

VILNIUS UNIVERSITY

Lina Bikeliene

CONNECTOR USAGE IN NATIVE AND NON-NATIVE LEARNERS' ENGLISH  
WRITING. CONTRASTIVE ANALYSIS

Summary of doctoral dissertation  
Humanities, Philology (O4 H)

Vilnius, 2012

This doctoral dissertation was written at Vilnius University in the Faculty of Philology in 2006-2012.

**Research supervisor:**

Assoc. Prof. Dr. Jonė Grigaliūnienė (Vilnius University, Humanities, Philology – 04 H)

The dissertation will be defended at the Council of Philology of Vilnius University.

**Chair:**

Prof. Dr. Meilutė Ramonienė (Vilnius University, Humanities, Philology – 04 H)

**Members:**

Prof. Habil. Dr. Ineta Dabašinskienė (Vytautas Magnus University, Humanities, Philology – 04 H)

Assoc. Prof. Dr. Nijolė Maskaliūnienė (Vilnius University, Humanities, Philology – 04 H)

Assoc. Prof. Dr. Jūratė Ruzaitė (Vytautas Magnus University, Humanities, Philology – 04 H)

Dr. Loreta Vaicekauskienė (Vilnius University, Humanities, Philology – 04 H)

**Opponents:**

Assoc. Prof. Dr. Inesa Šeškauskienė (Vilnius University, Humanities, Philology – 04 H)

Assoc. Prof. Dr. Violeta Kalėdaitė (Vytautas Magnus University, Humanities, Philology – 04 H)

The public defence of the dissertation is to be held in the meeting of the Council of Philology of Vilnius University at 12 o'clock on 29 June 2012 in the Faculty of Philology, Vilnius University.

Address: Universiteto St. 5, LT-01513 Vilnius.

The summary of the dissertation was sent out to relevant institutions on \_\_\_\_ May 2012.

The dissertation is available at the library of Vilnius University.

VILNIAUS UNIVERSITETAS

Lina Bikeliene

TEKSTO JUNGIMO PRIEMONIŲ VARTOSENA GIMTAKALBIŲ IR  
NEGIMTAKALBIŲ STUDENTŲ RAŠTO DARBUOSE ANGLŲ KALBA.  
GRETINAMOJI KONEKTORIŲ ANALIZĖ

Daktaro disertacijos santrauka  
Humanitariniai mokslai, filologija (04 H)

Vilnius, 2012

Disertacija rengta 2006-2012 metais Vilniaus universiteto Filologijos fakultete.

**Mokslinis vadovas:**

doc. dr. Jonė Grigaliūnienė (Vilniaus universitetas, humanitariniai mokslai, filologija  
04 H)

**Disertacija ginama Vilniaus universiteto Filologijos mokslo krypties taryboje:**

**Pirmininkas** – prof. dr. Meilutė Ramonienė (Vilniaus universitetas, humanitariniai  
mokslai, filologija – 04 H)

**Nariai:**

prof. habil. dr. Ineta Dabašinskienė (Vytauto Didžiojo universitetas, humanitariniai  
mokslai, filologija – 04 H)

doc. dr. Nijolė Maskaliūnienė (Vilniaus universitetas, humanitariniai mokslai, filologija  
– 04 H)

doc. dr. Jūratė Ruzaitė (Vytauto Didžiojo universitetas, humanitariniai mokslai, filologija  
– 04 H)

dr. Loreta Vaicekauskienė (Vilniaus universitetas, humanitariniai mokslai, filologija –  
04 H)

**Oponentai:**

doc. dr. Inesa Šeškauskienė (Vilniaus universitetas, humanitariniai mokslai, filologija –  
04 H)

doc. dr. Violeta Kalėdaitė (Vytauto Didžiojo universitetas, humanitariniai mokslai,  
filologija – 04 H)

Disertacija bus ginama viešame Filologijos mokslo krypties tarybos posėdyje 2012 m.  
birželio mėn. 29 d. 12 val. Vilniaus universiteto Filologijos fakulteto V. Krėvės  
auditorijoje.

Adresas: Universiteto g. 5, LT-01513, Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2012 m. gegužės \_\_\_\_ d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje.

