
1.3.1. Contribution of cohesive devices to essay quality	8
1.3.2. The frequency of cohesive devices	10
1.3.3. The use of connectors in the writing of learners of English and English natives	11
Research design and methodology	18
2.1. Rationale for the present study	18
2.2. Research questions	18
2.3. Model for classifying connectives	19
2.4. Selection of the connectors for observation and analysis	21
2.5. Defining overuse and underuse	22
2.6. On the vague distinction between connectors and conjunctions	23
2.7. The Academic Keyword List as a standard for linguistic correctness	24
2.8. Corpus compilation	26
2.8.1. Creating a corpus: meeting the ICLE criteria	26
2.8.2. The Corpus of Flemish Learners of English (CoFLE)	26

2.8.3. The Flemish-Dutch subcorpus of ICLE	28
2.8.4. The Louvain Corpus of Native English Essays (LOCNESS)	29
2.8.5. The retrieval software used	30
2.9. Statistical testing	30
Results	32
3.1. General study of adverbial connectors	32
3.1.1. Overall frequencies of adverbial connectors	32
3.1.2. Frequency of connectors per semantic category	32
3.2. Specific study of adverbial connectors	35
3.2.1. Most frequently used connectors	35
3.2.2. Overuse and underuse of connectors	38
3.2.3. The Academic Keyword List in CoFLE	45
3.3. Syntactic positioning of connectors	47
3.3.1. Clause-initial positioning	49
3.2.2. Clause-medial positioning	51
3.3.3. Clause-final positioning	52
Conclusion	55
Discussion	57
5.1. Discussion of the major findings	57
5.2. Pedagogical implications	59
5.3. Limitations	60
References	62
Appendix 1: essay topics	67
Appendix 2: on USB stick	68

List of tables

Table 1	The classification of connectors	21
Table 2	The 87 connectors chosen for analysis	22
Table 3	The adverbs in Paquot's (2010) <i>Academic Keyword List</i>	25
Table 4	The 'other' keywords in Paquot's (2010) <i>Academic Keyword List</i>	25
Table 5	Distribution of essays across BA years	27
Table 6	Distribution of first mother tongues	28
Table 7	Distribution of words and essays in my selection of LOCNESS	30
Table 8	Raw frequency and adjusted frequency of connectors in the three corpora	32
Table 9	Raw and adjusted frequencies of connectors per semantic category	33
Table 10	The 15 most frequently used connectors in all three corpora	35
Table 11	Connector overuse in the semantic categories of CoFLE	38
Table 12	Connector overuse in the semantic categories of FICLE	39
Table 13	Connector overuse in CoFLE	40-41
Table 14	Connector overuse in CoFLE where R_f in LOCNESS equals zero	41
Table 15	Connector underuse in CoFLE	42
Table 16	Connector underuse in CoFLE where R_f in CoFLE equals zero	42
Table 17	Connector overuse in FICLE	43
Table 18	Connector overuse in FICLE where R_f in LOCNESS equals zero	43
Table 19	Connector underuse in FICLE	44
Table 20	Connector underuse in FICLE where R_f in FICLE equals zero	45
Table 21	Connectors of the Academic Keyword List in the learner corpora	46-47
Table 22	Frequency of syntactic positioning of connectors in the three corpora	47
Table 23	Frequency of syntactic positioning of connectors after revision	48
Table 24	CIP and SIP of connectors in the three corpora	49
Table 25	Syntactic positioning of the 15 most used connectors in CoFLE	53
Table 26	Syntactic positioning of the 15 most used connectors in LOCNESS	53-54
Table 27	Syntactic positioning of the 15 most used connectors in FICLE	54

Introduction

Teachers, linguists and language aficionados know that English, despite its docile appearance, can in truth be quite challenging. When not taught properly, English can take on many non-standard forms, rendering clear and concise communication difficult. In Flanders, as elsewhere in the world, learners of English must be taught to avoid common mistakes and create cohesive and coherent texts. An essential linguistic cue to achieve this cohesion is by appropriately using connectors and conjunctions. They form what Halliday and Hasan (1979) consider to be one of the pillars of cohesion. In this study, connector usage by Flemish EFL learners is compared to that of English L1 learners.

A persistent problem I observed when teaching Flemish EFL learners is their inability to properly create coherent texts. All too often, they rely on a restricted set of connectors (e.g. Heino, 2010; Paquot, 2010) or use the wrong ones (Granger & Tyson, 1996). This study wishes to provide data and insight on connector use among Flemish EFL learners to ensure that future teaching practices and textbooks correctly identify and resolve problems among their learners concerning connector usage.

In the past two decades corpus linguistics has looked extensively at connector usage by EFL learners. No such research exists in Flanders, despite the ICLE (2002) containing an extensive subcorpus of essays written by native speakers of Belgian Dutch. This serves as an opportunity to compare a new and similar corpus to the older ICLE subset. With the help of these learner corpora, this study wishes to answer the following research question:

How do Flemish learners of English as Foreign Language use connectors in argumentative essays compared to English L1 learners?

To answer that question, a 43,899-word-strong *Corpus of Flemish Learners of English* (CoFLE) was compiled using 83 essays from undergraduate students at the KU Leuven. The Flemish subcorpus of the *International Corpus of Learner English* (or FICLE) was also included in the analyses. Both learner corpora were matched with the *Louvain Corpus of Native English Essays* (LOCNESS) to measure differences between L1 learners of English and EFL learners. With the help of grammars (e.g. Quirk et al., 1985) and previous research (e.g. Tapper, 2005), a list of connectors was drawn up for analysis.

The first chapter describes cohesion theory and looks at past research on the use of connectors in the writing of EFL and L1 learners. Chapter 2 goes into detail on the research design and methodology and Chapter 3 reveals the results this study has found. Chapter 4 discusses those results, arguing that caution should be exercised when interpreting data from research in corpus linguistics. Chapter 5 concludes the dissertation, looks at its implications for teaching English to Flemish learners and notes the possible limitations of this study.

Literature review

1.1. Cohesion and coherence

In 1976 Michael Halliday and Ruqaiya Hasan published *Cohesion in English*, a seminal book that would shape the academic discourse on cohesion and coherence. The authors start with the premise that communication happens through texts: spoken or written passages that form a unified whole (Halliday & Hasan, 1979). Halliday and Hasan (1979) and Eggins (1994) argue that a text is a written or spoken unit of meaning, not of form. A text becomes meaningful only when it is both coherent and cohesive (Adiantika, 2015). This interaction between cohesion and coherence is what Halliday and Hasan refer to as “texture”: the property of being a text. Together, cohesion and coherence form the crucial foundation for any language user wishing to communicate effectively and efficiently (Ghasemi, 2013).

Coherence is achieved by connecting ideas or clauses and making deductions through the surrounding social and cultural contexts (Adiantika, 2015; Eggins, 1994; Witte & Faigley, 1981). When the audience understands a text, when its meaning becomes relevant and consistent, the writer has created a coherent idea (Crossley, Kyle & McNamara, 2016; Reinhart, 1980; Zhang, 2000). Witte and Faigley define coherence as “semantic relations that allow a text to be understood and used” (p. 202). If a text is not sufficiently coherent, it “will almost certainly fail to communicate its intended message” (Bamberg, 1983, p. 417).

Cohesion is a set of semantic resources used to meaningfully relate parts of a text with what has gone before (Halliday & Hasan, 1979). These resources are the link between different semantic units or sentences, and hold a text together (Field & Yip, 1992; Witte & Faigley, 1981). Cohesion occurs when the interpretation of some element in the text is dependent on that of another: one presupposes the other and is dependent to it. Cohesion ties together a bit of discourse with what has immediately or remotely occurred before (Bamberg, 1983; Eggins, 1994; Halliday & Hasan, 1979). The reader can establish connections between ideas because of the presence or absence of linguistic cues that glue a text together (Crossley, Kyle & McNamara, 2016). To create cohesion, writers may use either local or global cohesion. Local cohesion (i.e. intersentential cohesion) connect parts of a sentence together, while global cohesive devices create “cohesion between larger chunks of texts” (Crossley et al., 2016, p. 1), for example between sentences or paragraphs. Local cohesion allows to look forward or backward in between clauses; global cohesion structures and organises a text (Yeqing, 2014).

Cohesion is crucial to texture to improve understanding, connectedness and consistency of sentences within a text (Hessamy & Hamedi, 2013). It is not, however, an item in particular that is cohesive; it is rather the relationship that a linguistic cue denotes between one clause and another that creates cohesion (Halliday & Hasan, 1979). It becomes particularly clear in Van Dijk’s (1980, in Bamberg, 1984) argument that local cohesion does not per se result in global coherence. Consider [1]:

[1] *I travelled home by train. Trains were first used in Belgium in 1835. Belgium is host to the European Parliament.*¹

Although the three sentences are tied by a form of lexical cohesion, namely repetition, the whole does not possess global coherence, because “the text lacks a semantic structure that relates all the propositions” (Bamberg, 1984, p. 307). Hence, it must be emphasised that a cohesive text is not by definition coherent (Carrell, 1982; Field & Yip, 1992; Ghasemi, 2013; Johns, 1986). It is context and the reader’s background knowledge that will render the text coherent; grammar and word choice will only make it cohesive.

1.2. Cohesive domains

Halliday and Hasan (1979) consider cohesion to rely on the meaning: without its semantic units, the text loses context — coherence —, rendering it illegible. To achieve cohesion, the authors recognise five cohesive device categories: reference, ellipsis, substitution, conjunction and lexical cohesion. The first four categories possess a grammatical function, the fifth a lexical one.

1.2.1. Reference

Reference occurs when the meaning of one item is not to be interpreted individually, but rather by relating it to another (Halliday & Hasan, 1979; Hessamy & Hamed, 2013; Sanders & Pander Maat, 2006). The referent — what the item refers to — can be present within the text (endophoric) or without (exophoric). Exophora heavily depend on the situation and context for them to be interpreted correctly. Endophoric references either refer back to an item (anaphora) or to an item that is yet to be mentioned (cataphora). In the phrase

[2] *The GP examined his last patient, after which he returned home for dinner. Although she too had a busy schedule, his wife waited for him to come back and eat together.*

the pronoun *he* is anaphoric, referring to *the doctor*. The pronoun *she*, contrariwise, alludes to the cataphoric referent *the doctor’s wife*. These two pronouns belong to the category of personal reference. Halliday and Hasan (1979) distinguish three categories: personals, demonstratives and comparatives.

Personal reference includes nouns (*Joni, car, etc.*), personal pronouns (*he, she, they, them, etc.*), possessive pronouns (*yours, theirs, etc.*) and possessive determiners (*my, your, etc.*). Halliday and Hasan point out that possessive pronouns are both referential and elliptical. The possessive pronoun *yours* in

[3] *I enjoy riding your bike. Yours is better at handling corners than mine.*

¹ This example is constructed and does not originate from sources mentioned earlier. This will henceforth be the case unless otherwise mentioned.

refers to *your* and elides *bike*. Similarly, *mine* refers to the exophoric *I* and elides the exophoric *my bike*.

Demonstrative reference is reference “by means of location” in time or space (Halliday & Hasan, 1979, p. 37) and includes demonstrative determiners (*this/that, these/those*) and adverbs (*here, there, then*) and the definite article *the*. Here, identifying the referent works on a scale of proximity: *this/these* and *here* point to something near; *that/those* and *there/then* designate something farther away. When demonstrative reference is usually used for verbal pointing (Halliday & Hasan, 1979), it is habitually used exophorically (see [4]). In that case, the reader relies on context to make out the intended meaning. Otherwise, demonstratives are also used to refer backwards and forwards in the text. *That* is always anaphoric (see [5]); *this* may either be anaphoric or cataphoric (see [6] for the latter) (Halliday & Hasan, 1979).

[4] *Rose, have a look at this.*

[5] *That man has been eyeing those garments suspiciously.*

[6] *Do you see the bridge in the distance? — This reminds me of a joke.*

Comparative reference, then, denotes similarity or identity, using comparative adjectives (*identical, better, less, etc.*) or adverbs (*likewise, so, more, less, etc.*). Adjectives function as a modifier to the noun phrase, while adverbs are adjuncts attached to an adjective, verb or other adverb. Therefore, in [7], *more* and *better* are adjectives, while in [8], *more* functions as an adverb.

[7] *You will have to present more facts if you want to make a better case.*

[8] *The rival products are more readily available than ours.*

1.2.2. Substitution

Substitution is a cohesive device whereby one item (the antecedent) is replaced by another which adds information to what has come before (Halliday & Hasan, 1979). It differs from reference in that substitution is concerned with wording rather than meaning: you replace words in a text to avoid repetition and embellish your writing (Adiantika, 2015; Hessamy & Hamedi, 2013). To this effect, Halliday and Hasan (1979) recognise three types of substitution: nominal, verbal and clausal substitution. These correspond to the functions which the substitute may take, i.e. a noun, a verb or a clause.

When applying nominal substitution, the writer replaces a countable noun phrase with (most commonly) *one* or *ones*. Crucial in understanding the difference between nominal reference and nominal substitution is the concept of redefinition (Halliday & Hasan, 1979). In the former, there is none: the presupposed item and its replacement are identical (e.g. *the man* → *he*). In the latter, some new specification is added. For example:

[9] *Did you buy red peppers? — No, I could only find yellow ones.*

In this case, the substitute differs from its antecedent: the substitution adds information which contrasts with the presupposed item.

Verbal substitution takes the form of *do* (or variations thereof) and its use resembles that of *one* in nominal substitution. It varies from the pronoun in that *do* may substitute not only a verb, but also a verb and other elements. In example [16], *do* substitutes *take him for your husband*.

[10] *Do you take him for your husband? — I do.*

Clausal substitution, then, is when the antecedent that gets replaced constitutes an entire clause (Halliday & Hasan, 1979). The substitutes are the words *so* and *not*. Clausal substitution may only be used in reported (or declarative), conditional [13] or modal [14] clauses. *Yes* and *no* may also be suitable substitutes for clauses, when you wish to answer a question in the affirmative or negative.

[11] *Is it going to rain tomorrow? — I believe so.*

[12] *Did you see that man running across the roof? — I did not.*

[13] *If you are found guilty, you face up to ten years in gaol. If not, you will walk free.*

[14] *Would you like another cup of tea? — I think not, thank you.*

1.2.3. Ellipsis

Ellipsis is the positive omission or non-expression of one or more words that are expected in the syntactic structure of a sentence (Adiantika, 2015; Cherchi, 1978; Halliday, 1994). The semantic element that is left out “may be presupposed at a subsequent place in the text . . . by saying nothing, where something is not required to make up the sense” (Halliday, 1994, p. 309). When using ellipsis, one literally elides a word or (a part of a) clause that is structurally necessary from a phrase (Halliday & Hasan, 1979). For example:

[15] *The woman ordered two drinks. The man ordered only one.*

[16] *Have you been cycling? — Yes, I have.*

[17] *Omar bought five t-shirts; Michelle, two.*

Because ellipsis closely resembles substitution (Halliday and Hasan (1979) see ellipsis as “substitution by zero” (p. 142)), this cohesive device falls, like substitution, into three categories: nominal, verbal and clausal.

Nominal ellipsis, first, leaves an antecedent, which functions as a noun or noun phrase, unsaid. In [15], *drink* is left out in the second clause. Second, verbal ellipsis elides within the verbal group. The elliptical form expresses finiteness, polarity (positive or negative), voice (active or passive) and tense (past or present) (Halliday & Hasan, 1979). In the response clause of [16], therefore, *have* expresses an active, positive, present perfect continuous form that is found in the first clause, i.e. *have been cycling*. Third, clausal ellipsis is used when there is no need to repeat the details from the preceding

clause. Example [17] shows that the second clause substitutes the first one by eliding the clause *bought t-shirts*.

1.2.4. Conjunction

Conjunction functions to connect two independent units (Hessamy & Hamedi, 2013). It establishes a link between what has already been said and what is yet to come (Adiantika, 2015; Field & Yip, 1992). This link is either coordinating or subordinating. This domain differs from the others, in that the semantic relations are explicitly marked, i.e. their sole purpose is to join ideas together (Sanders & Pander Maat, 2006; Witte & Faigley, 1981). Unlike the other cohesive devices that instruct readers to search for the tie in the text, conjunctions prepare them for what is to follow (Halliday & Hasan, 1979; Halliday, 1994).

Conjunctions are not to be confused with connectors. Conjunctions link clauses or phrases within a sentence; these clauses may be coordinate (*and bought* in [18]) or subordinate (*though* in [19]). Connectors are similar in function to conjunctions, though they link sentences instead of clauses or phrases, tying together all the units that together make up a text (*however* in [20]).

[18] *Mum and I went to the grocer's and bought vegetables for tomorrow's supper.*

[19] *He told me you could help me, though the opposite seems to be true.*

[20] *Stay home if it starts snowing tomorrow. If not, however, I expect you to be at work.*

Connectors appear in the form of adjuncts, of which three types exist: adverbs (*and, but, or, accordingly, therefore, etc.*), other compound adverbs (*however, therefore, nevertheless, etc.*) and prepositional expressions (*despite that, after, until, etc.*) (Eastwood, 2005; Halliday & Hasan, 1979). From those types, one differentiates connectors according to the semantic relationship they signal between two sentences in the text. There are several classifications of linking adverbials (e.g. Biber, Johansson, Leech, Conrad & Finegan, 1999; Greenbaum, 1996; Quirk, Greenbaum, Leech & Svartvik, 1985) that distinguish semantic classes other than Halliday and Hasan's (1979). While alternative groupings of the logical connectors provide an interesting insight into the subtleties of language, Halliday and Hasan's (1979) framework still offers the basis for any analysis.

All three types of conjunctive ties (adverbs, other compound adverbs, and prepositional expressions) are divided by Halliday and Hasan (1979) into four categories, depending on the semantic relations they establish between two items. These classes are: additive, adversative (or contrasting), causal and temporal.

Additive connectors add two clauses together in a text: they are the "and link" (Field & Yip, 1992, p. 15). Conjunctives that may be used include *and, or, moreover, in addition, similarly* and *for instance*.

Adversative or contrastive relations warn the reader that what is to follow is the opposite of what is expected and include conjunctives such as *but, yet, in fact, instead, however, rather* and *anyhow*.

Causal connectives join two sentences, one of which is cause, the other effect. This relation signals a reason or an argument (Field & Yip, 1992). You may find causal conjunctives such as *so, hence, as a result, because* and *in that case*. Causal connectives are also referred to as consequential (Martin, 1992), inferential (Biber et al., 1999) or resultive (Quirk et al., 1985) connectors.

Temporal connectors link sentences together by telling us when actions happen and in what sequence. Connectors for this category include *then, earlier, after that, later, this time, meanwhile* and *finally*.

1.2.5. Lexical cohesion

Lexical cohesion is a cohesive domain that establishes connections by the selection of vocabulary. The process concerns the repetition of a noun phrase which relates to the antecedent item (Adiantika, 2015). Sanders and Pander Maat (2006) write that elements which are considered to be lexically cohesive belong to the same lexical field. They then go on to say that the definition can be viewed more broadly: lexical cohesion is present if the pair of items have some perceivable lexicosemantic connection. This means that in [21], 'snowing' and 'coats' have some lexicosemantic connection.

[21] *It started snowing. The kids grabbed their coats.*

Halliday and Hasan (1979) divide lexical cohesion into reiteration and collocation. Reiteration comprises repetition (see *book* in [22]), (near-)synonymy (see *parrot* and *bird* in [23]) and superordinates (hyponymy and hypernymy) (in [23], *bird* is the hypernym, *parrot* the hyponym).

[22] *I have seen King's new book in the stores today. That book looks amazing.*

[23] *Amelia has a pet parrot, Rory. Her domestic bird can repeat up to thirty words.*

1.3. The foreign language learner perspective

1.3.1. Contribution of cohesive devices to essay quality

For an idea to be understood by the reader, for it to communicate its intended sense, that idea needs to be cohesive and coherent (Bamberg, 1983; Crossley, Kyle & McNamara, 2016; Halliday & Hasan, 1979). Though the studies that will be discussed in this paragraph distinguish in some way between 'good' and 'weak' compositions, there is no consensus on the definitions for those values. Being able to write well, Leńko-Szymańska (2015) argues, entails lexical fluency and grammatical correctness. Together with cohesion and coherence, these aspects make up what is considered to be 'good' writing (Liu & Braine, 2005). Cohesion alone, however, does not necessarily make a text meaningful. With the study of cohesion gaining mainstream appreciation in research papers, so did the hypothesis that the use of cohesive devices positively correlates with the quality of written essays.