# CONNECTOR USAGE IN NATIVE AND NON-NATIVE LEARNERS' ENGLISH WRITING. CONTRASTIVE ANALYSIS

## INTRODUCTION

**The object of the research** is text connectors. Inadequate use of these text linking devices that can highlight the writer's attitude towards the relationship between individual linguistic elements such as contrast, result, etc. (Quirk *et al.* 1991: 631–632, Biber *et al.* 1999: 875) is a wide-spread phenomenon even in advanced students' writing. This not only widens the gap between learner interlanguage and the target language but also, as Granger (1999: 192) points out, plays an important role in 'foreign-soundingness'. This study therefore was designed to analyse connector usage in non-native speakers' (NNSs) and native speakers' (NSs) written essays. A list of 129 connectors from Quirk *et al.* (1991: 631–632) and Biber *et al.* (1999: 875) was compiled. The problem of multifunctionality which is undoubtedly present in the study of text connectors was dealt with by classifying connectors according to their prototypical meaning.

**The goals of the dissertation** were set as follows:

1. To analyse text connector usage in Lithuanian learners' argumentative essays.
2. To compare and contrast connector usage in Lithuanian learners' essays to that in the target language, i. e. the English language of native speakers.
3. To describe the peculiarities of Lithuanian learners' connector usage in the context of learners with other mother-tongue backgrounds and to highlight the possible reasons for such usage.
4. To define the ways to improve Lithuanian EFL learners' use of connectors.

To achieve these goals, the following **research tasks** were set:

1. To compile a corpus of written English produced by Lithuanian EFL learners of advanced level.
2. To analyse quantitative connector usage parameters characteristic of both general and individual semantic categories of connectors.
3. To define sentence and text position of connectors in NNSs and NSs students' argumentative essays.
4. To perform a statistical significance analysis of the quantitative results.

5. To compare the results of connector usage in the Lithuanian EFL learners' argumentative essays to the results retrieved from corpora of learners with different mother-tongue backgrounds.
6. To analyse the reasons for the differences in Lithuanian and other EFL learners' connector usage.

**The novelty and relevance of the research.** Consistent EFL research started more than half a century ago but the scope was largely limited due to the manual analysis methods employed at the time. The first computerised corpora eliminated the size problem of the empirical material. The studies, however, were mainly concentrated on the analysis of general English. The situation largely changed with the start of the International Corpus of Learner English (ICLE) project at the end of the last century. It gave a start to wider possibilities for the analysis of EFL learner language and for establishing the features characteristic of one or several groups of EFL learners. In 2011, Hinkel (2011: 524) presented a list of languages that had EFL learner written discourse already analysed. The compiled list is not exhaustive but the reason for the absence of the Lithuanian EFL learner language in it is due to the fact that there have been few attempts so far to analyse the Lithuanian EFL learners' written discourse. This research, retrieving the data from a recently compiled Corpus of Lithuanian Learner English (LICLE), is one of the first attempts in Lithuania to address the topic of text connectors as one of the ways to ensure text cohesion in Lithuanian EFL students' written discourse. Hopefully it will play a significant role in the creation of educational materials for Lithuanian EFL learners as well as being a platform for further research to determine the role of an individual connector in text coherence and to determine the correlation between the use of text connectors and the level of an EFL learner.

The following **theses** are to be defended:

1. The quantitative analysis of the use of connectors retrieved from the NNSs and NSs corpora allows for determining the tendency of the Lithuanian EFL learners to overuse overt marking of semantic text linking relationships with the use of connectors.
2. The analysis of connectors in NSs and NNSs corpora shows that NNSs as well as NSs tend to overtly mark relationships of the same semantic categories: listing, contrast, result and apposition.

3. The deviation of the Lithuanian EFL learners' connector usage from the one observed in native students' essays can sometimes be considered as marking positive tendencies – connector usage in the LICLE corpus at times resembles the use in academic sections of BYU-BNC and COCA more than the use in the subcorpora of LOCNESS.
4. The Lithuanian EFL students' use of connectors does not coincide with the use of learners with other mother-tongue backgrounds. The effectiveness of teaching material oriented towards an unspecified general group of EFL learners therefore is questionable.
5. More detailed and accurate presentation of connectors in educational material could decrease the problems of overuse, underuse and misuse in the context of frequencies and sentence as well as text position of connectors.

### **Review of the earlier research.**

Numerous studies in learners' written discourse have been carried out on the topic of connectors as one of the most important text features that help in the processing of information (Schmied 2009: 1149). Connector studies can be classified into corpus (Liu 2008; Wu 2006; Chen 2006 *etc*) and non-corpus based (Bondi 2004; Barton 1995 *etc*). Corpus based studies of connectors can be further grouped according to the type of corpus used, e.g. EFL (Granger and Tyson 1996) or other types of corpora (Altenberg 2007). Classification of connector research can also be based on whether the use of connectors was studied in writing of EFL learners of one language group (Altenberg and Tapper 1998) or whether written work of EFL learners with different mother-tongue backgrounds was addressed (Ying 2007; Leńko-Szymańska 2007; Rica Peromingo 2008, Gilquin and Paquot 2007). Studies dedicated to one individual semantic category of connectors (Field 1993) can also be distinguished. One more important aspect of connector studies is the analysis of usage in texts of different disciplines (Fortuńo 2006; Pastor 2006, Tang and Ng 1995) or the comparison of connector usage in texts representing different EFL levels (Neff-van Aertselaer and Dafouz-Milne 2008; Fei 2006).