In their 1981 study, Witte and Faigley looked at high-rated and low-rated essays of English which were written by native speakers. They provide evidence that strong compositions are denser in cohesion: 31.7% of all words in high-rated compositions contribute to explicit cohesive ties compared to 20.4% in low-rated essays. Liu and Braine (2005) and Yang and Sun (2012) also report a positive correlation between the use of local cohesive devices and the writing quality of Chinese EFL students. Fei (2006) observes a significant correlation between the use of adverbial connectors and writing quality. More advanced learners employ, according to Yang and Sun (2012), more cohesive devices and do so more

accurately. Yang and Sun further argue that the writing competence of senior university students stabilises the more proficient they become, enabling them to be more cohesive and coherent. Jafarpur (1991) notes that the number of cohesive ties only accurately predicts essay quality with students who are more proficient in English. Jafarpur's (1991) findings are premised on the idea that the more experienced the writer is, the more often and the more varied they will use local cohesive devices in their essays, thus enhancing coherence. This assertion is reverberated by Jalilifar (2008), who found that Iranian graduate students and learners of English (who in this instance are more experienced than the junior and senior groups) used more discourse markers and did so more accurately. The more conjunctive ties an essay contains, the researcher claims, the higher its quality. Jalilifar (2008) perceives a direct and statistically significant relationship between the quality of the essays written by Iranian learners of English and the number of correctly used discourse markers.

Other studies showed a non-significant or even negative correlation in the number of cohesive devices and the quality of the essay. Crossley and McNamara (2012) suggest that more proficient writers compose texts containing fewer words than less accomplished EFL writers. Proficient second-language writers appear to produce texts that are linguistically sophisticated (displaying more lexical diversity, logical operators, etc.) and are less reliant on local cohesion. This negative correlation has previously been observed (McNamara, Crossley & McCarthy, 2010) in L1 English students. Crossley, Kyle and McNamara (2016) demonstrate that over time EFL learners experience, much like native learners of English, a growth in the use of local and global cohesive devices. Yet, when looking at the link between judgements of essay quality and cohesion patterns, their results were inconclusive.

Several other research papers indicate that no significant link exists between cohesion and essay quality. Johnson (1992) looked at English essays written by native speakers of Malay and English. She finds no differences in the number of cohesive devices used by writers of 'good' or 'weak' compositions. The two distinct evaluations were determined by the students' teachers of Malay and/or English. In other words, cohesion plays no role in writing quality. Dastjerdi and Samian (2011), Ghasemi (2013), Zhang (2000) and Witte and Faigley (1981), too, discern no relation between the number of cohesive devices and writing quality.

Instead of attributing poor essay quality to a lack of cohesive devices or good essays to an abundance of cohesion, Ghasemi (2013) instead ascribes unsatisfactory essays to the students' possible incapacity to appropriately communicate in English. The discrepancy between more and less advanced learners, he argues, could reasonably be attributed to their (in)ability to find and use the proper linguistic cues to create a cohesive and coherent text. Dastjerdi and Samian (2011) further argue that by overusing cohesive devices, learners may also misuse them. In doing so, they negatively impact the quality of their written compositions. According to Bamberg (1983; 1984), cohesion and coherence depend on the reader's expectations and their ability to fill in or bridge gaps. If these gaps are too big to overcome, the writer may have undercued, i.e. they have adopted too few cues to create connections in the text, or miscued, i.e. used the wrong cohesive devices.

The results of the research papers on this topic can be classified as “controversial,” according to Ghasemi (2013, p. 1617). Crossley et al. (2016, p. 4) describe them as “mixed”: while some have found a positive correlation between the number of cohesive devices and ‘good’ writing, others have demonstrated that there is no significant link between the two factors. Researchers therefore discourage English instructors to exclusively teach their students cohesion theory to successfully build a meaningful text (Carrell, 1982, Field & Yip, 1992; Ghasemi, 2013; Granger & Tyson, 1996; Johnson, 1992; Rahman, 2013). Texture can only be achieved through cohesion and coherence (Halliday & Hasan, 1979). Many do not believe that the quality of a text can be judged by analysing cohesion alone. Witte and Faigley (1981) believe cohesion and coherence to interact greatly, “but a cohesive text may be only minimally coherent” (Witte & Faigley, 1981, p. 200). Granger and Tyson similarly suggest that “no matter how much students study connectors or any other aspect of cohesion, an incoherent message will always remain so” (Granger & Tyson, 1996, p. 26). Ghasemi (2013) therefore suggests teachers of English should work with their pupils not on the quantity of cohesive devices, but on their quality. It is more important that the learners know how to use them correctly. Cohesive devices are only one way to make a text more cohesive.

1.3.2. The frequency of cohesive devices

Although researchers, among whom Johnson (1992), Witte and Faigley (1981) and Yang and Sun (2012), have discovered no evidence to link quality of compositions with the number of cohesive ties, they have noted certain trends in the types of cohesive devices used by learners of English. Analysing what cohesive devices EFL students choose may better our understanding of language use and where the learners experience problems.

In research mapping the different categories of cohesive devices that learners of English use in their expository or academic dissertations, it appears that EFL students most commonly apply lexical cohesion to tie together parts of discourse (e.g. Adiantika, 2015; Dastjerdi & Samian, 2011; Johns, 1980; Khalil, 1989; Ghasemi, 2013; Liu & Braine, 2005; Rahman, 2013; Sanczyk, 2010; Zhang, 2000). Also widely used are conjunctions and reference, both of which either take second or third place (e.g. Dastjerdi & Samian, 2011; Johns, 1980; Khalil, 1989; Liu & Braine, 2005; Rahman, 2013; Sanczyk, 2010; Zhang, 2000). Ellipsis and substitution are found infrequently in the writing of non-native speakers, if at all. Even when looking at written discourse other than discourse essays, similar patterns emerge. When Johns (1980) analysed the cohesive ties in business letters, reports and textbooks, she found lexical cohesion to be the most common lexicogrammatical cohesive device used (46% for letters, 79% for reports and textbooks).

Proficiency is another factor that plays a role in the way EFL students make their text cohesive. Rahman (2013) has observed that first-year learners mostly rely on lexical cohesion (47.8%), while third-year non-native speakers, reference and lexical cohesion (each at 39%) are used just as much. The researcher attributes this to the capacity of the third-year students to vary their vocabulary and use fewer repetitions. Yang and Sun (2012) report that the more advanced the student is, the greater the number of local cohesive ties employed to link sentences together. The authors have also found that the highly proficient students would more often use those devices correctly. Yeqing (2014) found that

Chinese English-major-undergraduates, who have a lower proficiency in English, tend to overuse adverbial connectors compared to postgraduates following the same course. This same overall tendency applies to English L1 students, who mostly resort to lexical cohesion as well: about two-thirds of all cohesive ties used fall into that category, with students with a lower proficiency employing more lexical ties than high-proficiency learners (Witte & Faigley, 1981). In Rahman's (2013) study, then, native speakers integrate more cohesive devices and rely less on lexical cohesion (33%) than first-year (47.8%) or third-year (39%) Arabic EFL learners.

Researchers have theorised how these apparent discrepancies between native and non-native students, and between more or less proficient students, can be explained. Adiantika (2015), Yang and Sun (2012) and Zhang (2000) among others suggest that EFL students overusing lexical cohesion could be due to their lack of lexical repertoire. Accordingly, repetition is by far the most prevalent lexical cohesive tie found in the essays of EFL learners (Adiantika, 2015; Khalil, 1989; Rahman, 2013; Zhang, 2000). Lexical cohesion is the go-to cohesive domain for foreign-language learners, because it allows them to rely on reiteration patterns that are easy to reproduce: repetition, synonymy and superordinates (Halliday & Hasan, 1979; Rahman, 2013). More skilful English students, native learners or not, tend to use more grammatical cohesion, such as conjunctive ties (Witte & Faigley, 1981). It is for this reason that the above-mentioned studies (e.g. Khalil, 1989; Rahman, 2013; Witte & Faigley, 1981; Zhang, 2000) were less likely to find conjunction or reference cohesive ties and detected almost no cases of substitution or ellipsis. It seems that the frequency of cohesive devices and its distribution across the different domains may be a universal characteristic among EFL learners, as the results showed among Indonesian learners of English (Adiantika, 2015), Iranian students (Dastjerdi & Samian, 2011), Arab learners (Khalil, 1989), Chinese learners of English (Liu & Braine, 2005; Zhang, 2000) and Polish students (Sanczyk, 2010).

1.3.3. The use of connectors in the writing of learners of English and English natives

1.3.3.1. The frequency of semantic types of connectors

When Tapper (2005) analysed 279 essays written by Swedish EFL students, she found that the semantic category of contrastive connectives was the one most frequently used. The American students whose compositions were used for comparison, preferred these connectors over the others as well. The other most used semantic categories are, in decreasing order, resultive, clarifying, additive, corroborative and transitional.

Appel and Szeib's (2018) analysis of essays composed by Arabic learners of English found that they most often used listing connectives. The other semantic categories are, by order of popularity, resultive, contrastive, additive and appositional.

Chinese and Cantonese learners of English, it seems, favour contrastive, resultive, listing and additive connectors (Appel & Szeib, 2018; Field & Yip, 1992; Yeqing, 2014; Xu & Liu, 2012; Zhang, 2000). Lee (2004) observed a similar pattern in Korean students, who most frequently use contrastive and resultive connectors; additive and causal, too, are regularly used (Ha, 2014; Ha, 2016; Lee, 2004).

French learners of English, add Appel and Szeib (2018), use, in decreasing order of frequency, corroborative, listing, resultive and contrastive the most. Lithuanian learners of English will use listing, contrastive, appositive and resultive connectors most frequently (Bikelienė, 2008a).

With Iranian learners, Hayjian (2015) observes that resultive ties are most frequently used. They are then followed by listing, appositive and contrastive connectives. This differs from Dastjerdi and Samian's (2011) research, where additive ties account for over half of all conjunctive devices used, followed by causal ties, adversatives and temporal connectors. This research ties in with Jalilifar's (2008), which investigated the use of discourse markers in composition writings of Iranian learners of English. The study shows that 73% of said discourse markers were elaborative (similar to the additive conjunctives found in Dastjerdi and Samian's (2011) work). They are followed by inferential, contrastive and causal markers. It is clear, therefore, that although similarities may exist in the studies done with Iranian EFL learners, there is no uniformity even in one country. The discrepancies could be ascribed to the difference in size of the corpora or to the different programmes in which the participating students are enrolled.

The semantic categories that are used most frequently by learners of English worldwide vary depending on their mother tongue. Learners of a language are never impervious to the ramifications of their native language (e.g. Yeqing, 2014). They will inevitably pick up, for better or for worse, grammatical and syntactical structures and apply them to the target language, making it easier for them to construct sentences. Learners, Granger (1998, as cited in Hajjyan, 2015) writes, "cannot be regarded as phraseologically virgin territory: they have a whole stock of prefabs in their mother tongue" (p. 166). This is also known as L1 transfer. The frequency of connectors can therefore not be seen as an interlanguage feature: mother-tongue influence may be a possible explanation for foreign learners of English overusing connectors. Leńko-Szymańska adds that, when writing in a foreign language, learners will apply L1 transfer of linguistic features as well as rhetorical conventions, for instance text structure, explicitness or intertextuality. When those patterns are produced by every foreign learner of English in their respective L1, then differing uses in connector frequency will inevitably be found.

Another explanation for learners' propensity to use certain semantic categories over others, might lie in the genre of the text in which they are writing — in this case argumentative essays. When writing argumentative essays, students will apply a certain pattern to bring their composition to a satisfactory end. In *A Genre Description of the Argumentative Essay*, Ken Hyland (1990) outlines the basic structure that is found in argumentative essays. He recognises three stages in argumentative essays: thesis, argument and conclusion. He further identifies a plethora of moves used in the rhetoric of writers, some of which include claim, support, consolidation and marker. The latter, for examples, serves to indicate "addition, contrast, condition, specificity, etc. by adverbial connectives . . . indicating changes in the discussion" (Hyland, 1990, p. 73). These so-called moves in the writing of argumentative essays means that students are going to rely, consciously or not, on certain set structures that need connectors to create cohesion between ideas, sentences and paragraphs. This happens whether students are learning English as their first or as a foreign language. Connectors which for instance mark addition

(e.g. also or moreover), contrast (e.g. but, however or although) or corroboration (e.g. of course) are likely to appear in students' writings, no matter if they are English learners or native speakers. Hence why Tapper (2005) and Lee (2004) identified contrastive connectors to be the most frequent category with Swedish and Korean EFL learners respectively, or why Appel and Szeib (2018) found listing and corroborative connectors to be used most often by Arab and French students.

1.3.3.2. Overuse and underuse of connectors

Conjunctions and connectors are, together with reference, the most frequent grammatical cohesive devices found in English discourse writing (e.g. Field & Yip, 1992; Khalil, 1989; Liu & Braine, 2005; Rahman, 2013; Zhang, 2000). Several studies have compared the use of connectors by learners of English and native speakers in academic essays (e.g. Altenberg & Tapper, 1998; Appel & Szeib, 2018; Bolton, Nelson & Hung, 2002; Chen, 2006; Field & Yip, 1992; Granger & Tyson, 1996; Ha, 2014; Ha, 2015; Hajjyan, 2015; Heino, 2010; Lee, 2004; Narita, Sato & Sugiura, 2004; Tapper, 2005). They aimed to discover the overuse or underuse patterns of conjunctive ties into which the writing of foreign language learners may fall.

No clear definition exists of overuse or underuse (Appel & Szeib, 2018; Chen, 2006; Ha, 2016). Comparing frequency figures between native speaker (NS) and non-native (NNS) corpora enables researchers to note where differences in the use of cohesive devices arise. If the NNS corpus contains more of a certain cohesive tie than its native counterpart, there is overuse of that particular connector. Conversely, when the NNS corpus comprises fewer hits than the NS one for a cohesive device, there is underuse of that one item. Overuse and underuse are not marks of composition quality; instead, they only indicate trends in the learners' writing from which researchers have made varying deductions and assumptions. Shea (2009) notes that these terms are devoid of connotations which signal problematic use; Heino (2010) and Tapper (2005) say they do not indicate incorrect usage but are only used as "descriptive labels" (Heino, 2010, p. 2).

Contrary to their hypothesis, Granger and Tyson (1996) report that French EFL learners do not overuse or underuse connectors in their compositions compared to native peers. These findings are reverberated by Bikelienè (2008b), who, when looking at the use of resultive connectors by Lithuanian learners of English, observed a slight but non-significant underuse of that category, except for the connector *therefore*, which appears to be significantly underused (Bikelienè, 2008a; 2008b). However, some kinds of conjunctive ties are applied more often by EFL learners than by English native speakers in Granger and Tyson's (1996) study. Connectors that confirm, exemplify and add to the arguments were overused. The researchers identify contrasting connectors as being underused by French EFL students. Ha (2014; 2016) and Narita, Sato and Sugiura (2004) report similar findings, noting that Korean (and Japanese, in the latter study) learners tend to underuse contrasting connectors such as *yet* and *instead*. Likewise, Chen (2006) noted that Taiwanese EFL learners underuse adversative connectors.

Heino's (2010) research into connector usage by Swedish learners of English shows, like Altenberg and Tapper's (1998) study, a significant underuse of connectors compared to native speakers. In contrast, Tapper (2005) did find a general overuse of connectors. Swedish learners of

English, Tapper (2005) reports, overuse clarifying and corroborative connectors. Altenberg and Tapper noted in 1998 that Swedish learners overused appositive conjunctive ties. Tapper (2005) speculates that corroborative logical connectives may be a “shared learner language feature” (Tapper, 2005, p. 124), on account of Granger and Tyson’s (1996) study, which found this overuse also among French and, to a lesser extent, German learners of English. Tapper (2005) bases this argument on the premise made by Altenberg and Tapper (1998) that the genre of academic essays by its very nature necessitates more corroborative connectors to be used than any other genre. This would explain the overuse of corroborative connectives by Swedish EFL learners. Their claim is countered by stronger arguments put forward by Appel and Szeib (2018) and Granger and Tyson (1996). The French students that Appel and Szeib (2018) surveyed showed a tendency to overuse the corroborative connectives *in fact* and *indeed*, which was also reported by Granger and Tyson (1996). This, the latter argue, is because of the French’s inclination to use *en fait* and *en effet*, which are used as stylistic enhancers in their mother tongue. However, when used in English, *in fact* and *indeed* both carry a corroborative connotation.

In 2015 Hajiyan analysed the connector usage with Iranian university learners of English. The researcher found that the students use significantly more connectors than the native speakers from LOCNESS. The Iranian learners overused connectors in every category analysed save the contrastive type.

Ha (2014) predicted that Korean students would also overuse connectors when writing essays. Like Granger and Tyson’s hypothesis (1996), hers was also disproved: no overall overuse of connectors was detected. Ha notes that additive connectors such as *moreover* are used significantly more often by non-native than native speakers. This is in line with Field and Yip (1992), Ha (2016), Lee (2004), Chen (2006), and Narita et al. (2004), who similarly observe these trends in students of English from Hong Kong, Korea, Taiwan and Japan respectively. Fei (2006), Yeqing (2014) as well as Xu and Liu (2012), contrary to the aforementioned studies on the use of connectors by Asian learners of English, report that Chinese learners of English overuse adverbial connectors in their compositions. Bolton, Nelson and Hung (2002) and later Milton and Tsang (2003) found that Hong Kong learners of English in both cases overused a variety of logical connectors compared to published academic English. The latter study emphasises that native English-speaking learners of English, too, were observed to overuse connectors.

In 2016 Ha studied the use of linking adverbials in the writing of Korean university EFL learners. She finds that the learners overuse adverbials in “all the semantic categories” (Ha, 2016, p. 1099), in particular temporal, additive and causal linking adverbials. Only contrasting connectives are underused. The author attributes this overuse to Korean students wishing to create surface logicity, i.e. superficially linking clauses instead of developing ideas “in more subtle ways” (Ha, 2016, p. 1096).

While Ha (2014) looked at the use of connectors in general, Ha (2016) focuses on a subset of those connectors (Eastwood, 2005): linking adverbials. This explains why Ha (2014) on the whole negates the overuse hypothesis for connectors, while Ha (2016) firmly supports it in linking adverbials. Furthermore, Ha (2014) notes that Korean EFL learners do overuse certain additive adverbials, namely *moreover*, *besides* and *furthermore*, which Ha (2016) confirms.

Hungarian learners of English were found to overuse adverbial connectors, especially resultive and contrastive ones (Tankó, 2004). Bikelienė (2008a) found that Hungarian and Lithuanian students alike overuse connectors, with Lithuanians tending to overuse listing, appositive and contrastive connectors, while underusing resultive and transitional ones. The researcher reports that while the compositions written by learners and natives contain nearly the same number of adverbial connectors, the learners' essays lack variety. The few types of connectors found in the EFL learners' work indicates frequent repetition or, in other words, connector overuse. This is not the case for Lithuanian learners of English, Bikelienė (2008a) reports. Tankó (2004) ascribes this comparative overuse to the way Hungarian language works. The author explains that unlike English, Hungarian "does not require the overt marking of relations between linguistic units of the texts" (Tankó, 2004, p. 177). Xu and Liu (2012) echo Tankó's (2004) findings, explaining that students are taught that texts need conjunctive ties in order to be logical, cohesive and coherent. Subsequently, the learners will use connectors on purpose, to achieve "surface logicity" (Chen, 2006, p. 125). The author further adds that teachers time and again teach their students to express textual cohesion explicitly, by emphasising the use of adverbial connectors in their writing. Ha (2016) warns that Korean students should be made clear that connectors are but one way to achieve coherence in writing; they are not the end-all-be-all of textual cohesion. Rather, they are part of a larger framework that glues texts together.

Surface logicity becomes apparent in Chen's (2006) study. The researcher notes that Taiwanese learners of English use conjunctive ties to join sentences together without attempting to establish coherence. This means that a less experienced student is likely to write sentences one after the other, without them having a logical connection: there may be intersentential cohesion, but there is no coherence. Some semantic relations are implicit and need not be expressed explicitly. Overusing connectors may make a text cohesive, but it does not per se lead to coherence (Witte & Faigley, 1981; Yeqing, 2014). Chen's (2006) argument is reverberated by Tapper (2005), who says that connectors indicate coherence relations without creating them. Likewise, Appel and Szeib (2018) observe that Arabic writers of English overuse additive linking adverbials to try to make their text coherent "where no deep logical connections exist" (Appel & Szeib, 2018, p. 8). In that same study, Chinese students are found to overuse one contrastive adverbial, viz. *however*, whereas French EFL learners overuse connectors that exemplify arguments, thus tying in with the results of Granger and Tyson's (1996) study.

Leńko-Szymańska (2008; 2015) notes that overuse of linking expressions may not be inherent to EFL learners. Rather, the researcher points out, it may be a general feature to be observed in any learner of English, whether native or not. Leńko-Szymańska's (2008) research demonstrates that both EFL students and native writers overuse connectors when compared to expert native English speakers (journalists etc.). This is further corroborated by Bolton, Nelson and Hung (2002), who report that overuse is a prominent feature both of EFL learners (in this instance from Hong Kong) and native (British) students. Leńko-Szymańska's (2008; 2015) juxtaposition shows that connector overuse by EFL learners is relative to (in this case British) native *learners*. When comparing both to *expert* writers, the discrepancy seems smaller and less significant (if at all). This view serves to balance the idea that the writing of EFL learners drastically differs in quality from that of published academic writing (Shea, 2009).

To sum up, it may be said that overuse and underuse of connectives are not universal — or interlanguage — features among learners of English; rather, they are language specific. Some researchers have found conclusive evidence that essays written by the participants contain connective overuse (e.g. Appel & Szeib, 2018; Chen, 2006; Fei, 2006; Field & Yip, 1992; Ha, 2014; Narita, Sato & Sugiura, 2006; Tapper, 2005), others have found the opposite — underuse — to be true (e.g. Altenberg & Tapper, 1998; Bikelienè, 2008b; Hajiyan, 2015; Heino, 2010). Others still, like Granger and Tyson, 1996, found neither overuse or underuse to apply to their research. Yet, despite the differences, a general pattern emerges: learners of English seem to rely on a restricted set of connectors (Heino, 2010; Paquot, 2010), which they as a result will overuse greatly to establish textual cohesion, even if it only creates “surface logicity on a piece of writing where no deep logicity exists” (Crewe, 1990, p. 320, as cited in Paquot, 2010, p. 175).

1.3.3.3. Syntactic positioning of connectors

In English, connectors can be placed in different positions in a sentence. It is most common to find connectors in sentence-initial position, though certain connectors, such as *however*, are most typically placed in medial position; others, while few, such as *though*, have a preferred final position (Quirk et al., 1985). It has been ascertained in several studies, for instance Field and Yip (1992), Lee (2004) and Ha (2014) that non-native learners of English use a greater number of conjunctive ties than native-speaker learners. Which semantic classifications they favour over others and which categories of connectors they over- or underuse, has been shown in 1.4.2.1 and 1.4.2.2 respectively. It appears that these features are not universal; on the contrary, they are strongly dependent on the learner’s mother tongue and their proficiency. The position of these connectors has also been studied by many researchers (e.g. Altenberg & Tapper, 1998; Bikelienè, 2008b; Dastjerdi & Samian, 2011; Field & Yip, 1992; Granger & Tyson, 1996; Ha, 2014; Hajiyan, 2015; Narita, Sato & Sugiura, 2004; Xu & Liu, 2012; Zhang, 2000) who have found strikingly similar results.

According to Dastjerdi and Samian (2011) and Hajiyan (2015), Iranian learners of English prefer using connectors in sentence-initial position when compared to native speaker students. These learners put connectors such as *however* or *nevertheless* in sentence-initial position instead of the more usual medial position. Paquot (2010) writes that in academic style, while it is customary to place connectors in sentence-medial position, learners of English do not favour it. This finding is echoed by Hajiyan’s (2015) participants, who preferred the sentence-initial position to give the impression of cohesion. Connectors occurred in sentence-final position in the latter’s study, but infrequently so. This, Hajiyan (2015) writes, is because the Iranian EFL learners felt unsure about putting conjunctive ties in that position.

Chinese EFL learners, according to Xu and Liu (2012) and Zhang (2000), choose to insert connectors in sentence-initial position over any other. As in Hajjyan’s (2015) case, the researchers found *however* to be overused in clause-initial position, where it accounts for 80% of the use of that particular discourse marker. In Zhang’s (2000) study, *however*, *moreover* and *therefore* were used

exclusively in sentence-initial position. Native learners, Xu and Liu (2012) report, use *however* merely 20% of the time in that position.

Hungarian learners of English will equally write connectors in sentence-initial position, Tankó (2004) reports. This trend is further observed with Japanese students (Narita, Sato & Sugiura, 2004), Hong Kong learners of English (Field & Yip, 1992), Korean EFL learners (Ha, 2014; Lee, 2006), Lithuanian EFL students (Bikelienė, 2008b), French learners (Granger & Tyson, 1996), Swedish learners (Altenberg & Tapper, 1998; Heino, 2010) and Polish EFL learners (Sanczyk, 2010).

From these studies we can conclude that learners of English as a Foreign Language, regardless of their mother tongue, seem to prefer the use of connectors in sentence-initial position over medial and final positions. Students, Xu and Liu (2012) explain, will adopt this tendency, having been taught that using connectors explicitly organises texts and creates cohesion. Narita, Sato and Sugiura (2004) add that learners of English want to build explicit links between the preceding and upcoming propositions. They build a bridge that acts as an immediate tie in a text (Appel & Szeib, 2018). Because the learners are instructed to use connectors in their academic writing, they will feel safer having used them at the unmarked sentence-initial position (Hajiyan, 2015; Paquot, 2010). Ha (2014) agrees, adding that the use of connectors in ISP might be an attempt to forge an immediate cohesive link. It could also be, according to Narita et al. (2004), that learners might not be aware of the difference between adverbial connectors and conjunctions, “let alone the flexibility of connector-positioning” (Narita et al., 2004, p. 1174).

It is clear that the tendency of learners of English to put connectors in sentence-initial position is a universal one. Already in 1996, Granger and Tyson speculated that this feature might not be language-specific. Two decades of research have shown that, indeed, the positioning of connectors is an interlanguage characteristic.

Research design and methodology

2.1. *Rationale for the present study*

Previous studies on conjunctive ties in the essays of learners of English as a Foreign Language have shown that although overall trends may exist, there are still significant differences depending on the L1 or region of those students. The *International Corpus of Learner English* has previously compiled essays written by Dutch learners (including some written by Belgian native speakers of Dutch). Given that, to the best of my knowledge, no previous research has been done on the use of connectors in the English essays of Flemish university students, not even for the Dutch component of the ICLE, there is a need to map the problems in that group's EFL writing. The rationale for this study is fourfold.

First, there is the type of composition. In essay writing, one habitually distinguishes four types: argumentative, narrative, descriptive and expository texts (Zhang, 2000). Argumentative essays are a common occurrence in the students' academic career, as they are often required to structure their arguments with relevant theses and antitheses (Ghasemi, 2013). This need for linguistic appropriateness further makes argumentative writing an adequate measurement tool of language expertise (Yang & Sun, 2012). Therefore, this study aims to investigate how Flemish learners of English use connectors in argumentative compositions and how that usage differs from that among native speakers.

Second, connectors are likely to appear frequently in argumentative essays (Granger & Tyson, 1996), since learners have to build a cohesive and coherent text in order to persuade the reader. It is expected that this study will find instances of connector overuse and underuse in the writing of EFL learners.

Third, a plethora of studies has found that learners of English favour using conjunctive ties in sentence-initial position. Authors such as Granger and Tyson (1996), Narita et al. (2004), Tankó (2004) and Bikelienè (2008b) argue that this behaviour is universal. This research hopes to add weight to this statement or to disprove it.

Fourth, despite the ICLE containing essays written by Flemish students (marked as DBAN1, DBAN2 and DBAN3 in the subcorpus), no research has been done that provides an overview of the inferences that can be drawn based on the input from the Dutch component of the ICLE. Comparing this corpus and my own could furnish interesting results. Either the data is similar and shows that connector usage in English in Flanders has not changed in 20 years, or the data yields different results; the discrepancies will then be analysed and explained. Moreover, with the help of Paquot's (2010) *Academic Keyword List* (AKL), which provides a list of words suitable for academic prose, I hope to find that Flemish EFL learners display register awareness when writing essays.

2.2. *Research questions*

Overall, this study seeks to identify the use of connectors in academic argumentative essay writing by Flemish EFL students. To do so, this research wishes to answer the following research question:

How do Flemish learners of English as a Foreign Language use connectors in argumentative essays compared to English L1 learners?

The main question is further divided into five questions, each of which deals with part of the research problem:

- (1) What connectors do Flemish EFL learners use in argumentative essay writing?
- (2) With what frequency do Flemish EFL learners use connectors compared to native speakers?
- (3) How does connector usage differ between Flemish EFL learners and native-speaker practice?
- (4) In what syntactic clausal position do Flemish EFL learners place connectors?
- (5) How do the results compare to the Flemish subcorpus of the ICLE?

My first hypothesis states that like previous research on learner use of connectors (see Section 1.3.3.2), this study will find Flemish learners of English to use connectors with greater frequency than their native counterparts. Furthermore, when looking at individual connectors, Flemish learners of English will use a narrower variety of connectors. On the other hand, the native-speaker corpus and learner corpora will all share certain connectors among their most popular ones, because these connectors are frequently used in any context. What's more, I expect learners to overuse connectors in most semantic categories. Lastly, I hypothesise that, since it is seen as an interlanguage feature (see Section 1.3.3.3), Flemish learners of English will overwhelmingly prefer to place connectors in sentence-initial or clause-initial position.

Even though coherence is necessary when writing an argumentative essay (the goal is after all to communicate an idea, to persuade the reader), this study looks at explicit cohesion through connector use, the glue that joins clauses and paragraphs together — the formal connection. Cohesion allows for quantitative empirical data; coherence not so much². It is important to note that this work merely looks at the usage of cohesion and its correctness in academic essays, *not* at the quality of the essay writing. Hence, in the discussion of the data the terms *overuse* and *underuse* do not imply a qualitative assessment of the participants' work but merely an observation of quantitative differences.

2.3. Model for classifying connectives

In Section 1.2.4 a short description of conjunctions and connectors was presented, based on Halliday and Hasan's (1979) classification of these ties. As was said in that paragraph, there are many more classifications that go into further details. Quirk et al.'s *A Comprehensive Grammar of the English Language* (1985), for instance, distinguishes seven categories of what the authors describe as

² Several studies have attempted to establish a framework that could score coherence in written and spoken texts. Most are based on the study by Grice (1975, in Khalil, 1989), which, although not designed to explain coherence, offers a solid basis for that very purpose. Grice proposes the Cooperative Principle, consisting of four maxims. Quantity (#1) requires a text to be as informative as is required. Quality (#2) means a text must not contain falsehoods. Relation (#3) is concerned with relevance. Manner (#4), then, implies orderliness, brevity and clarity. Although these maxims are by no means irrelevant to the composition of academic essays, the author believes them to be subject to the assessor's opinion. Using Grice's maxims in this paper is not the essence of the study ahead.

conjuncts: listing, summative, appositional, resultive, inferential, contrastive and transitional, compared to Halliday and Hasan's (1979) four. Martin (1992) distinguishes four classes, too, but divides them up with more precision than Halliday and Hasan (1979). Biber et al. (1999), in the *Longman Grammar of Spoken and Written English*, report six different categories (closely resembling Quirk et al.'s (1985) model) which groups the logical connectors: enumeration/addition, summation, apposition, result/inference, contrast/concession and transition. Greenbaum's (1996) *Oxford English Grammar*, lastly, discerns six categories of what he calls logical connectives: listing, apposition, result, reformulation, contrast and concession.

Evidently, every grammar book adopts a different take on the classification of connectors. All of the previously posited options present models with which one delineates one's work and which classify logical connectors. For the present study several classifications will be combined, viz. Quirk et al.'s (1985) division of connectors, complemented with elements from Tapper (2005). These models will help me outline my work and give me a clear framework with which to proceed.

Tapper (2005) presents a reworked version of Quirk et al.'s (1985) and Martin's (1992) models of connectives. Tapper argues that both books provide categories which are more fine-tuned than Halliday and Hasan's (1979). She also adds corroborative connectors (i.e. *in fact, actually, as a matter of fact, indeed*) as a separate category, as they do not feature in either Quirk et al. (1985) or Martin (1992). Lastly, Tapper removes temporal connectors, reasoning that these connectors are external to argumentative texts.

The present study adds further slight modifications to the two models mentioned above, by removing the categories of transitional and discoursal connectors. These types of connectors are mainly observed in narrative texts (Heino, 2010) and are therefore not relevant to argumentative essays. The final classification of the connectors is presented in Table 1.

- (1) Listing and additive
 - (a) Enumerative (e.g. *first(ly), second(ly), finally, last(ly), next, then*)
 - (b) Equative (e.g. *correspondingly, likewise, in the same way, similarly*)
 - (c) Reinforcing (e.g. *furthermore, also, moreover, what is more, in addition*)

- (2) Appositive
 - (a) Specific
 - (i) Exhaustive (e.g. *that is, i.e., in other words, namely, viz.*)
 - (ii) Non-exhaustive (e.g. *for instance, for example, namely*)

 - (b) Generality
 - (i) Local (e.g. *in general, generally*)
 - (ii) Particular (e.g. *particularly, specifically*)

- (3) Summative (e.g. *altogether, overall, then, therefore, thus, to sum up, to conclude, to resume*)

- (4) Contrastive
 - (a) Reformulatory (e.g. *better; rather, more precisely*)
 - (b) Replacive (e.g. *alternatively, alias, better, worse*)
 - (c) Comparative (e.g. *in comparison, on the one hand ... on the other hand*)
 - (d) Antithetic (e.g. *conversely, instead, by contrast, contrariwise, on the contrary*)
 - (e) Concessive
 - (i) Dismissive (e.g. *anyway, anyhow, besides*)
 - (ii) Counterexpectation (e.g. *however, nevertheless, yet, still, though*)

- (5) Resultive
 - (a) Concluding (e.g. *as a result, therefore, thus, as a consequence*)
 - (b) Inferential (e.g. *in that case, else, otherwise, in other words*)
 - (c) Explanatory (*after all*)

- (6) Corroborative (e.g. *in fact, indeed, actually, as a matter of fact*)

Table 1: The classification of connectors

2.4. Selection of the connectors for observation and analysis

Since this study aims to identify patterns of connector usage among Flemish learners of English, the frequency of connectors will be looked at according to their semantic classification (see Section 2.3).

My initial decision was to create an exhaustive list of connectors which would mainly derive from Quirk et al. (1985) and Greenbaum (1996). The resulting list of connectors was found to be too extensive. Including all the connectors would not make sense, since some connectors barely feature or do not appear at all in either of the three corpora in this study (e.g. *in consequence, in second place, subsequently, notwithstanding, in any events, alias, altogether, correspondingly*). Other connectors proved irrelevant for academic essays (e.g. *a, b, c*: technical use especially). Therefore, the logical step was to opt for a corpus-driven approach, whereby the selection of connectors depends on those I would find in the corpora analysed. After careful examination, a final list of 87 connectors was selected for this study. They are to be found in Table 2.

Listing and additive	first(ly), second(ly), third(ly), in the first place, in the third place, next, first of all, finally, lastly, then, to begin with, likewise, similarly, in the same way, and, furthermore, also, besides, moreover, what is more, in addition, additionally, above all
Appositive	that is (to say), i.e., in other words, namely, for instance, for example, e.g., in general, generally, on the whole, particularly, in particular, specifically
Summative	overall, then, therefore, thus, in conclusion, in sum, to conclude, briefly, in short
Contrastive	rather, meanwhile, but, in comparison, (on the one hand...) on the other hand, on the contrary, in/by contrast, conversely, instead, at first, then, in any case, anyway, besides, however, nevertheless, nonetheless, only, yet, still, (al)though
Resultive	as a result, therefore, thus, so, as a consequence, consequently, accordingly, somehow, hence, now, in that case, otherwise, after all
Corroborative	in fact, indeed, actually, as a matter of fact, of course, apparently, in effect, admittedly

Table 2: The 87 connectors chosen for analysis

2.5. Defining overuse and underuse

In Section 1.3.3.2, I mentioned that, to the best of my knowledge, there is no clear definition of overuse or underuse (Appel & Szeib, 2018; Chen, 2006; Ha, 2016). This subsection is not an attempt at providing a precise definition; instead, I want to find out what methodologies, which have previously been used, are feasible for this study.

In its easiest form, connector overuse occurs when a connector (or an entire semantic category) appears more frequently in a learner corpus than in a native one. Conversely, we may then speak of connector underuse when foreign learners of English use fewer instances of a connector than do the native speakers. To compensate for corpus size, one can look at adjusted instead of raw frequencies. This method is found in Chen (2006), Bolton, Nelson and Hung (2002), Fei (2006) and Tapper (2005). Granger and Tyson (1996), on the other hand, compared similarly sized samples from a native and non-native corpus and looked at the raw frequencies of connectors. In her 2016 study, finding that marking overuse seemed to be somewhat random, Ha decided to interpret overuse as “the minimum ratio . . . between the adjusted frequency of the NNS corpus and that of the NS corpus is . . . 1.5” (Ha, 2016, p. 1095). In other words, when dividing the adjusted frequency from learner corpus by that of the native speaker corpus, you should obtain 1.5 or more to categorise it as overuse. With regards to underuse, the author set the cut-off point at 0.95. Using the same calculation as before, everything figuring below that number then qualifies as underuse. Appel and Szeib (2018) decided on a seemingly more arbitrary method. If a discrepancy of more or less 14 occurrences per 1,000 sentences was detected, they would mark that connector as being overused or underused. A one-way ANOVA test would assess significance.

Despite the varying interpretations of what constitutes overuse or underuse, researchers agree that the two terms should be devoid of negative connotations (Shea, 2009). They describe a phenomenon (Heino, 2010; Tapper, 2005), but do not necessarily make it good or bad.

For this study, I decided to adopt Ha's (2016) method, because the researcher set fixed cut-off points for overuse and underuse. In doing so, I avoid haphazardly determining what constitutes overuse or underuse.

2.6. *On the vague distinction between connectors and conjunctions*

Only connectors will be analysed in English learners' texts, not conjunctions. Conjunctions and connectors are frequently confused; the items get used interchangeably in learner texts.

Grammatically speaking, conjunctions, as was clarified in Section 1.2.4, connect two independent units into one entity. They, in effect, bind clauses or phrases *within* a sentence. Connectors, on the other hand, link separate sentences to create cohesion. They join ideas *across* sentences. So, while the two terms perform grammatically similar functions, it is what they bind that makes them different. Consider the following uses of *yet*, where, in example [24], *yet* is used as a conjunction that ties two previously separate clauses into one sentence. Example [25], conversely, relates two ideas without turning two sentences into one.

[24] *Almost all existing sex and AIDS education classes stress chastity, **yet** half the nation's high school girls are sexually active; 16% have had four or more partners.* (ICLE-US-MRQ-0021.1)

[25] *All such past attempts have foundered on the thorny issue of national interest as present attempts at creating a common agricultural policy have shown. **Yet**, British fears do not seem [sic] to be particularly influenced by the historical perspective.* (ICLE-BR-SUR-0005.3)

In practice, the rules that define connectors are somewhat vaguer than that. The English language allows sentences to go on and on. While this practice is not strictly wrong, it renders texts incoherent. A well-crafted mix of main clauses and subclauses creates one long sentence that contains several (if not many) substructures. If an adverb functioning as a connector is found to link two clauses together, but not two sentences, ought one to classify it as a conjunction or a connector? After all, the only difference between that clause and a sentence is a capital letter and a punctuation mark. Consider examples [26] and [27] below.

[26] *I believe that the public has a right to be informed about anything and everything that they want to be informed about, and people want to be informed about the death penalty; therefore, media should have access to report on executions.* (ICLE-US-MRQ-0027.1)

[27] *I believe that the public has a right to be informed about anything and everything that they want to be informed about, and people want to be informed about the death penalty. Therefore, media should have access to report on executions.*

Had the writer chosen to separate the clauses before and after the semicolon in [26], the result would have been [27]. In both cases, *therefore* performs the same function as a resultive connector. Yet, by adhering to the grammatical rules, *therefore* in [26] does not qualify as a connector.

Where to draw the line? That is the question. For practical reasons, I will consider typical conjunctions such as *but* or *and* to be connectors when they are found at the head of a sentence. I am aware that my results might therefore include certain false positives (i.e. a word that I consider to be a connector acting as a conjunction). This study, however, wants to reflect Flemish learners' writing and their use of connectors. If they do use a conjunction as a connector (and actually considering them to be connectors), it would be unwise to leave those out and risk skewing my results. Indeed, were I to omit them in my calculations, I feel like I would be actively leaving out useful data. For this reason, I decided to analyse certain connectors in sentence-initial position only, because that is the only syntactic position in which they function as connectors (e.g. *and*, *but*, *so*). For others (such as *therefore* in [26] and [27]), I will analyse the connectors in sentence-initial, clause-initial, clause-medial and clause-final positions.

2.7. *The Academic Keyword List as a standard for linguistic correctness*

When do words fit an academic context and when ought they be left out? Conventional grammars and dictionaries provide an indication of the registers in which a word may or may not be used (for instance formal, neutral and informal). They do not, however, provide a specific context in which words are to be used or avoided. Paquot's (2010) monograph on academic vocabulary proposes a pertinent list of words that researchers should hope to find in academic discourse.

In her book *Academic Vocabulary in Learner Writing: From Extraction to Analysis* Paquot (2010) designs a keyword list of academic words, intending to improve Coxhead's (2000) previous approach, a corpus-based *Academic Word List* (AWL). It is an improvement on Coxhead's (2000) previous work on that same topic. Paquot (2010) devised, based on the premise that an academic word need not be a little-used one, a list of 930 potential academic keywords, aptly named *Academic Keyword List* (AKL). Her criteria somewhat match Coxhead's (2000), in that both require the words to be well-distributed across the corpora, to be frequently used and to appear in professional and student academic writing. Of those 930 keywords, 38% are nouns, 25% verbs, 19% adjectives, 9% adverbs and 8% others (determiners, conjunctions, prepositions, pronouns, articles and ordinal numbers). More than half of all the keywords found in the AKL do not feature in Coxhead's (2000) but are instead present in the 2,000 most frequent words in English. This, Paquot (2010) stresses, shows the importance of general vocabulary in specialised academic prose.

It is this list which is of interest to the present study, as it will contain connectors. Paquot's (2010) keywords (see "Academic Keyword List", n.d.) in those categories are to be found in Tables 3

and 4. When looking at the 87 adverbs and 75 other word classes, one recognises 40 possible connectors such as *accordingly, although, consequently, despite, for example, indeed, in general, secondly, therefore*, etc.

Adverbs (87 items)

above, accordingly, accurately, adequately, also, approximately, at best, basically, clearly, closely, commonly, consequently, considerably, conversely, correctly, directly, effectively, e.g., either, equally, especially, essentially, explicitly, extremely, fairly, far, for example, for instance, frequently, fully, further, generally, greatly, hence, highly, however, increasingly, indeed, independently, indirectly, inevitably, initially, in general, in particular, largely, less, mainly, more, moreover, most, namely, necessarily, normally, notably, often, only, originally, over, partially, particularly, potentially, previously, primarily, purely, readily, recently, relatively, secondly, significantly, similarly, simply, socially, solely, somewhat, specifically, strongly, subsequently, successfully, thereby, therefore, thus, traditionally, typically, ultimately, virtually, wholly, widely

Table 3: The adverbs in Paquot's (2010) Academic Keyword List

Others (75 items)

according to, although, an, as, as opposed to, as to, as well as, because, because of, between, both, by, contrary to, depending on, despite, due to, during, each, even though, fewer, first, former, from, for, given that, in, in addition to, in common with, in favour of, in relation to, in response to, in terms of, in that, in the light of, including, its, itself, latter, less, little, many, most, of, or, other than, per, prior to, provided, rather than, same, second, several, since, some, subject to, such, such as, than, that, the, their, themselves, these, third, this, those, to, unlike, upon, versus, whereas, whether, whether or not, which, within

Table 4: The 'other' keywords in Paquot's (2010) Academic Keyword List

Paquot's (2010) list of potential academic keywords provides a benchmark with which to compare students' writing in this study and elsewhere. In her monograph, the author does not provide a list of connectors that learners of English should use; it merely shows what words should be present. Of those, some are connectors. By looking at the connectors in this list and putting them side by side to my list of connectors to be analysed, I will be able to identify those connectors that are used appropriately by EFL learners and those that belong to another register than the academic one. It must be added, however, that the texts used to compile the AKL are academic in the truest sense of the word: theses, dissertations and reports. The essays I will analyse are argumentative and written in an academic style but are not strictly academic. In the allotted time, the participating students could not (and neither did I expect them to) compose a text that satisfied the rigorous criteria of an academic text. Analysing their writing with Paquot's (2010) AKL should as a consequence not be seen as a definitive judgement of the quality of academic writing. Rather, the AKL serves as a tool to measure register sensitivity. It could indicate how well learners of English are acquainted with the different connectors that are used in a more specialised (i.e. academic) setting.

2.8. Corpus compilation

In order to answer the questions raised in Section 2.2, I decided to adopt a corpus-based approach. Despite the ICLE containing a Dutch from Belgium subcorpus (see Section 2.8.3), I decided to compile a more recent corpus of argumentative essays written by Flemish learners of English. The resulting Corpus of Flemish Learners of English (CoFLE) uses learner criteria established by the Centre for Corpus Linguistics at UC Louvain. This enables a comparison between corpora that contain homogeneous properties (age, genre, proficiency and so forth). I will be using the native-speaker LOCNESS to juxtapose the students' essays to my own corpus of non-native speakers' texts.

2.8.1. Creating a corpus: meeting the ICLE criteria

When creating the non-native corpus, I deemed it best to meet the criteria of the *International Corpus of Learner English* (see e.g. "Corpus collection guidelines", n.d.; Granger, Dagneaux, Meunier & Paquot, 2009; Granger & Tyson, 1996). This way, the data under investigation for this present study would be collected similarly to other research which has employed the ICLE. It is "essential" (Granger & Tyson, 1996, p. 18) that the data should be comparable.

The first criterion laid down is that the participating students should learn English as a Foreign Language (EFL), not as a second language (ESL). If researchers do not follow this rule, Granger and Tyson (1996) say, it could result in skewed results, particularly concerning language transfer. A Qualtrics survey found that no participants in the present study learned English as a second language.

Second, the learners should all be of advanced proficiency in English. This criterion is vaguely defined and states that proficient students are those who study at a university level and are able to communicate fluently in English (Granger et al., 2009). Those eligible are learners of English who study linguistics, have been taught English for several years, use English for Academic Purposes (EAP), and whose level of English is advanced so as to write an academic text in English.

Third, Granger and Tyson (1996) draw attention to text type. The researchers claim that many features of language are genre-sensitive. To draw reliable conclusions on the research that has been carried out, the text genre ought, therefore, to be consistent. To analyse cohesion, coherence and connector usage, it speaks for itself that the study should be made up of argumentative essays. The compositions may be timed or untimed, with or without reference tools (such as dictionaries, thesauri, grammars, etc.) and should contain 500 words or more.

Fourth, there should be a native speaker corpus of similar writing. In a similar line of argument, it is necessary to compare one's results to a corpus that meets the same criteria as one's own. For this reason, I chose to work with the *Louvain Corpus of Native Essay Writing* (LOCNESS) and the Flemish subcorpus of the *International Corpus of Learner English* (FICLE). For more on these corpora, see 2.8.3 and 2.8.4.

2.8.2. The Corpus of Flemish Learners of English (CoFLE)

CoFLE was compiled for the purposes of this dissertation and consists of 83 essays written by Flemish undergraduate students at the KU Leuven, at the campuses in Brussels and Antwerp. The corpus contains a total of 43,899 word tokens and 4,156 word types. The essay with the lowest number of

words contained 303 of them; the one with the highest word count, 924. Essays contained an average of 517.23 words (SD = 85.692).

The assignment was timed: the students were to complete it within 45 minutes. Before starting to compose their piece, the students were instructed on what they were expected to do, i.e. write an argumentative essay. They had to write a text that was in favour of or against the topic of their choice. The actual purpose of this study was not revealed to the participants, so as not to make them aware of their connector usage. Then the students were presented with the different topics of which they were to pick one (see Appendix 1). Students were told they could use reference tools such as online (translation) dictionaries, thesauri, etc. If the students wanted to obtain more information on the topic, they were also allowed to consult the Internet.

Years of study of participants		
	Frequency	Percent
BA1	46	55.4
BA2	24	28.9
BA3	13	15.7
TOTAL	83	100.0

Table 5: Distribution of essays across BA years

Forty-eight respondents stem from KU Leuven Campus Brussels and are made up of 11 first-year, 24 second-year and 13 third-year Bachelor students in Applied Language Studies (Table 5). The spread across the three phases of the curriculum does not pose a potential problem, as this distribution is a normal one. The higher the curriculum year, the fewer students continue to study a certain course. This distribution therefore reflects real life and does not pose a problem for this study.

Initially, I collected 54 essays, but six had to be removed from the list of eligible participants. In four cases the students were international students from Lithuania or Spain; they could therefore not be integrated in a Flemish learner corpus. Two more did not meet the essay length requirement I set up (see later).

The other 35 participants are from KU Leuven Campus Antwerp and are all first-year Bachelor students in Applied Language Studies. Initially, there were 36 participants, but one participant's essay could not be included due to its shortness. Including students from another campus does not pose problems, either. Their educational backgrounds are the same and they follow the same course. Including these students does not affect the desired degree of homogeneity.

Belgium is a culturally diverse country. Of the 83 participants, 69 considered Dutch to be their first mother tongue (see Table 6). Fourteen other students are bilingual and speak more than one language on a day-to-day basis. The combinations observed are Flemish with French (n = 7), English (n = 1), Spanish (n = 1), Turkish (n = 1), Arabic (n = 1), Urdu (n = 1), Russian (n = 1) and a Tamazight language

(n = 1). Some of the fourteen students indicated speaking Dutch as their other first language, others did not³. Still, all fourteen have been educated throughout their lives in Flemish schools. By having graduated in Flemish secondary schools, they are deemed fit to enrol in Flemish universities. Consequently, I consider them valid respondents for the present study.

First mother tongue		
	Frequency	Percent
Dutch	69	83.1%
French	7	8.4%
English	1	1.2%
Spanish	1	1.2%
Turkish	1	1.2%
Arabic	1	1.2%
Russian	1	1.2%
Urdu	1	1.2%
Tamazight language	1	1.2%
TOTAL	83	100%

Table 6: Distribution of first mother tongues

Regarding the length of the essays, I took some liberty to allow essays which contained slightly fewer than 500 words. While encoding the data, I noticed that some of the essays did not strictly meet the 500-word minimum, but their content still rendered them usable for the current analysis. Having read the essays, I could see that the text was structured properly and that students had, within the 45 minutes, written as much as they could. Since the mean length of the essays across the corpus still exceeded 500, I do not consider the shorter essays to be problematic to this study. Three essays were nevertheless removed: the three participants (CoFLE-KUL-BRU-013, CoFLE-KUL-BRU-018, CoFLE-KUL-ANT-035) produced essays that were respectively 121, 236 and 248 words long, which I deemed too inadequate to be included in the learner corpus.

2.8.3. The Flemish-Dutch subcorpus of ICLE

The Dutch component of the *International Corpus of Learner English* was compiled by professors from the University of Nijmegen, Vrije Universiteit Amsterdam, Rijksuniversiteit Groningen and Katholieke Vlaamse Hogeschool-Antwerpen. The latter is a Belgian university, whereas all others are Dutch universities.

³ In the survey, some respondents did not indicate Dutch as their first language. After conversing with them, I ascertained that all actually *do* speak Dutch to a level of native control. Because these participants are plurilingual (sometimes speaking five or six languages fluently), they chose to indicate the foreign languages they speak (e.g. Moroccan, Punjabi, Chechen or a Ghanaian language) instead of Dutch, which, again, they master.

The Dutch subcorpus of ICLE contains 248 argumentative essays which total 210,501 words, or an average of 849 words per essay. The subcorpus is further divided into Dutch from Belgium (151 essays, 87,079 words which equal 41.4% of the total Dutch subcorpus) and Dutch from The Netherlands (94 essays, 121,730 words which equal 57.8% of the total Dutch subcorpus).

Initially, the Dutch from Belgium (or Flemish) component of the ICLE (or FICLE) contained another three compositions written by students from the Université Catholique de Louvain, located in French-speaking Louvain-la-Neuve. Even though these students speak Dutch as their mother tongue, they have benefited from a different educational framework than those who read at Flemish universities. The three students (whose essays are in total 1,692 words long, which equal 0.80% of the total Dutch subcorpus) will therefore, not be included in the present study. The same rule applies to the Dutch students. They, too, benefited from a different education than Flemish learners; their data does not reflect the years of education Flemish university students have undergone.

While data from Flemish learners of English has indeed been gathered in ICLE, it has not been, to the best of knowledge, presented in a study that solely focussed on that target group. The lack of research on Flemish university learners of English could allow me, as was previously explained in Section 2.1, to look into the data of both my own corpus and that of the Dutch component of the ICLE. From there on, results or further hypotheses might be drawn from the similarities and/or discrepancies that will emerge between corpora that contain texts from students with a similar learning curve of English.

2.8.4. The Louvain Corpus of Native English Essays (LOCNESS)

LOCNESS is a corpus consisting of native English essays. Its goal is to serve as a control corpus for the ICLE. Because the two corpora contain compositions that meet the same criteria, they are fit to be juxtaposed and used for comparison.

In total, the native English essays corpus consists of 326,838 words, of which 150,591 are argumentative essays written by American university students. Another 18,826 words come from 16 compositions on literary subjects written by American students. These 16 essays will not be included in my selection: they are defined by the makers of LOCNESS as being mixed (i.e. some are argumentative, but not all). Another 96,888 words make up the British university students' essays (LOCNESS-UK-U). Most of these are expository essays on literary topics; 33 essays, which make up 19,048 words, are argumentative. Lastly, 60,428 words are from British A-level argumentative essays (LOCNESS-UK-A), which are not suited for comparison within the scope of this study. The reason why is explained below.

For the present research, the American-written *argumentative* essays (not essays treating literary topics) of the native speaker corpus will be used, as well as the part of the British component of LOCNESS which contains argumentative essays. I do not believe that combining British English and American English essays will influence my results, since the aim of this study is to compare essays written by foreign language learners to those composed by native learners of English. The country of origin of the native speakers is therefore of no importance.

The American subcorpus, then, contains 150,591 words from 175 argumentative essays, or an average of 861 words per essay. The collected compositions originate from various universities: Marquette University (WI), Indiana University at Indianapolis (IN), Presbyterian College (SC), the University of South Carolina (SC) and the University of Michigan (MI). The section of the British component of LOCNESS (LOCNESS-UK-U) that will be included in the native-speaker corpus contains 33 argumentative essays and contains 19,048 words. On average, therefore, each essay contains roughly 577 words. For most of the essays taken from LOCNESS-US, the following properties are true: they are untimed essays of 500 words or longer and students were allowed to use reference tools. The 33 argumentative essays from LOCNESS-UK-U were not strictly timed and students used no reference tools.

I reject most of the British components of LOCNESS for a couple of reasons. First, a part of the British subcorpus is made up of A-level argumentative essays. While these learners are only one year younger than the youngest participants in my study, I withhold them from LOCNESS on the grounds that they have not yet enjoyed some of the extra knowledge offered at university. Age is not the discriminating factor here; rather, the formal distinction between secondary and tertiary education is. Second, most of the texts written by British university students pertain to either literature or argumentation. While the latter is relevant to this study, the former is not and will therefore be disregarded. To achieve a higher degree of homogeneity between the native and non-native-speaker corpus, level of education and topic, among other things, may not influence the results.

	LOCNESS-UK-U	LOCNESS-US	TOTAL
Number of essays	33	175	208
Number of words	19,048	150,591	169,639

Table 7: Distribution of words and essays in my selection of LOCNESS

2.8.5. The retrieval software used

For the present study, I decided to compile the bundle of essays into a corpus using the concordance software AntConc (Version 3.5.7; Anthony, 2018). Since it is a free-to-use software, it was chosen over other similar programmes like WordSmith Tools. In an email exchange, professor Sylviane Granger told me that AntConc allows for the same type of extraction as the pay-to-use WordSmith Tools (S. Granger, personal communication, December 15, 2018). AntConc allows the user to easily generate word, keyword and concordance lists. The latter makes it possible to extract connectors manually and observe them in context. These features are also to be found in WordSmith Tools (Yeqing, 2014).

2.9. Statistical testing

I opted to use log-likelihood (LL) to calculate statistical significance. Conventional methods such as Pearson's chi-squared test do not take corpus size into consideration, which could possibly result in skewed data and false positives. Log-likelihood allows corpus size into its calculations, effectively

making it a suitable alternative to a chi-square value. This effective method of calculating statistical significance is prevalent in previous corpus linguistics research (e.g. Bikeliené, 2008b; Buysse, 2012; Narita, Sato & Sugiura, 2004)

Log-likelihood is calculated on the online wizard provided by Paul Rayson (University of Lancaster). On his website, Rayson writes that “the higher the G2 value, the more significant is the difference between two frequency scores” (Rayson, n.d., “Log-likelihood and effect size calculator”). The higher the G2 (or LL) value, the higher the statistical significance. To reach the minimal percentile for statistical significance ($p < 0.05$), the LL value needs to be equal to or greater than 3.84. These are the minimum LL values required per percentile, according to Rayson (n.d.):

- 95th percentile; 5% level; $p < 0.05$; critical value = 3.84
- 99th percentile; 1% level; $p < 0.01$; critical value = 6.63
- 99.9th percentile; 0.1% level; $p < 0.001$; critical value = 10.83
- 99.99th percentile; 0.01% level; $p < 0.0001$; critical value = 15.13

Results

3.1. General study of adverbial connectors

3.1.1. Overall frequencies of adverbial connectors

Based on the selection of connectors made (see Section 2.4), a total of 651 connectors were identified in the learner corpus (CoFLE) and 1,307 can be found in the native-speaker corpus (LOCNESS). FICLE contains 1,222 connectors. To enable comparison between the three corpora, the adjusted frequency was calculated. Learners in CoFLE make use of 148.29 connectors per 10,000 words; native speakers participating in LOCNESS employ 77.05 connectors and learners in FICLE, 140.33. These results are also shown in Table 8.

	CoFLE	LOCNESS	FICLE
Total word tokens	43,899	169,639	87,079
Total connector frequency	651	1,307	1,222
Adjusted frequency per 10,000 words ⁴	148.29	77.05	140.33

Table 8: Raw frequency and adjusted frequency of connectors in the three corpora

Both the Corpus of Flemish Learners of English (adj.f = 148.29) and FICLE (adj.f = 140.33) contain a higher relative frequency of connectors than LOCNESS (adj.f = 77.05). This density of connectors is almost twice as high in the learner corpora than in the native one. There is a statistically significant difference between the connector frequency in CoFLE and that in LOCNESS ($p < 0.001$; LL 170.98) and between the connector frequency in FICLE and that in LOCNESS ($p < 0.001$; LL 222.27). This confirms the hypothesis that Flemish learners of English use connectors with a greater frequency than their native counterparts. There is no significant difference between the two learner corpora (LL 1.29).

It results that Flemish learners of English use more connectors in their essays than do native English learners. This may arise from a desire by foreign language learners to take control of their writing by explicitly writing down connectors. There are several ways to achieve cohesion without the use of connectors; it seems that Flemish learners nonetheless favour them as a way to confidently ensure textual cohesion in their compositions. These results are in line with previous studies which analysed connector usage in EFL learners (e.g. Ha, 2016; Hajjyan, 2015; Tapper, 2005).

3.1.2. Frequency of connectors per semantic category

After looking at the overall frequency with which connectors appear in the three corpora, the frequency of connectors in the different semantic categories was analysed. As can be seen in Table 9, connectors from all semantic categories are present in the writings of Flemish learners of English (both in CoFLE as in FICLE) and in the native corpus.

⁴ The adjusted frequency will henceforth always be per 10,000 words

	CoFLE			LOCNESS			FICLE		
	Rf	Adj.f	% ⁵	Rf	Adj.f	%	Rf	Adj.f	%
Listing & additive	179	40.78	27.50%	288	16.98	22.04%	359	41.23	29.38%
Appositive	56	12.76	8.60%	133	7.84	10.18%	88	10.11	7.20%
Summative	41	9.34	6.30%	59	3.48	4.51%	36	4.13	2.95%
Contrastive	234	53.30	35.97%	545	32.13	41.70%	475	54.55	38.87%
Resultive	94	20.73	13.98%	170	10.02	13.01%	161	18.49	13.18%
Corroborative	50	11.39	7.68%	112	6.60	8.57%	103	11.83	8.43%
TOTAL	651	148.29	100%	1,307	77.05	100%	1,222	140.33	100

Table 9: Raw and adjusted frequencies of connectors per semantic category

In all three corpora, the **contrastive** connectors are used most frequently. In CoFLE their adjusted frequency totals 53.30 (equalling 35.97%, or over one third of all connectors). In FICLE the adjusted frequency is slightly higher, equalling 54.55 (38.87% of all connectors present). In the native speaker corpus, then, contrastive connectors make up 32.13 words in every 10,000 (or 41.70% of all connectors). Contrastive connectors appear significantly more in CoFLE ($p < 0.0001$; LL 38.96) and in FICLE ($p < 0.0001$; LL 69.48) than they do in LOCNESS.

Although the share of contrastive connectors in LOCNESS is strikingly higher than those of the two learner corpora, the frequency per 10,000 words is approximately one and a half times smaller. This is because, as indicated in Section 3.1.1, connectors in essays written by native speakers are less densely present than in those written by Flemish learners of English. This difference is also notable in for instance listing and additive, corroborative or appositive connectors (see below). Looking at the adjusted frequencies offers depth to the results which the percentages alone could not do. It allows us to juxtapose corpora and indicate usage difference between the three.

The second-most popular category of connectors, again across all three corpora, is the **listing and additive** one. In CoFLE it accounts for 40.78 connectors in every 10,000 words and 27.50% of all connectors used. A nearly similar frequency is found in the other learner corpus, FICLE, where listing and additive connectors make up 29.38% of all connectors and are represented 41.23 times every 10,000 words. The difference between the two learner corpora is not significant (LL 0.01). In LOCNESS the percentage is somewhat lower, at 22.04%. The adjusted frequency of these connectors in the native speaker corpus is 16.98. Here, too, Flemish learners of English use significantly more listing and additive connectors than native-speaker learners ($p < 0.0001$; LL 77.17 for CoFLE; $p < 0.0001$; LL 125.79 for FICLE).

⁵ This percentage denotes the share of that semantic category in all the connectors found in either corpus.

Third are the **resultive** connectors. Once more, this category takes the same spot in all three corpora, learner and native speakers alike. In CoFLE resultive connectors are responsible for 13.98% of all connectors used and have an adjusted frequency of 20.73. For learners in FICLE, this adjusted frequency is not significantly lower (LL 1.26), equalling 18.49 and amounting to 13.18% of all connectors used. In LOCNESS, resultive connectors occur 10.02 times in every 10,000 words; they constitute 13.01% of all connectors found in that corpus. Once more, we find a statistical significance in the usage of resultive connectors between participants in CoFLE ($p < 0.0001$; LL 31.86) and FICLE ($p < 0.0001$; LL 30.38) compared to native speakers in LOCNESS.

The fourth most frequently used semantic category of connectors in CoFLE is the **appositive**, which has an 8.60% share in all the connectors found in the compositions. Appositive connectors occur 12.76 times every 10,000 words. A similar share is found in FICLE, where these connectors make up 7.20% of the connectors found there; the adjusted frequency is 10.11. Native speakers, conversely, use fewer appositive connectors than learners. LOCNESS counts a lower adjusted frequency ($\text{adj.f} = 7.84$), though these connectors make up a larger percentage of the total number of connectors (10.18%). Despite the smaller frequency variation, participants in CoFLE still use significantly more connectors than those in LOCNESS ($p < 0.01$; LL 8.69). No statistically significant difference was found between FICLE and LOCNESS in this category (LL 3.35).

Corroborative connectors take the fifth place in the Corpus of Flemish Learners of English, though this position differs in the other corpora. With an adjusted frequency of 11.39 and a 7.68% share in the total connector count, it sits just above the summative and just below the appositive categories. The differences in raw frequency are small, too: 50 corroborative connectors to 56 appositive ones. In FICLE corroboration features slightly more often, appearing 11.83 times in every 10,000 words, totalling 8.43% of all connectors. The percentage found in LOCNESS is practically similar: 8.57%. The adjusted frequency in the native speaker corpus, however, is lower, reaching 6.60 occurrences per 10,000 words. The higher adjusted frequencies found in CoFLE ($p < 0.01$; LL 9.51) and FICLE ($p < 0.0001$; LL 17.85) are significantly higher than that of LOCNESS. The observed frequencies in the learner corpora are not significantly different from one another (LL 0.05).

Finally, the **summative** category is the one which is least used in CoFLE. In that corpus, these connectors have an adjusted frequency of 9.34; the total number of connectors consists of 6.30% of those in the summative category. For FICLE, that percentage is slightly lower, at 2.94%, with an adjusted frequency of 4.13. This frequency is lower still in LOCNESS, being 3.48. Summative connectors have a 4.51% share in the total number of connectors present in the native speaker corpus. They appear significantly less in the NS corpus than they do in CoFLE ($p < 0.0001$; LL 21.50), but the difference is not significant when compared to FICLE (LL 0.66). Summative connectors are the connectors where a statistically significant discrepancy can be found between CoFLE and FICLE ($p < 0.001$; LL 12.61).

3.2. Specific study of adverbial connectors

3.2.1. Most frequently used connectors

CoFLE	Rf	% ⁶	LOCNESS	Rf	%	FICLE	Rf	%
But	70	10.75%	However	206	15.76%	But	290	23.73%
However	63	9.68%	Also	112	8.57%	And	124	10.15%
Also	46	7.07%	(Al)though	109	8.34%	Also	100	8.18%
And	41	6.30%	But	102	7.80%	So	85	6.96%
For example	38	5.84%	For example	68	5.20%	Of course	60	4.91%
Therefore (resultive)	29	4.45%	And	48	3.67%	However	40	3.27%
Of course	27	4.15%	Therefore (resultive)	42	3.21%	(Al)though	34	2.78%
So	26	3.99%	So	41	3.14%	For example	34	2.78%
On the one / other hand	25	3.84%	In fact	40	3.06%	Therefore (resultive)	32	2.62%
(Al)though	20	3.07%	Yet	35	2.68%	On the one / other hand	31	2.54%
In conclusion	15	2.30%	Of course	30	2.30%	First of all	22	1.80%
Yet	13	2.00%	On the one / other hand	26	1.99%	First(ly)	22	1.80%
Second(ly)	12	1.84%	Then (listing and additive)	20	1.53%	Indeed	21	1.72%
Moreover	12	1.84%	Now	20	1.53%	For instance	20	1.64%
First(ly)	11	1.69%	Therefore (summative)	17	1.30%	Then (listing and additive)	16	1.31%
TOTAL	448	68.82%	TOTAL	916	70.08%	TOTAL	931	76.19%

Table 10: The 15 most frequently used connectors in all three corpora

In CoFLE five out of the fifteen most popular connectors are **contrastive**: *but*, *however*, *on the one / other hand*, *(al)though* and *yet*. These five connectors are the first, second, ninth, tenth and twelfth most commonly found connectors in CoFLE. Together, they account for 29.34% — or just over a quarter — of all connectors found in that corpus.

FICLE contains one contrastive connector fewer in the list of the most used ones, i.e. *yet*. The connector *but* in itself makes up 23.73% of all connectors used. In this learner corpus, it became apparent that, contrary to stylistic guidelines and preferences in the English language (which states that

⁶ This percentage represents that connector's share of all the connectors in that corpus.

sentences should not be prefaced by coordinating conjunctions), *but* is often used by learners as a connector. In total, these four connectors add up to 32.32% of all connectors used.

LOCNESS contains the same five contrastive connectors in its top 15 as CoFLE. The five connectors together add up to 36.57% of all connectors used.

Interesting but not surprising to note is that the most-used connector in all three corpora is a contrastive one. Because of the intrinsic structure of an argumentative essay, contradiction is likely to appear more often than other ways to express a nuanced opinion.

Next to the contrastive category, **listing and additive** connectors form a recurrent category among the fifteen most found connectors. As was expounded in Section 3.1.2, the presence of these two categories is down to the nature of essays written in the argumentative mode. It is not surprising, therefore, to observe that in CoFLE, ten connectors in this category are either contrastive ($n = 5$) or listing and additive ($n = 5$). FICLE, counts nine of its fifteen most common connectors in either the contrastive ($n = 4$) or the listing and additive ($n = 5$) category. LOCNESS makes do with eight (three listing and additive, five contrastive).

In CoFLE the additive connectors to be found are *also*, *and*, *second(ly)*, *moreover* and *first(ly)*. These connectors respectively feature in third, fourth, thirteenth, fourteenth and fifteenth position of occurrence and add up to 18.74% of all connectors in the Flemish learners' compositions.

FICLE equally contains five additive connectors, though not the same as the other learner corpus. These are, in decreasing order of popularity: *and*, *also*, *first of all*, *first(ly)* and *then*, amounting to 23.24% of all connectors used.

LOCNESS, then, contains three additive connectors in its top 15 ranking: *also*, *and* and *then*, which have a 13.77% share of all the connectors used. Connectors such as *first(ly)*, *second(ly)* and *moreover* do not appear in top 15 of the native-speaker corpus, though they are present in the non-native-speaker corpora. The great difference in frequency results in connector overuse by Flemish learners of English (see Section 3.2.2).

Resultive connectors, albeit appearing fewer times in the list of fifteen most used connectors than the two previous categories, still seem to be used frequently. In CoFLE two such connectors emerge: *therefore* and *so*. Together, they make up 8.45% of all connectors used, manifesting themselves 29 and 26 times respectively.

FICLE contains the identical pair of resultive connectors, namely *so* and *therefore*. *So* features 85 times in total, *therefore*, 32. The two make up 9.57% of all connectors to be found in this corpus.

The native speaker corpus contains three resultive connectors among the most frequently used ones. They are *so*, *therefore* and *now*. *So* appears 41 times; *therefore* occurs one fewer time, 42; *now* is only used 20 times. Their share in the totality of connectors used stands at 7.88%.

Therefore not only features as a frequently-used resultive connector in all three corpora, but it also exists as a common **summative** connector in LOCNESS. Although FICLE contains connectors from that category, too, none of them have made it onto its list of most common connectors.

In CoFLE *in conclusion* is the only summative connector in the top 15 most-used connectors. It appears 15 times and accounts for 2.30% of all the connectors in this corpus.

In LOCNESS *therefore* is the only connector from the summative category present among the fifteen most popular connectors. It appears fifteenth in that list and occurs 17 times, or 1.30% of all connectors used.

The second-to-last category to appear in the most frequently used connectors is the **corroborative** one. While both FICLE and LOCNESS count two such connectors in their top 15, CoFLE only has one: *of course*. The corroborative connector *of course*, moreover, is found in all three corpora, in fifth (FICLE), seventh (CoFLE) and eleventh (LOCNESS) place. In CoFLE *of course* occurs 27 times, accounting for 4.15% of all connectors. FICLE has 60 occurrences of that connector and another 21 of the corroborative connector *indeed*. Together, they are responsible for 6.63% of all connectors used in this corpus. The corroborative connectors *of course* and *in fact*, both present in the top 15 most-used connectors in LOCNESS, occur 70 times, or 5.36% of all connectors.

Lastly, **appositive** connectors in the top 15 appear once each in CoFLE and LOCNESS and twice in FICLE. In all three corpora, we find *for example*; *for instance* is only present in FICLE.

In the Corpus of Flemish Learners of English, *for example* is present 38 times, adding up to 5.84% of the total number of connectors. FICLE counts 34 instances of *for example* and 20 of *for instance*; in total, they make up 4.42% of all connectors for that corpus. LOCNESS has 68 instances of *for example*, which means 5.20% of its connectors.

In total, the fifteen most frequently found connectors account for the majority of connectors that are found in the two learner corpora and the native speaker one. The fifteen most popular connectors in CoFLE make up 68.82% of all connectors found in that corpus; for LOCNESS, that percentage lies higher at 70.08%; FICLE finds 76.19% of all its connectors in just fifteen of them.

Of the fifteen connectors found in the Corpus of Flemish Learners of English, eleven featured in the native speaker corpus (LOCNESS); eleven appeared in the other learner corpus, FICLE. This means that there is a 73.34% overlap between the fifteen most-used connectors that are found in CoFLE and LOCNESS, and an equal 73.34% overlap between those of CoFLE and FICLE. The similarity between LOCNESS and FICLE equals 73.34%, too: eleven of the fifteen most popular connectors in the learner corpus were also present in the top 15 of the native-speaker corpus. It must be noted, however, that these percentages only reflect the number of connectors that can be found between several corpora and not their frequency.

Ten out of the fifteen most frequently occurring connectors, or 66.67%, are shared across all three corpora. These are *however*, *also*, *but*, *on the one / other hand*, *for example*, *of course*, *and*, *so*, *therefore* (resultive) and *(al)though*.

While the frequencies differ per corpus (*however*, for instance, has frequencies varying between 3.24% and 15.62%), the similarities that the three corpora present are striking. These similarities are not defined by languages, but come as a result of the mode those essays are set in. In Section 1.3.3.1, I already described the structure of argumentative essays as defined by Hyland (1990), which is found in texts written by native speakers and EFL learners alike. For this reason, connectors which mark addition, contrast or corroboration, for example, are likelier to be used than others. This is why words such as *however*, *of course*, *so*, etc., which are ubiquitous in the English language, are likely to appear in all three corpora.

3.2.2. Overuse and underuse of connectors

To analyse data for this section of the study, I decided, as was mentioned in Section 2.5, to work with cut-off points, after Ha's (2016) paper. For a connector to qualify as being overused, its ratio of occurrences in the non-native speaker corpus to the native speaker corpus is set at 1.5 or higher. We may speak of underuse if that ratio is equal to or lower than 0.95. Certain other connectors did not appear in either the learner corpora or the native-speaker corpus. In these cases, I could apply the same method as above: dividing zero by any number equals zero and dividing by zero is impossible. Those results would not yield any useful information. I consequently decided to consider those connectors to be overused if $\text{CoFLE/FICLE} - \text{LOCNESS} \leq 5$. Conversely, connectors are considered underused if $\text{CoFLE/FICLE} - \text{LOCNESS} \geq -5$.

3.2.2.1. Overuse and underuse per semantic category

In all semantic categories, Flemish learners of English in CoFLE overuse connectors in their academic essays. Overall, then, and as is shown in Table 11, the non-native students overuse connectors when compared to L1 English learners.

	CoFLE	LOCNESS	
	Adj.f	Adj.f	NNS/NS ratio
Listing and additive	40.78	16.98	2.40
Appositive	12.76	7.84	1.63
Summative	9.34	3.48	2.69
Contrastive	53.30	32.13	1.66
Resultive	20.73	10.02	2.07
Corroborative	11.39	6.60	1.73
TOTAL	148.29	77.05	1.92

Table 11: Connector overuse in the semantic categories of CoFLE

These results are in line with what has been observed in Section 3.1.2, i.e. that Flemish learners of English use more connectors across all semantic classifications. By juxtaposing the adjusted

frequencies in Table 11 and using a cut-off point of 1.5, it is visible that Flemish learners of English overuse all categories.

Despite **contrastive** connectors being the most frequently used connectors in all three corpora, it is not, with its ratio of 1.66 just surpassing the required minimum, the most overused connector in CoFLE. That honour goes to the lesser-used **summative** category, which has an adjusted frequency of 9.34 occurrences per 10,000 words, compared to 3.48 in LOCNESS, giving it an overuse ratio of 2.69. Interestingly, the category of summative connectors may occur the least often, yet it is overused the most. This is because the smaller the numbers divided, the greater the fluctuations they will yield. The **corroborative**, with a ratio of 1.73, the **appositive**, with a ratio of 1.63, are also overused by Flemish learners of English. More heavily overused connectors are the **resultive** connectors, with an NNS/NS ratio of 2.07, and the **listing and additive** ones, being used by participants in CoFLE 2.40 times for every occurrence in LOCNESS.

Overall, Flemish learners of English in CoFLE significantly overuse connectors ($p < 0.0001$; LL 170.98): for every connector used in the native speaker corpus, Flemish learners use 1.92 of them. This same trend is reverberated when comparing LOCNESS with FICLE (see Table 12 below). In the latter non-native-speaker corpus, the overall connector overuse ratio is at 1.82 ($p < 0.0001$; LL 222.27). Indeed, the connector density in CoFLE and FICLE are practically similar, being at 148.29 and 140.33 occurrences per 10,000 words respectively. A log-likelihood calculation shows us that the differences in the two adjusted frequencies is not statistically significant (LL 1.29).

	FICLE	LOCNESS	
	Adj.f	Adj.f	NNS/NS ratio
Listing and additive	41.23	16.98	2.43
Appositive	10.11	7.84	1.29
Summative	4.13	3.48	1.19
Contrastive	54.55	32.13	1.70
Resultive	18.49	10.02	1.84
Corroborative	11.83	6.60	1.79
TOTAL	140.03	77.05	1.80

Table 12: Connector overuse in the semantic categories of FICLE

Two categories are not significantly overused in FICLE, namely the **appositive** (LL 3.35) and **summative** one (LL 0.66). In FICLE the **listing and additive** connectors that overused most with an NNS/NS ratio of 2.43. Whereas **resultive** and **corroborative** connectors were only slightly overused in CoFLE, these two categories appear more frequently in the writing of participants in FICLE with overuse

ratios of 1.84 and 1.79 respectively. **Contrastive** connectors also get overused by that group, having a ratio of 1.70.

This study ties in with Tapper's (2005), who theorized that the overuse of **corroborative** connectors may be a universal or interlanguage feature. I have found corroborative connectors to be overused in both CoFLE and FICLE, thus substantiating Tapper's (2005) research.

No underuse is detected in semantic categories of either of the two non-native-speaker corpora.

3.2.2.2. Comparison between CoFLE and LOCNESS

In the Corpus of Flemish Learners of English, 38 individual connectors (see Table 13 below) were found to be overused: all of them had an NNS/NS ratio of 1.5 or more. Of those 38, 11 connectors figured in the listing and additive category, 4 in the appositive, 4 in the summative, 10 in the contrastive, 7 in the resultive and 2 in the corroborative.

Twelve of the fifteen most frequent connectors identified in Section 3.2.1 were also found to be overused by Flemish learners of English. In Table 13, these connectors are in bold. Apart from *(al)though* and *yet*, which are underused (see Table 15), all but one figure in this list. The contrastive connector *however* surprisingly does not appear in the list, as it is neither overused or underused (NNS/NS ratio of 1.18).

		CoFLE	LOCNESS	
		Adj.f	Adj.f	NNS/NS ratio
Listing and additive	lastly	1.37	0.18	7.73
	moreover	2.73	0.35	7.73
	second(ly)	2.73	0.53	5.15
	first(ly)	2.51	0.59	4.25
	furthermore	2.05	0.53	3.86
	first of all	2.51	0.71	3.54
	and	9.34	2.83	3.30
	third(ly)	0.68	0.24	2.90
	besides	1.59	0.65	2.46
	in addition	1.59	0.94	1.69
	also	10.48	6.60	1.59
Appositive	e.g.	0.91	0.12	7.73
	in general	0.68	0.12	5.80
	for example	8.66	4.01	2.16
	specifically	0.23	0.12	1.93
Summative	to conclude	1.37	0.12	11.59
	in conclusion	3.42	0.59	5.80
	therefore	2.28	1.00	2.27
	then	0.91	0.59	1.55

Contrastive	nevertheless	1.82	0.29	6.18
	on the contrary	0.68	0.12	5.80
	conversely	0.23	0.06	3.86
	on the one / other hand	5.69	1.53	3.72
	but	15.95	6.01	2.65
	in/by contrast	0.46	0.18	2.58
	besides	0.23	0.12	1.93
	only	1.14	0.65	1.76
	instead	1.37	0.83	1.66
	still	0.68	0.41	1.66
Resultive	as a consequence	0.91	0.06	15.46
	as a result	1.82	0.71	2.58
	therefore	6.61	2.48	2.67
	so	5.92	2.42	2.45
	now	2.28	1.18	1.93
	thus	1.59	0.94	1.69
	consequently	0.46	0.29	1.55
Corroborative	of course	6.15	1.77	3.48
	indeed	1.82	0.77	2.38

Table 13: Connector overuse in CoFLE

		CoFLE	LOCNESS	
		Rf	Rf	Difference
Contrastive	nonetheless	6	0	+6

Table 14: Connector overuse in CoFLE where Rf in LOCNESS equals zero

Some connectors are found with a high overuse ratio, such as *lastly* (NNS/NS ratio of 7.73), e.g. (7.73), *to conclude* (11.59) and *as a consequence* (15.46). Many others have a smaller overuse ratio, usually between 2 and 4. The connectors which have a high ratio of overuse, are those which in general are not much used in either corpus. In the case of *as a consequence*, the connector has an adjusted frequency of 0.91 in CoFLE to 0.06 in LOCNESS. In the learner corpus, *as a consequence* appears four times; in the native-speaker corpus, once. So, while the NNS/NS ratio is high, the raw and adjusted frequencies are low. By contrast, the contrastive connector *but* has an adjusted frequency of 15.95 in CoFLE and 6.01 in LOCNESS, giving it an NNS/NS ratio of 2.65. As was said earlier, it is normal for smaller number to yield larger fluctuations; this has to be kept in mind when interpreting overuse and underuse data.

One connector, the contrastive *nonetheless*, was found to be used six times in CoFLE, but not in LOCNESS, which indicates overuse of that particular connector.

		CoFLE	LOCNESS	
		Adj.f	Adj.f	NNS/NS ratio
Listing and additive	finally	0.68	0.77	0.89
	likewise	0.23	0.35	0.64
Appositive	namely	0.23	0.35	0.64
	in particular	0.23	0.35	0.64
Contrastive	(al)though	4.56	6.43	0.71
Resultive	hence	0.23	0.47	0.48
Corroborative	in fact	1.59	2.36	0.68
	apparently	0.23	0.35	0.64

Table 15: Connector underuse in CoFLE

		CoFLE	LOCNESS	
		Rf	Rf	Difference
Appositive	that is (to say)	0	8	-8
	particularly	0	6	-6
Contrastive	then	0	8	-8

Table 16: Connector underuse in CoFLE where Rf in CoFLE equals zero

CoFLE contains 8 connectors which are underused when compared to the native-speaker corpus (see Table 15). Of those eight, one connector, *(al)though*, figures among the fifteen most popular connectors in CoFLE and has an underuse ratio of 0.71. The lowest underuse value is that of the contrastive connector *hence*, appearing only 0.48 times in CoFLE for every occurrence in LOCNESS. In other words, it is used twice as much by native learners of English than Flemish learners.

The connectors *that is (to say)*, *particularly* and the contrastive *then* were found to be underused, too, as they do not appear in CoFLE but are used eight times (or six for *particularly*) in LOCNESS.

3.2.2.3. Comparison between FICLE and LOCNESS

In FICLE 30 connectors are identified as being overused. Of those 30, 9 belong to the listing and additive category, 4 to the appositive, 2 to the summative, 9 to the contrastive, 4 to the resultive and 2 to the corroborative. The full list is available in Table 17.

Eleven of the fifteen most-used connectors in FICLE, identified in Section 3.2.1, were found to be overused. They have been made bold in Table 17. Another two were underused and two more did not appear to be either overused or underused. (Those are *for example* and the resultive *therefore*, which, with an NNS/NS ratio of 0.97 and 1.48 respectively, did not make the threshold as underused or overused connectors).

		FICLE	LOCNESS	
		Adj.f	Adj.f	NNS/NS ratio
Listing and additive	and	14.24	2.83	5.03
	moreover	1.61	0.35	4.55
	first(ly)	2.53	0.59	4.29
	first of all	2.53	0.71	3.57
	second(ly)	1.84	0.53	3.46
	besides	1.49	0.65	2.30
	finally	1.61	0.77	2.10
	also	11.48	6.60	1.74
	then	1.84	1.18	1.56
Appositive	e.g.	0.57	0.12	4.87
	in general	0.57	0.12	4.87
	in other words	1.03	0.24	4.38
	for instance	2.30	0.88	2.60
Summative	to conclude	0.92	0.12	7.79
	then	1.26	0.59	2.14
Contrastive	on the contrary	1.15	0.12	9.74
	but	33.30	6.01	5.54
	nevertheless	1.26	0.29	4.29
	still	1.49	0.41	3.62
	on the one / other hand	3.56	1.53	2.32
	instead	1.72	0.83	2.09
	meanwhile	0.11	0.06	1.95
	then	0.92	0.47	1.95
	besides	0.23	0.12	1.95
Resultive	as a consequence	0.34	0.06	5.84
	so	9.76	2.42	4.04
	after all	1.38	0.71	1.95
	accordingly	0.11	0.06	1.95
Corroborative	of course	6.89	1.77	3.90
	indeed	2.41	0.77	3.15

Table 17: Connector overuse in FICLE

		FICLE	LOCNESS	
		Rf	Rf	Difference
Listing and additive	in the first place	7	0	+7

Table 18: Connector overuse in FICLE where Rf in LOCNESS equals zero

Here, too, the importance of the overuse ratio differs from one connector to the next. Some have a high NNS/NS ratio, such as *on the contrary* (9.74), *to conclude* (7.79) and *as a consequence* (5.84). Similarly to what was found in 3.2.2.2, the connectors with the highest overuse ratio are those that do not occur many times in FICLE or LOCNESS. In the case of the contrastive connector *on the contrary*, we see that it appears 10 times in FICLE (adj.f = 1.15) and twice only in LOCNESS (adj.f = 0.12). It is difficult to draw conclusions on the overuse of those connectors, since they do not appear frequently enough in either corpus. The exception to this is *but*, which, as the most used connector in FICLE, appears 33.3 times per 10,000 words. In LOCNESS it is also a popular connector, where it has an adjusted frequency of 6.01, or about five times smaller than in the learner corpus. As a consequence, its NNS/NS ratio is 5.54. The same is true for *and*, which has an overuse ratio of 5.03, meaning that participants in FICLE use the additive connector five times more often than participants in the native-speaker corpus.

		FICLE	LOCNESS	
		Adj.f	Adj.f	NNS/NS ratio
Listing and additive	third(ly)	0.11	0.24	0.49
Appositive	i.e.	0.57	0.65	0.89
	particularly	0.23	0.35	0.65
	that is (to say)	0.23	0.47	0.49
	in particular	0.11	0.35	0.32
Summative	in conclusion	0.34	0.59	0.58
	therefore	0.57	1.00	0.57
	thus	0.46	0.88	0.52
Contrastive	rather	0.46	0.59	0.78
	in/by contrast	0.11	0.18	0.65
	at first	0.11	0.18	0.65
	(al)though	3.90	6.43	0.61
	however	4.59	12.14	0.38
	yet	0.57	2.06	0.28
Resultive	thus	0.69	0.94	0.73
	now	0.69	1.18	0.58
	hence	0.23	0.47	0.49
Corroborative	admittedly	0.11	0.18	0.65
	in fact	1.49	2.36	0.63
	actually	0.34	0.94	0.37

Table 19: Connector underuse in FICLE

		FICLE	LOCNESS	
		Rf	Rf	Difference
Listing and additive	likewise	0	6	-6
	in addition	0	16	-16

Table 20: Connector underuse in FICLE where Rf in FICLE equals zero

FICLE underuses 20 connectors, more than CoFLE does (see Table 19). Of those, two figured in their top 15 most-used connectors: *(al)though*, with an NNS/NS ratio of 0.61, and *however*, with an underuse ratio of 0.38. *However* figures as the fourth-most underused connector in FICLE. Other connectors which have an even lower NNS/NS ratio are *actually* (0.37), *in particular* (0.32) and *yet*, which, with an underuse ratio of 0.28, is the most underused of all connectors in this corpus.

Likewise and *in addition* are both observed to be underused by participants in FICLE: they do not appear at all in that corpus, but do in LOCNESS, six and sixteen times respectively.

3.2.3. The Academic Keyword List in CoFLE

In Section 2.7, Paquot's (2010) *Academic Keyword List* and its relevance for this study were discussed. The AKL enables me to look at the connectors that the Flemish learners of English have used and judge whether they are suited for academic purposes. Using the AKL is not a way to judge essay quality, but merely a way to observe language register sensitivity.

Twenty-six connectors from the AKL appear in LOCNESS; 23 in CoFLE and 21 in FICLE (see Table 21). The AKL connectors in the listing and additive category have a total frequency of 45 in CoFLE, 53 in FICLE and 46 in LOCNESS and make up 25.14%, 14.76% and 15.97% respectively of the total number of connectors in that semantic category. The percentage of appositive connectors in the corpora that also figure in the AKL is 94.64% in CoFLE, 79.55% in FICLE and 82.71% in LOCNESS. For summative connectors, that is 36.59% (CoFLE), 25% (FICLE) and 54.24% (LOCNESS); for contrastive connectors, 38.03% (CoFLE), 16.84% (FICLE) and 60% (LOCNESS); for resultive connectors, 45.05% (CoFLE), 27.33% (FICLE) and 42.35% (LOCNESS); for corroborative connectors, finally, 16% (CoFLE), 20.39% (FICLE) and 11.61% (LOCNESS).

The numbers above show that the connectors, marked up by Paquot (2010) as potential academic keywords, are found in the writings of Flemish learners of English. In some categories, especially the appositive one, that number was high; in others, such as listing and additive, that number was lower. The discrepancy has to do with how varied the connectors are in each category. The contrastive and listing ones are by far the categories with the most types of connectors, while there are fewer words to express apposition. As a result, those connectors which do convey apposition, because of their smaller numbers, are more likely to be found in the AKL than contrastive connectors, of which dozens exist, explaining the high percentage in the former and the relatively lower in the latter.

In total, the connectors found in Paquot's (2010) AKL make up 38.25% of all connectors found in CoFLE, 22.67% of all those found in FICLE and 45.91% of all those found in LOCNESS. I expected Flemish

learners of English to employ several connectors found in the AKL and do so frequently. It is not surprising to find that other connectors still make up the majority of connectors present in the corpora. While Paquot's (2010) *Academic Keyword List* is a handy tool to assess whether learners use academic words in their writing, it is more difficult to consider certain connectors to be academic. Throughout their secondary education, Flemish (and by extension other foreign) learners of English will learn to write cohesive texts and linking ideas, sentences and paragraphs through connectors. They will be taught a basic set, allowing them to express their thoughts coherently. It is only as their fluency grows that their vocabulary will, too. Flemish learners at the tertiary level of education (i.e. university and college level) will acquire more words still, which, because of their precise meaning, will be used parsimoniously. They will fit an academic context well and are required to create a coherent complex text. These newly-acquired connectors (such as *hence* and *for instance*) are not words that get used often, though connectors such as *however*, *also*, *for example*, *therefore*, etc. will appear more frequently. Academic texts (or in this case, argumentative essays) will contain a good deal of connectors found in Paquot's (2010) *Academic Keyword List*, but the less frequent a word in the English language, the less it will be found in writings, even if academic.

The AKL is a useful tool to reflect the degree to which a text may be said to be academic, but it fails to do so when the only things one looks at are connectors. Connectors are elementary parts of a text, its building blocks. Few connectors can be seen as niche or advanced. Indeed, most are known to learners of different proficiencies. Teaching students more connectors is useful only to increase their vocabulary and avoid repetition, but picking connectors from the AKL will not render a text more academic per se.

As discussed in Section 2.7 these texts are not strictly academic. The fact that these learners of English use 26 different connectors which are also present in Paquot's (2010) AKL shows that the students have the required register sensitivity when writing argumentative essays. They know when to adapt their vocabulary and adopt a greater variety of words to express themselves. When comparing the essays written by Flemish EFL learners to those from LOCNESS, it appears that native speakers use more connectors from AKL, indicating they possess a higher register sensitivity still.

		CoFLE	FICLE	LOCNESS
Listing and additive	first(ly)	11	22	10
	second(ly)	12	16	9
	third(ly)	3	1	4
	moreover	12	14	6
	in addition	7	0	16
	above all	0	0	1
Appositive	namely	1	3	6
	for instance	4	20	15
	for example	38	34	68
	e.g.	4	5	2
	in general	3	5	2

	generally	1	0	3
	particularly	0	2	6
	in particular	1	1	6
	specifically	1	0	2
Summative	therefore	10	5	17
	thus	5	4	15
Contrastive	conversely	1	0	1
	however	63	40	206
	only	5	6	11
	(al)though	20	34	109
Resultive	therefore	29	32	42
	thus	7	6	16
	consequently	2	3	5
	accordingly	0	1	1
	hence	1	2	8
Corroborative	indeed	8	21	13
TOTAL		249	277	600

Table 21: Connectors of the Academic Keyword List in the learner corpora

3.3. Syntactic positioning of connectors

In Section 2.6 a distinction was made between connectors that link up sentences and those that connect clauses. In some cases, it was argued, a connector which does not figure at the beginning of a sentence but at the start of a clause within a sentence, may still be considered a connector. For that reason, I have chosen to analyse the connector in initial (IP), clause-medial (CMP) and clause-final (CFP) positions, instead of the habitually more prevalent sentence-initial (SIP), sentence-medial (SMP) and sentence-final (SFP) syntactic positions. The initial position will further be divided into sentence-initial position (SIP) and clause-initial position (CIP).

	CoFLE		LOCNESS		FICLE	
	Freq.	% ⁷	Freq.	%	Freq.	%
IP	486	74.65%	972	74.37%	921	75.37%
CMP	157	24.12%	327	25.02%	282	23.08%
CFP	8	1.23%	8	0.61%	19	1.55%
TOTAL	651	100%	1,307	100%	1,222	100%

Table 22: Frequency of syntactic positioning of connectors in the three corpora

⁷ This percentage is the share which a certain syntactic positioning has of all the connectors found.

The syntactic positioning of the connectors in the three corpora is shown in Table 22. As is visible, learners of English, native speakers and EFL students alike, overwhelmingly place connectors in the initial position. The results are consistent across all three corpora.

In the Corpus of Flemish Learners of English, connectors occur 486 times in initial position (IP), which equals 74.65% of all the connectors used. Another 24.12%, or 157 occurrences, are placed in clause-medial position (CMP). Only 8 connectors find themselves in clause-final position (CFP), amounting to 1.23% of all connectors found in that corpus.

Of the 1,222 connectors found in FICLE, 921 instances, or 75.37%, are found in IP. The clause-medial position occupies 23.08% of all connectors in that corpus, counting 282 occurrences. Here, too, the CFP represents a small percentage, at only 1.55% (or 19 connectors).

The native speaker corpus LOCNESS, then, contains 972 connectors which are used in IP; this equals 74.37% of all connectors. The medial position is used 327 times, or 25.02% of all connectors found in the corpus. The final position has again a small share, with 8 occurrences, or 0.61%.

Choosing where to place a connector, is not always a matter of preference. Some connectors can only function in IP; when placed elsewhere, they perform another function or have another meaning. Of those included in my analysis, these are *first(ly)*, *second(ly)*, *third(ly)*, *next*, *to begin with*, *in the same way*, *and*, *what is more*, *that is (to say)*, *i.e.*, *e.g.*, *to conclude*, *briefly*, *but*, *anyway*, *yet*, *still*, *(al)though*, *so*, *now* and *in that case*. Looking at the distribution of syntactic positioning of connectors which can adopt more than one position offers a more nuanced image across all three corpora. The consequent results are found in Table 23⁸.

	CoFLE		LOCNESS		FICLE	
	Freq.	% ⁹	Freq.	%	Freq.	%
IP	278	64.80%	604	67.56%	332	55.43%
CMP	143	33.33%	283	31.66%	251	41.90%
CFP	8	1.86%	7	0.78%	16	2.67%
TOTAL	429	100%	894	100%	599	100%

Table 23: Frequency of syntactic positioning of connectors after revision

The results from Table 22 confirm our hypothesis that Flemish learners of English mostly place connectors in initial position. What's more, by eliminating from our calculations those connectors which can only figure in IP, we observe that the initial position remains the preferred spot for Flemish and native learners alike to place their connectors. The proportion of connectors in IP does, however, diminish in Table 23. This phenomenon is all too normal, since we have removed high-frequency connectors which were only used in IP.

⁸ Henceforth, calculations will, unless mentioned otherwise, be made with the numbers from Table 22 (i.e. those containing all connectors). They represent a truer image of connector usage by Flemish EFL learners.

⁹ This percentage is the share which a certain syntactic positioning has of all the connectors found.

This research confirms the theory that the syntactic positioning of connectors is not language-specific; rather, it seems that this feature is universal. This study further highlights the fact that native speakers of English preponderously use connectors in IP, just like Flemish EFL learners. It would therefore seem that both foreign learners of English and native learners alike feel placing connectors in the clause-initial position is a “safe bet” (Paquot, 2010, p. 179).

Tables 25, 26 and 27 show the syntactic positioning of the fifteen most frequently used connectors in all three corpora¹⁰. Because this list contains the majority of all connectors for the three corpora (see Table 10 in Section 3.2.1), it is to be expected that the distribution of frequencies among the syntactic positions of the connectors in that list resembles the frequencies found in the corpora in their entirety. Indeed the fifteen most frequently found connectors in CoFLE are distributed across the syntactic positions as follows: 77.45% in IP (n = 347), 22.10% in CMP (n = 99) and 0.45% in CFP (n = 2). Compare those to the subdivision in Table 24: the distribution is remarkably similar.

This similitude between top 15 and the entire corpus is equally visible in FICLE and LOCNESS. In the former, 76.37% of all the 15 most-found connectors are placed in initial position (n = 711). Another 22.66% are placed in CMP (n = 211) and a further 0.97% are written in the clause-final position (n = 9). The native speaker corpus boasts analogous percentages, with 74.02% of the most frequent connectors falling in initial position (n = 678). In CMP, that share equals 25.55% (n = 234), while in CFP, that number is a minute 0.44% (n = 4).

3.3.1. Clause-initial positioning

The clause-initial positioning mentioned above can further be divided into two separate categories: connectors which are exclusively sentence-initial and clause-initial connectors. In the former, we find connectors that start a sentence; in the latter, we distinguish connectors that are at the head of the clause but *not* at the beginning of a sentence. This differentiation helps to see whether learners of English prefer to use connectors at the start of a clause or at the start of a sentence.

	CoFLE			LOCNESS			FICLE		
	Rf.	% IP ¹¹	% all ¹²	Rf.	% IP	% all	Rf.	% IP	% all
SIP only	424	87.24%	65.13%	895	92.08%	68.48%	854	92.73%	69.89%
CIP only	62	12.76%	9.52%	77	7.92%	5.89%	67	7.27%	5.48%
TOTAL	486	100%	74.65%	972	100%	74.37%	921	100%	75.37%

Table 24: CIP and SIP of connectors in the three corpora

¹⁰ These calculations are based on Table 22, i.e. **with** the connectors which figure in IP only. Leaving them out would distort the results too greatly.

Due to their size, they have been moved to the bottom of this chapter.

¹¹ This percentage denotes the ratio between SIP or CIP and the all connectors used *in initial position*.

¹² This percentage shows the share SIP or CIP possesses in all connectors used *in that corpus*.

The results from the analysis show that learners of English, both native speakers and foreign language learners mostly place their connectors in sentence-initial positioning. As can be seen in Table 24, the SIP is where 65.13% of all connectors in CoFLE are to be found. Another 9.52% are found in CIP only position, or 62 occurrences. Of all the connectors in initial position, SIP has an 87.24% share, while the CIP only connectors make up 12.76%.

In FICLE we observe similar statistics. 69.89% of all connectors found in that learner corpus are written at the head of a sentence, while 5.48% of all connectors find themselves at the beginning of a clause that is not also the beginning of a sentence. The ratio of connectors within the initial position stands at 92.73% for SIP to 7.27% for CIP.

Similar results are also found in LOCNESS. English native learners place 68.48% of all connectors in sentence-initial position, compared to 5.89% in clause-initial position. Of all connectors in initial position, then, 92.08% figure in SIP, while 7.92% of them are in CIP only.

These results mean that Flemish learners of English and native learners alike write more sentences that begin with a connector than they write clauses within sentences that begin with a connector. Learners of English are likelier to separate different ideas into different sentences, rather than joining them together into two independent clauses separated by punctuation marks (e.g. the semicolon).

Certain connectors are used more frequently than others in initial position. The most striking examples are, as was seen previously, *and*, *but*, *so* and *yet*.

It is at times frowned upon to start a sentence with *and* or *but*, but, while stylistically discouraged by some teachers, it does not constitute a grammatical mistake. Therefore, these two words cannot, unlike most others, function as a connector outside SIP. Across all three corpora, they are placed in initial position 100% of the time (see Tables 25, 26 and 27). *And* appears 41 times in CoFLE (adj.f = 9.34); in FICLE, it shows up 124 times (adj.f = 14.24), in LOCNESS, it occurs 48 times (adj.f = 2.83). *But* is used 70 times in CoFLE (adj.f = 15.95); in FICLE, 290 times (adj.f = 33.30); in LOCNESS, 102 times (adj.f = 6.01). From these adjusted frequencies, we can observe that native speakers of English are more often made clear that these connectors should not be used at the head of a sentence. Flemish learners of English still make this 'mistake' and are prone to starting their sentences with *and* or *but*.

In CoFLE certain other connectors from the top 15 frequently occur in initial position. These include *however* (69.84% of its occurrences are in IP), *on the one / other hand* (80%), *of course* (74.07%), the resultive connector *therefore* (65.52%), *(al)though* (90%) and *so* (100%).

In the case of LOCNESS, *however* also prominently features in IP, with 69.42% of that connector being found at the head of a clause or sentence. *Also* appears just over half the time in IP (54.46%), *(al)though* does so 66.97% of the time. Others include *for example* (58.82%), *therefore* (71.43% for the resultives, 76.47% for the summatives) and others.

FICLE counts 20 of its 22 instances of *first of all* in initial position, equalling 90.91%. Other connectors found prevalently in IP include *(al)though* (61.76%), the resultive *therefore* (71.88%), *on the one / other hand* (77.42%), *first(ly)* (81.82%) and the listing and additive *then* (75%).

The above-mentioned connectors are commonly placed in initial position but are not the only ones placed there. Their presence among the fifteen most popular connectors in either of the three corpora, however, indicates that there is a clear trend. Learners of English, be they native speakers or foreign language learners, will predominantly place connectors in initial position. Sometimes, their placement in IP is necessary for them to be analysed as a connector. *And*, *but* and *so*, for instance. As Buysse (2012) writes, resultive *so* functions as a “coordinating conjunction or an adverb and hence always introduces a main clause” (p. 1776). Mostly, though, learners place them in IP to explicitly mark a relation of result or inference with the prior adjacent clause. It is up to the learner to let go of that need for safety and be willing to diversify their writing. Some connectors nevertheless do feature in other positions than the initial one.

3.2.2. Clause-medial positioning

Clause-medial positioning appears substantially less than initial position in either of the three corpora — about one in four is placed in CMP (see Table 22 above). In CoFLE 157 connectors are placed in CMP, which equals 24.12% of all connectors. For LOCNESS, the share of CMP is 25.02% (n = 327). FICLE, then, has 282 instances (or 23.08%) in medial position.

Most of the connectors found in CMP are from the fifteen most-frequent connectors in either of the three corpora. For the Corpus of Flemish Learners of English, 99 out of 157 connectors in clause-medial position are in the top 15. Connectors that are frequently found in that position include *also* (where 78.26% of its tokens are in CMP), *however* (30.16% of all its connector tokens), *for example* (50%) and the resultive connector *therefore* (34.48%).

The top 15 in LOCNESS contains 234 out of the 347 connectors found in CMP in the native-speaker corpus. Some connectors that are commonly found in that position include *however* (29.13%), *also* (45.54%), *for example* (41.18%), *(al)though* (32.11%), *in fact* (37.5%) and *on the one / other hand* (38.46%). Interestingly, *actually* is a connector which did not make the top 15 in LOCNESS, occurring 16 times. 13 of those instances are found in CMP, which equals 81.25% of all the tokens of that particular connector.

For the fifteen most-used connectors in FICLE, too, the share of CMP connectors is great, with 211 out of 282 belonging to the clause-medial position in that list. Similarly to the previous two corpora, recurring connectors in CMP are *however* (57.5%), *also* (81%), *for example* (55.88%), *indeed* (55.67%), *for instance* (65%) and *(al)though* (38.24%). In this learner corpus, the summative *then*, which does not appear in the top 15 of FICLE, was found to be used 10 out of 11 times in CMP, or 90.91% of all its occurrences.

Some of the connectors found in clause-medial position have also been found in clause- or sentence-initial position. This is because some connectors function equally well in either position, but, because of the English learners' propensity to place connectors in initial position, they will be found more often in the latter place. Certain connectors, such as *however*, *also*, *(al)though*, *for example*, *of course* and *therefore* have a greater tendency to be used in CMP than others. Still, while this handful of connectors

are frequently used in clause-medial position, many more find themselves placed in IP (see Tables 22 and 23). The reality does not change: connectors prevail at the head of a clause or sentence.

It is clear from the numbers above (which can be found in detail in Tables 25, 26 and 27 below) that Flemish learners of English and native learners alike use the clause-medial position to the same degree, about one in four times. A similar distribution of connectors was found in the top 15 most-used connectors across all three corpora.

3.3.3. Clause-final positioning

Of all the syntactic positions in the learner and native speaker corpora, the clause-final position (CFP) appears least often. In CoFLE only 8 instances of connector placed in CFP are to be found, which equals 1.23%. In the case of LOCNESS, too, we find 8 tokens in CFP, or 0.61%. FICLE contains 19 connectors in CFP, or 1.55% of that corpus.

Because of their rare occurrences in learner writing, it does not make sense to analyse the list of the fifteen most frequently occurring connectors and see how many times the CFP appears. Instead, for each corpus, I will give the list of all connectors that emerge from the learners' essays in clause-final position.

In the Corpus of Flemish Learners of English, *for example*, and *in general* each contain two tokens placed in clause-final position. *In any case*, *in the first place*, *actually* and the summative connector *then* add one token each to the small share CFP has.

The percentage of connectors figuring in CFP in LOCNESS is lower still than in CoFLE, at 0.61% only. *However* makes up three of the eight tokens found in CFP; *indeed* makes up for two, while connectors *in particular*, *(al)though*, and *in effect* all have one token in clause-final position.

In FICLE the share of CMP is slightly higher than in the other two corpora, with 1.55% (n = 19) of all connectors being in that position, of which *of course* occurs the most (n = 5). Connectors that appear twice in CFP are *in the first place*, *that is (to say)*, *on the contrary* and the contrastive *then*. There are six connectors which appear only once in clause-final position: *for example*, the summative *then*, *however*, *(al)though*, *accordingly*, *indeed*.

Contrary to the initial and clause-medial positions, it is difficult to make a statement as to what connectors will appear more often than others in clause-final position. Indeed, since connectors placed in clause-final position occur less than 2% in either of the three corpora, it seems to me that which connectors are used in CFP are random. It is the personal choice of writers, I believe, not a general trend, that got these connectors in the lists above. Despite some connectors being said to appear more in clause-final position, there are not enough occurrences of them in this present study to support this evidence.

CoFLE	Rf	IP	% ¹³	CMP	%	CFP	%
But	70	70	100.00%	0	0.00%	0	0.00%
However	63	44	69.84%	19	30.16%	0	0.00%
Also	46	10	21.74%	36	78.26%	0	0.00%
And	41	41	100.00%	0	0.00%	0	0.00%
For example	38	17	44.74%	19	50.00%	2	5.26%
Therefore (resultive)	29	19	65.52%	10	34.48%	0	0.00%
Of course	27	20	74.07%	7	25.93%	0	0.00%
So	26	26	100.00%	0	0.00%	0	0.00%
On the one / other hand	25	20	80.00%	5	20.00%	0	0.00%
(Al)though	20	18	90.00%	2	10.00%	0	0.00%
In conclusion	15	15	100.00%	0	0.00%	0	0.00%
Yet	13	13	100.00%	0	0.00%	0	0.00%
Second(ly)	12	12	100.00%	0	0.00%	0	0.00%
Moreover	12	12	100.00%	0	0.00%	0	0.00%
First(ly)	11	10	90.91%	1	9.09%	0	0.00%
TOTAL	448	347	77.45%	99	22.10%	2	0.45%

Table 25: Syntactic positioning of the 15 most used connectors in CoFLE

LOCNESS	Rf	IP	% ¹⁴	CMP	%	CFP	%
However	206	143	69.42%	60	29.13%	3	1.46%
Also	112	61	54.46%	51	45.54%	0	0.00%
(Al)though	109	73	66.97%	35	32.11%	1	0.92%
But	102	102	100.00%	0	0.00%	0	0.00%
For example	68	40	58.82%	28	41.18%	0	0.00%
And	48	48	100.00%	0	0.00%	0	0.00%
Therefore (resultive)	42	30	71.43%	12	28.57%	0	0.00%
So	41	41	100.00%	0	0.00%	0	0.00%
In fact	40	25	62.50%	15	37.50%	0	0.00%
Yet	35	35	100.00%	0	0.00%	0	0.00%
Of course	30	21	70.00%	9	30.00%	0	0.00%
On the one / other hand	26	16	61.54%	10	38.46%	0	0.00%

¹³ This percentage shows how much of that connector is in a given syntactic position.

¹⁴ This percentage shows how much of that connector is in a given syntactic position.

Then (<i>listing and additive</i>)	20	14	70.00%	6	30.00%	0	0.00%
Now	20	16	80.00%	4	20.00%	0	0.00%
Therefore (<i>summative</i>)	17	13	76.47%	4	23.53%	0	0.00%
TOTAL	916	678	74.02%	234	25.55%	4	0.44%

Table 26: Syntactic positioning of the 15 most used connectors in LOCNESS

FICLE	Rf	IP	%¹⁵	CMP	%	CFP	%
But	290	290	100.00%	0	0.00%	0	0.00%
And	124	124	100.00%	0	0.00%	0	0.00%
Also	100	19	19.00%	81	81.00%	0	0.00%
So	85	85	100.00%	0	0.00%	0	0.00%
Of course	60	33	55.00%	22	36.67%	5	8.33%
However	40	16	40.00%	23	57.50%	1	2.50%
(Al)though	34	20	61.76%	13	38.24%	1	2.94%
For example	34	14	41.18%	19	55.88%	1	2.94%
Therefore (<i>resultive</i>)	32	23	71.88%	9	28.13%	0	0.00%
On the one / other hand	31	24	77.42%	7	22.58%	0	0.00%
First of all	22	20	90.91%	2	9.09%	0	0.00%
First(ly)	22	18	81.82%	4	18.18%	0	0.00%
Indeed	21	6	28.57%	14	66.67%	1	
For instance	20	7	35.00%	13	65.00%	0	0.00%
Then (<i>listing and additive</i>)	16	12	75.00%	4	25.00%	0	0.00%
TOTAL	931	711	76.37%	211	22.66%	9	0.97%

Table 27: Syntactic positioning of the 15 most used connectors in FICLE

¹⁵ This percentage shows how much of that connector is in a given syntactic position.

Conclusion

In this study, I looked at the frequencies of connectors in Flemish learner corpora and calculated the densities of these connectors per semantic classification as well as those of the fifteen most popular connectors. Next, I determined whether Flemish learners of English overuse and underuse connectors in argumentative essay writing. I further observed where in a clause Flemish learners of English were more likely to place connectors. Not only did I examine my own learner corpus, I also included FICLE in my analyses. Doing so would help generalise my findings if the results were similar or discover why results differed between the corpora. In this chapter, I will recapitulate the major findings of this study.

First, this research found that Flemish learners of English use more connectors than English native speakers. In a 43,899-word-strong corpus called the Corpus of Flemish Learners of English (CoFLE), 651 connectors were identified. In other words, connectors in CoFLE have an adjusted frequency of 148.29 per 10,000 words. A similar adjusted frequency was also observed in the second non-native corpus (FICLE), which comprises 140.33 connectors per 10,000 words. Compared to LOCNESS' adjusted frequency of 77.05, we see close to a doubling of connectors in the argumentative essays of Flemish learners of English.

Next, I looked at the frequency of connectors per semantic category. Here, too, Flemish learners of English use more connectors per 10,000 words than do the native speakers. All three corpora, both those of non-native as of native learners of English, share the same three most-used categories: the contrastive, listing and additive, and resultive ones. The other three categories (appositive, corroborative and summative) feature in all three corpora, though their share is smaller compared to the first three aforementioned. These results match the distribution of semantic categories in the top 15 most used connectors.

Overall, Flemish EFL learners overuse connectors in their academic essay writing: for every connector used in LOCNESS, CoFLE contains 1.92 connectors (or 1.80 for FICLE). What's more, Flemish learners of English in CoFLE overuse connectors in every semantic category. In FICLE they overuse all categories save the appositive and summative ones. No semantic category is found to be underused in either learner corpus.

Learners participating in CoFLE overuse 38 and underuse 8 individual connectors. In the most extreme cases of overuse or underuse, where the NNS/NS ratio was particularly high (*lastly*, *to conclude*, and *as a consequence*) or low (e.g. *anyway* and *in fact*), the frequencies of those connectors was relatively low to begin with. High-frequency connectors such as *also* or *for example* get overused, too, though the ratios are not nearly as high (or low, in case of underuse) as words that have a low frequency in either the non-native or the native-speaker corpus. Smaller frequencies are more likely to yield larger fluctuations when divided by one another. Therefore, the 'worst' cases of overuse and underuse are usually not as telling as high-frequency connectors which are slightly less over- or underused. In those

cases, we may say that their overuse is a feature shared by Flemish learners of English. The same conclusions can be applied to the overuse pattern found in FICLE. Participants in this learner corpus have overused 30 connectors and underused 20 connectors.

Flemish learners of English mainly place their connectors in initial position. Previous studies (see Section 1.3.3.3) have claimed this phenomenon is a universal one. The results from this study confirm this statement. What's more, this research has also found that L1 learners of English also place most of their connectors in IP.

These results have implications for English teaching in Flanders. At present, teachers greatly focus on cohesion through connectors. While this method leads to successful cohesive texts, it also leads to learners overusing connectors and placing them mainly in sentence-initial position. The cohesive ties achieved through the use of connectors are explicit. Teachers should start to educate their students on the plenitude of implicit cohesion methods.

Discussion

5.1. Discussion of the major findings

Throughout Chapter 3, I compared CoFLE and FICLE to LOCNESS as well as to each other. Unsurprisingly, the two corpora are similar in many regards. Both corpora have a nearly identical connector density (148.29 per 10,000 words for CoFLE, 140.33 for FICLE).

The overwhelming preference to use contrastive, listing and additive connectors and, to a marginally lesser extent, resultive connectors is no great surprise. Argumentative essays are by their very nature going to contain words that indicate contrast, disagreement, addition and/or results. It is the task of students to propose arguments and counterarguments to their statement and, in doing so, reach a satisfactory conclusion. To achieve this feat, they must use connectors from the three aforesaid categories. It is possible to conjecture that this distribution of percentages across the semantic categories may show a lack of vocabulary variety on the part of Flemish learners of English. Indeed, the percentages seem to be more spread out in LOCNESS than in either CoFLE or FICLE.

The fifteen most common connectors are responsible for the majority of individual connector types found in all three corpora. In total, CoFLE and FICLE contain 67 and 68 different connector types respectively. For the native-speaker corpus, that amount equals 77. This means that the native learners of English show a greater language variety when applying connectors in their texts. Yet, the three corpora have many connectors in common; after all, an argumentative essay will share similarities, whatever the language of origin of the learner. Their arguments need consolidation and support and are achieved through connectors and conjunctions that indicate addition, contrast, results, corroboration, etc. The more frequent connector is used in the English language (see Browne, Culligan & Phillips' 2013 *New General Service List*), the likelier it is to be found in the present essays. Indeed, connectors such as *but*, *and*, *also*, *so* and *however* will be present in either corpus, simply because they are ubiquitous in English in the first place.

In both the Flemish learner corpora and in the LOCNESS, some of the high-frequency connectors are in fact conjunctions which were placed at the start of a sentence, effectively making them connectors. Why are some of these conjunctions so prevalent in connectors position? *And* is the third most-common word in the English language (Browne, Culligan & Phillips, 2013). While it acts mainly as a conjunction, it can occasionally also appear as a connector. The same applies to *but*, which appears 25th in Browne et al.'s (2013) *New General Service List*¹⁶. When using *and* or *but* in clause-initial position to connect two separate sentences, they become connectors. According to Eastwood (2005), placing *and*, *or*, *but* or *so* at the start of a sentence has become a standard feature in English text. Although English writing style used to discourage this usage, "this attitude is now less common" (Eastwood, 2005, p. 322nb). Even though *and*, *but* or *so*, for instance, function as a conjunction, Flemish learners of English and

¹⁶ The NGSL was created 60 years after West's original 1953 *General Service List* and aims just like its predecessor to map core vocabulary in the English language.

native speakers alike seem to use it as a connector in their essays, despite both groups being taught to avoid this practice. As long as they produce cohesive texts, pupils are usually not told to edit the affected sentences.

There are slight discrepancies to be observed in the way Flemish learners of English in CoFLE and FICLE overuse connectors in their academic essay writing. Although non-native learners in both corpora overuse connectors to a similar degree, the categories in which they do so and the intensity with which they overuse them vary. Differences in connector (over)use within the same EFL learner group have been observed in other studies. For instance, Heino (2010) and Altenberg and Tapper (1998) found a general underuse of connectors amongst Swedish learners of English, whereas Tapper (2005), who also examined Swedish EFL learners, reports a general overuse of connectors. That said, the differences observed between CoFLE and FICLE are not as drastic as with the Swedish researchers. As was said before, the overall trend of overuse here is similar in either learner corpus; only when looking at the semantic categories do we find some different overuse patterns that may be indicative of diverging personal preferences by Flemish EFL learners.

However is an interesting case to analyse separately. In several studies looking at connector frequency this connector has been found to be either overused (e.g. Appel & Szeib, 2018; Ha, 2016; Shea, 2009) or underused (e.g. Granger & Tyson, 1996). What makes it interesting in this case are the contradictory results *within* the same L1 group, i.e. Flemish EFL learners. In CoFLE *however* is neither overused or underused, having an NNS/NS ratio of 1.18. In FICLE it gets heavily underused, with an NNS/NS ratio of merely 0.38. How is this difference to be explained? *But* appears 290 times in FICLE, which represents 59.92% of all the contrastive connectors (n = 484). It is possible, therefore, that participants in this learner corpus decided to mark contradiction by means of connectors other than *however*, something which participants in CoFLE decided against. Indeed, *but*, appearing 70 times in the latter corpus, represents 29.79% of all contrastive connectors. Its smaller share in CoFLE enables other connectors, such as *however*, to be used more often — a possible reason why CoFLE does not overuse *however*.

In the learner corpora, several connectors were observed to have high underuse or overuse ratios. In FICLE, *in particular* has an underuse ratio of 0.32; in CoFLE, *likewise* has one of 0.64. The connector *to conclude* has an overuse ratio of 7.79 in FICLE; in CoFLE, *as a consequence* has one of 15.46. These four connectors and many more besides have a frequency smaller than 1 in 10,000 words. Connectors with higher frequencies do not have such high overuse or underuse ratios. The most extreme cases of underuse barely figure in either corpora in the first place. Therefore, it is safe to say that the most outspoken cases of underuse in this corpus are not the most relevant for our research. Instances of overuse or underuse where the adjusted frequencies of the connectors are higher (e.g. 4.56 in CoFLE versus 6.43 in LOCNESS for *(al)though*), we may conclude that overuse and underuse are linked to the linguistic habits of the learners and is not a mere happenstance. When interpreting

overuse and underuse data, it is important to keep in mind that the lesser-used a connector, the higher the chances are of observing large overuse ratios.

Finally, this study found that learners of English, both Flemish and native learners, mostly place their connectors in initial position. Researchers across different countries have previously found their learners of English to prevalently use connectors in IP (e.g. Altenberg & Tapper, 1998; Bikelienè, 2008b; Dastjerdi & Samian, 2011; Granger & Tyson, 1996; Ha, 2014; Hajiyan, 2015; Lee, 2006; Narita, Sato & Sugiura, 2004; Xu & Liu, 2012). Several of these researchers, like Granger and Tyson (1996), have claimed that the trend in learner writing of using connectors in initial position is a universal feature. The results of this study corroborate this statement: with practically three quarters of all connectors featuring in IP in CoFLE and FICLE, Flemish learners of English do indeed prefer to start clauses or sentences with a connector. However, some connectors cannot be placed elsewhere but IP (e.g. *and*, *but*, *yet*, *so*, etc.) and consequently distort the results. Indeed, it would not be fair to say that learners favour placing connectors in IP if, for some of those connectors, they have no choice but to place them there. I therefore filtered out these connectors and found that the connectors were more evenly distributed. Connectors in IP now make up 64.80% of all connectors in CoFLE and 55.43% in FICLE. We can conclude that while there is still a tendency to place connectors in IP, the proportion is less pronounced than when keeping IP-only connectors in our calculations.

When looking into the patterns of syntactic positioning in EFL learners' writing, I found no study ascertaining that native-speaker learners, too, prefer placing connectors in IP. I was consequently surprised to find that LOCNESS contains a practically similar percentage of connectors placed in IP: 74.37% (or, with IP-only connectors filtered out, 67.56%). To me, it indicates that the differences we observe between native speakers and learners of English are not that great after all. In Paquot's 2010 book *Academic Vocabulary in Learner Writing*, the author claims that the initial position acts a "safe bet" (p. 179) to mark the link between one idea and what has come before. Authors such as Xu and Liu (2012) or Narita et al. (2004) write that foreign language learners learn to organise texts by using connectors to link ideas and, by being taught to use them at the start of a sentence, do not become aware of the flexibility of connector-positioning. By putting forward such claims, the researchers imply that native learners of English do not face such problems. This my results repudiate. If indeed learners prefer the initial position, then so do learners of English as a First Language: the distribution of connectors per syntactic position in this study is the same with Flemish EFL learners as their native speaker counterparts. Positioning connectors at the start of the sentence is an inherent feature not only in essays of EFL learners, but also in those of native learners of English.

5.2. Pedagogical implications

Flemish learners of English usually correctly employ connectors when writing essays. They are able to establish cohesive relations and create a coherent whole which the reader will understand. Several improvements, however, are called for. It was noticed in this study that Flemish EFL learners, when compared to English L1 learners, tend to overuse connectors in their essay writing. What's more, both

L1 and EFL learners also mostly place them at the start of a sentence. This behaviour overuse stems from the methodology used by teachers. Cohesion, much like vocabulary acquisition and grammar assimilation, is taught explicitly. Although explicit cohesive ties are a useful way to map out the desired text structure, the creation of textual cohesion goes beyond the explicit. This study presents pedagogical implications to curb those problems.

First of all, students should be taught the different ways to connect ideas (i.e. addition, contrast, corroboration, etc.). Learners should identify connectors which are overused and find out how it negatively impacts the legibility (Appel & Szeib, 2018). Better still, learners could be encouraged to forego connectors altogether in a first draft and instead rely on other methods to establish cohesion, such as ellipsis, reference or lexical cohesion. Later, they would decide whether cohesion in certain passages could benefit from the addition of a connector or not. Overemphasizing the importance of connectors leads students to overuse them. They should be made aware that using more connectors does not equal a more cohesive text: they help link ideas in a text but are not the only tool at a writer's disposal.

Second, learners of English, both non-native — in this case Flemish — and native, should be taught how to position their connectors in their sentences. All too often, they will start off with one, so as to explicitly mark the link between what has gone before and what has come after. Much like in the previous point, teachers should point out the alternatives to this method. Doing so will increase the learners' awareness of their writing style. By avoiding repetition and varying the syntactic positioning, they will improve the flow of their text, thus creating a more coherent and cohesive whole.

Third, context is crucial when learning about cohesion. In *Spell It Out*¹⁷, David Crystal (2012) argues that words should never be taught out of context; instead, they should be learned in association with other words. Learners should be able to identify the role a word plays in the bigger picture. For every connector, its semantic and grammatical contexts should be considered.

Lastly, learners should be presented authentic linguistic content. I strongly believe in the power of linguistic immersion. The best way to learn a language is to be surrounded by it. All too often, curricula do not stimulate pupils to delve deeper than the books they are given. Constant exposure to a language is the way forward. In my own teaching, I have always encouraged pupils to communicate in English, to myself and to others. Allow them to make mistakes but be ready to correct them and give them structure for the future. Projects and portfolios are a great way to submerge pupils in the language and be introduced to authentic content such as English literature, television, radio and podcasts. Doing so, they will see when and how connectors are used — and when and how they are not. Then, teachers can show them the underlying rules and stylistic guidelines.

5.3. Limitations

Although this research aims to be thorough and extensive, some limitations should be remarked, which could guide future research in a certain direction.

¹⁷ Although pertaining to spelling, the methodological aspects brought up in Crystal's book are equally pertinent when it comes to cohesion.

First, the corpus analysed is limited in size. Further research could increase its size by finding more participants in Flemish universities. What's more, I agree with Hessamy and Hamed (2013), who argue that independent tasks as were produced for this learner corpus are merely a snapshot of the students' capabilities and it does not properly capture their academic writing competence. When writing for an independent task (as was the case for all three corpora in this study), students lack topic familiarity which will affect their performance. A lack of background knowledge effectively acts as a hindrance for fluent writing. Later research could decide to opt for integrated tasks instead of independent ones.

Second, this research is merely quantitative in its execution and does not delve into the individual essays the students have written. To better understand how Flemish learners of English use and misuse connectors in their essays, a qualitative follow-up research is needed. While an interesting facet of textual cohesion, analysing connector misuse did not fit the scope of my study. Moreover, when reading individual essays written by the participants, I did not notice many instances where connectors usage would qualify as incorrect. Instead, I would detect inappropriate use (e.g. placing connectors from the wrong semantic field), which, again, did not fit the scope of my study. A qualitative approach to CoFLE and FICLE could yield interesting answers to relevant questions. How does Flemish influence English EFL learners' writing? Can we speak of L1 transfer? What connectors get frequently misused? How could EFL learners avoid inappropriate use of connectors?

Third, I decided to base my analysis on an existing list of 87 connectors, stemming mainly from classifications by Greenbaum (1996) and Quirk et al. (1985). More connectors exist in the English language than could have been examined, however. Future research could opt for a different methodology by carefully picking out the connectors present in the learner corpus and creating a list from those hits. For instance, it could use Paquot's (2010) *Academic Keyword List* as a starting point and analyse strictly academic texts (e.g. theses or dissertations) to see which connectors from the AKL feature in those texts.

Finally, this corpus study compared argumentative essays written by non-native and native learners of English. As pointed out earlier in Section 4.1, however, the results of this research could be seen under a different light if compared to writings of expert L1 writers. In doing so, one could distinguish the linguistic pitfalls still present in learner writing, native and non-native alike. Almost all research in corpus linguistics has focused on the differences between EFL and L1 learners only. This methodology — consciously or not — discards the notion that corpora containing L1 learner texts might (and likely do) give a false idea of what 'correct' English looks like. A similar written corpus by expert L1 writers would offer a better way to distinguish the correct from the incorrect. Putting these three groups side by side should become standard practice in the future. This realisation comes too late for this study, but I hope that future research will bring more nuance to their results by considering L1 learners to be just that: learners. Learners, native or not, are not free of mistakes.

Despite this research containing some limitations, I do hope this explorative study has added to the existing research on the use of connectors in argumentative essay writing by EFL learners and that it will serve as a reference for future studies.

References

- Academic Keyword List. (n.d.). Retrieved December 21, 2018, from <https://uclouvain.be/en/research-institutes/ilc/cecl/academic-keyword-list.html>
- Adiantika, H. N. (2015). Cohesive Devices in EFL Students' Expository Writing. *English Review: Journal of English Education*, 4(1), 1–9. <https://doi.org/10.25134/erjee.v4i1.316>
- Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80–93). London: Longman.
- Anthony, L. (2018). AntConc (Version 3.5.7) [Computer Software]. Tokyo: Waseda University. Retrieved from <http://www.laurenceanthony.net/software>
- Appel, R., & Szeib, A. (2018). Linking adverbials in L2 English academic writing: L1-related differences. *System*, 78, 115–129. <https://doi.org/10.1016/j.system.2018.08.008>
- Bamberg, B. (1983). What Makes a Text Coherent? *College Composition and Communication*, 34(4), 417–429. <https://doi.org/10.2307/357898>
- Bamberg, B. (1984). Assessing Coherence: A Reanalysis of Essays Written for the National Assessment of Educational Progress, 1969-1979. *Research in the Teaching of English*, 18(3), 305–319. Retrieved from <https://www.jstor.org/stable/40171021>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Bikelienė, L. (2008a). Connector usage in advanced Lithuanian Learners' English Writing. In B. Lewandowska-Tomaszczyk (Ed.), *Corpus Linguistics, Computer Tools, and Applications - State of the Art* (pp. 741–755). Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang. Retrieved from <https://goo.gl/xYFJB9>
- Bikelienė, L. (2008b). Resultive connectors in Advanced Lithuanian Learners' English Writing. *Kalbotyra*, 59(3), 30–37. Retrieved from <http://www.journals.vu.lt/kalbotyra>
- Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics*, 7(2), 165–182. <https://doi.org/10.1075/ijcl.7.2.02bol>
- Browne, C., Culligan, B., & Phillips, J. (2013). New General Service List. Retrieved from <http://www.newgeneralservicelist.org>
- Buysse, L. (2011). Magali Paquot: Academic Vocabulary in Learner Writing. From Extraction to Analysis. *Applied Linguistics*, 32(3), 356–359. <https://doi.org/10.1093/applin/amr012>
- Buysse, L. (2012). So as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, 44(13), 1764–1782. <https://doi.org/10.1016/j.pragma.2012.08.012>
- Carrell, P. (1982). Cohesion Is Not Coherence. *TESOL Quarterly*, 16(4), 479–488. <https://doi.org/10.2307/3586466>
- Chen, C. W. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1), 113–130. <https://doi.org/10.1075/ijcl.11.1.05che>

- Cherchi, L. (1978). L'ellipse comme facteur de cohérence. *Langue Française*, 38, 118–128.
<https://doi.org/10.3406/lfr.1978.6123>
- Corpus collection guidelines. (n.d.). Retrieved December 17, 2018, from
<https://uclouvain.be/en/research-institutes/ilc/cecl/corpus-collection-guidelines.html>
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213.
<https://doi.org/10.2307/3587951>
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1–16. <https://doi.org/10.1016/j.jslw.2016.01.003>
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35(2), 115–135. <https://doi.org/10.1111/j.1467-9817.2010.01449.x>
- Crystal, D. (2012). *Spell It Out: The Singular Story of English Spelling*. Profile Books.
- Dastjerdi, H. V., & Samian, S. H. (2011). Quality of Iranian EFL learners' argumentative essays: Cohesive devices in focus. *Mediterranean Journal of Social Sciences*, 2(2), 65–76. Retrieved from <https://goo.gl/nDsvrq>
- Eastwood, J. (2005). *Oxford Learner's Grammar: Grammar Finder*. Oxford: Oxford University Press.
- Egins, S. (1994). *An Introduction to Systemic Functional Linguistics*. London: Pinter.
- Fei, D. (2006). The effect of the use of adverbial connectors on Chinese EFL learners' English writing quality. *CELEA Journal (Bimonthly)*, 29(1), 105–111.
- Field, Y., & Yip, L. M. O. (1992). A Comparison of Internal Conjunctive Cohesion in the English Essay writing of Cantonese Speakers and Native Speakers of English. *RELC Journal*, 23(1), 15–28. <https://doi.org/10.1177%2F003368829202300102>
- Ghasemi, M. (2013). An Investigation into the Use of Cohesive Devices in Second Language Writings. *Theory and Practice in Language Studies*, 3(9), 1615–1623. <https://doi.org/10.4304/tpls.3.9.1615-1623>
- Granger, S., Dagneaux, E., Meunier, F., & Paquot, M. (2009). *International Corpus of Learner English v2 (Handbook + CD-Rom)*. Louvain-la-Neuve: Presses universitaires de Louvain. Retrieved from <https://uclouvain.be/en/research-institutes/ilc/cecl/iclev2.html>
- Granger, S., & Tyson, S. (1996). Connector Usage in the English Essay Writing of Native and Non-Native EFL Speakers of English. *World Englishes*, 15(1), 17–27. <https://doi.org/10.1111/j.1467-971X.1996.tb00089.x>
- Greenbaum, S. (1996). *The Oxford English Grammar*. Oxford: Oxford University Press.
- Ha, M.-J. (2014). A Corpus-Based Study on Korean EFL Learners' Use of English Logical Connectors. *International Journal of Contents*, 10(4), 48–52. <https://doi.org/10.5392/IJoC.2014.10.4.048>
- Ha, M.-J. (2016). Linking adverbials in first-year Korean university EFL learners' writing: a corpus-informed analysis. *Computer Assisted Language Learning*, 29(6), 1090–1101. <https://doi.org/10.1080/09588221.2015.1068814>

- Hajiyani, M. (2015). *Connectors in Contrast : A Comparative Study of Iranian-Learner and Native-Speaker Usage*. Albert-Ludwigs-Universität. Retrieved from <https://freidok.uni-freiburg.de/data/10015>
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar* (2nd ed.). London: Arnold.
- Halliday, M. A. K., & Hasan, R. (1979). *Cohesion in English* (3rd impr.). London: Longman.
- Heino, P. (2010). *Adverbial Connectors in Advanced EFL Learners ' and Native Speakers ' Student Writing*. University of Stockholm. Retrieved from <http://www.diva-portal.org/smash/get/diva2:325443/FULLTEXT01.pdf>
- Hessamy, G., & Hamed, S. (2013). A Comparison of the Use of Cohesive Devices in EFL Learners' Performance on Independent vs. Integrated Writing Tasks. *Study in English Language Teaching*, 1(1), 121–146. <https://doi.org/10.22158/selt.v1n1p121>
- Hyland, K. (1990). A Genre Description of the Argumentative Essay. *RELC Journal*, 21(1), 66–78. <https://doi.org/10.1177/003368829002100105>
- Jafarpur, A. (1991). Cohesiveness as a Basis for Evaluating Compositions. *System*, 19(4), 459–465. [https://doi.org/10.1016/0346-251X\(91\)90026-L](https://doi.org/10.1016/0346-251X(91)90026-L)
- Jalilifar, A. (2008). Discourse Markers in Composition Writings : The Case of Iranian Learners of English as a Foreign Language. *English Language Teaching*, 1(2), 114–122.
- Johns, A. M. (1980). Cohesion in Written Business Discourse: Some contrasts. *The ESP Journal*, 1(1), 35–43. [https://doi.org/10.1016/0272-2380\(80\)90008-6](https://doi.org/10.1016/0272-2380(80)90008-6)
- Johnson, P. (1992). Cohesion and Coherence in Compositions in Malay and English. *RELC Journal*, 23(2), 1–17. <https://doi.org/10.1177%2F003368829202300201>
- Khalil, A. (1989). A study of cohesion and coherence in ESL students' writing. *System*, 17(3), 359–371. <https://doi.org/10.1080/08351818409389208>
- Lee, E.-J. (2004). A Corpus-Based Analysis of the Korean EFL Learners' Use of Conjunctive Adverbials. *English Teaching*, 59(4), 283–301.
- Leńko-Szymańska, A. (2008). Non-native or non-expert? The use of connectors in native and foreign language learners' texts. *Acquisition et Interaction En Langue Étrangère*, 27, 91–108. Retrieved from <http://aile.revues.org/4213>
- Leńko-Szymańska, A. (2015). Abundance, However, is not Always Desirable: Connectors in Polish EFL Learners' Texts. In M. B. Paradowki (Ed.), *Productive Foreign Language Skills for an Intecultural World. A Guide (not only) for Teachers* (pp. 237–253). Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang. <https://doi.org/10.1002/ajmg.a.37417>
- Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33(4), 623–636. <https://doi.org/10.1016/j.system.2005.02.002>
- Martin, J. R. (1992). *English Text: System and Structure*. Amsterdam: Benjamins. <https://doi.org/10.1075/z.59>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic Features of Writing quality. *Written Communication*, 27(1), 57–86. <https://doi.org/10.1177/0741088309351547>

- Milton, J., & Tsang, E. S. (1993). A Corpus-based study of logical connectors in EFL students writing: directions for future research. In R. Pemberton & E. S. Tsang (Eds.), *Studies in Lexis* (pp. 215–246). Hong Kong: HKUST Language Center. Retrieved from <http://hdl.handle.net/1783.1/1083>
- Narita, M., Sato, C., & Sugiura, M. (2004). Connector Usage in the English Essay Writing of Japanese EFL Learners. In M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa, & R. Silva (Eds.), *Proceedings of the Fourth International Conference on Language Resources and Evaluation* (pp. 1171–1174). Lisbon: European Language Resources Association. Retrieved from <http://www.lrec-conf.org/proceedings/lrec2004/pdf/48.pdf>
- Paquot, M. (2010). *Academic Vocabulary in Learner Writing: From Extraction to Analysis*. London: Continuum International Publishing Group.
- Quirk, R., Greenbaum, S., Leech, G. N., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language* (1st, 2nd ed.). London: Longman.
- Rahman, Z. A. A. A. (2013). The use of cohesive devices in descriptive writing by Omani student-teachers. *SAGE Open*, 3(4), 1–10. <https://doi.org/10.1177/2158244013506715>
- Rayson, P. (n.d.). Log-likelihood and effect size calculator. Retrieved from <http://ucrel.lancs.ac.uk/llwizard.html>
- Reinhart, T. (1980). Conditions for Text Coherence. *Poetics Today*, 1(4), 161–180. <https://doi.org/10.2307/1771893>
- Sanczyk, A. (2010). *Investigating argumentative essays of English undergraduates studying in Poland as regards their use of cohesive devices*. University of Oslo.
- Sanders, T. J. M., & Pander Maat, H. L. W. (2006). Cohesion and Coherence: Linguistic Approaches. In *Encyclopedia of Language & Linguistics* (pp. 591–595). <https://doi.org/10.1016/B0-08-044854-2/00497-1>
- Shea, M. (2009). A Corpus-Based Study of Adverbial Connectors in Learner Text. *MSU Working Papers in SLS*, 1(1), 1–13. Retrieved from <https://bit.ly/2UZkkGf>
- Tankó, G. (2014). The use of adverbial connectors in Hungarian university students' argumentative essays. In J. M. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 157–181). Amsterdam: Benjamins. <https://doi.org/10.1075/scl.12.13tan>
- Tapper, M. (2005). Connectives in advanced Swedish EFL learners' written English – preliminary results. In F. Heinat & E. Klingvall (Eds.), *Working papers in linguistics* (pp. 115–144). Lund: The Department of English, Lund University. Retrieved from <https://bit.ly/2GnnYI4>
- Witte, S. P., & Faigley, L. (1981). Coherence, Cohesion, and Writing Quality. *College Composition and Communication*, 32(2), 189–204. <https://doi.org/10.2307/356693>
- Xu, Y., & Liu, Y. (2012). The Use of Adverbial Conjuncts of Chinese EFL Learners and Native Speakers—Corpus-based Study. *Theory and Practice in Language Studies*, 2(11), 2316–2321. <https://doi.org/10.4304/tpls.2.11.2316-2321>
- Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 23(1), 31–48. <https://doi.org/10.1016/j.linged.2011.09.004>

- Yeqing, L. (2014). *A Corpus-Based Contrastive Study of Adverbial Connectors in English Majors' Academic Theses*. Xi'an International Studies University. Retrieved from <https://goo.gl/z7Q14Y>
- Zhang, M. (2000). Cohesive Features in the Expository Writing of Undergraduates in Two Chinese Universities. *RELC Journal*, 31(1), 61–95.
<https://doi.org/10.1177%2F003368820003100104>

Appendix 1: essay topics

The topics below stem from the *International Corpus of Learner English v2 Manual* by Granger, Dagneaux, Meunier and Paquot (2009):

- Crime does not pay.
- The prison system is outdated. No civilised society should punish its criminals; it should rehabilitate them.
- Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.
- A man/woman's (salary) should be commensurate with their contribution to the society they live in.
- The role of censorship in Western society.
- Karl Marx said that religion was the opium of the masses. If he were alive today, he would replace religion with the internet.
- All armies should consist entirely of professional soldiers: there is no value in a system of military service.
- George Orwell wrote, "All men are equal: but some are more equal than others." How true is this today?
- Money is the root of all evil.
- 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?
- In our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion?
- Feminists have done more harm than good to the cause of women.

Appendix 2: on USB stick

The following files are to be found on the USB stick included with this dissertation:

- The entire master's thesis;
- The *Corpus of Flemish Learners of English*;
- The spreadsheet containing the data of all the connectors analysed;
- The SPSS encoding data of the participants;
- The survey filled in by the participants is available through the following URL:
https://kuleuven.eu.qualtrics.com/jfe/form/SV_6x5beVSAjnzEWAR