In Lithuanian linguistics, connectors, or words carrying their function, were mainly analysed in the Lithuanian language (Murinienė 2005; Bieliniskienė 2009; Bitinienė 2009; Alaunienė and Valskys 2009; Bartkutė 2007 *etc.*). The Lithuanian learners' EFL language and connector usage in particular due to objective reasons – the LICLE corpus was completed only in autumn 2010 – have hardly been addressed yet.

One of the most detailed studies of connectors in Lithuanian EFL language was carried out by Burneikaitė (2008, 2009a, 2009b) and Bikelienė (2008a, 2008b, 2009, 2010). Though both authors note the overuse of connectors in the Lithuanian EFL learners' writing, their treatment of connector usage by NSs differs. Burneikaitė (2008) considers it to be not the norm but the manifestation of different cultural and academic writing traditions. Bikelienė (2008a, 2008b, 2009, 2010), on the contrary, follows Leech (1998: xv) in the belief that for quantitative research and comparison of NNSs language it is of vital importance to set a standard or norm according to which NNSs corpus data is evaluated. Following this tradition, the deviation from NSs results is considered to be a deviation from the norm. An important insight made in Burneikaitė's (2008) study concerning the different role of academic writing traditions in different universities of Lithuania cannot be explicated due to the nature of the corpus used for the present work. It would however be a very interesting aspect for future research.

In Juknevičienė's (2011) PhD thesis dedicated to the study of lexical bundles, the use of connectors in the Lithuanian EFL learners' and NSs language was addressed to a certain extent. Though the author's study included only a small number of connectors (*in conclusion, in addition, for example; on the one hand, on the other hand* etc.), the results indicate the already observed tendency of connector overuse in the Lithuanian EFL language.

Unlike in Lithuanian, in foreign linguistics, connector studies have enjoyed research popularity. Though the topic of this work presupposes greater interest in EFL connector studies, such research has not been limited to the English language only (e. g. Slovene and Croatian [Bulc 2005]).

One of the pioneering connector studies was the work carried out by Crismore (1980). In her study, the author examined how five logical connectors were used by one hundred students of different English language proficiency levels. The results indicating the problems of connector usage in NNSs writing were later confirmed by a number of linguists (Crewe 1990; Granger and Tyson 1996; Altenberg and Tapper 1998; Tankó 2004 *etc.*).

Crismore's (1980) attempt to classify connectors according to their level of complexity was original and has not been repeated in other studies. The explanations of complexity according to the phonetic structure of connectors – *moreover* as a multi-



syllable connector is more difficult to learn than a one syllable *still* – though original are rather subjective. How can one objectively classify connectors into easier and more difficult to learn? Even if this task is accomplished, can the number of syllables be considered as playing the decisive role? The use of corpora for the retrieval of connectors does not allow the researchers to duplicate Crismore's study asking the learners to provide synonyms for the connectors studied. All these reasons must have determined the abandoning of Crismore's methodology in future studies.

Before the wide use of corpora, connector research like the study of Crismore (1980) was of a very narrow scope: Connor's (1984) generalisation is based on the analysis of six, Khalil's (1989) – twenty, essays. In order to get more accurate results it was of vital importance to broaden the scope of empirical material. In this context an important role is played by the ICLE project, part of which is the LICLE corpus. ICLE not only gave a start to the broad analysis of learner English but also allowed the comparison of the learner language to the target corpus of native students' essays – LOCNESS. Granger and Tyson's (1996) study on the use of connectors in French EFL and NSs essays has set the tone for future research on the subject. Their work has been cited and used as a comparison by a number of linguists (Altenberg and Tapper 1998, Narita *et al.* 2004 *etc.*). The authors raised a hypothesis of connector overuse in NNSs written discourse. The study, however, rejected it since no significant evidence was found to account for the overall overuse or underuse of connectors in the French EFL learners' essays: one connector in NNSs essays corresponds to 1.09 connectors in NSs essays. The results clearly contradict the initial hypothesis – the French EFL learners tend to slightly underuse connectors in their writing. The authors also noted some significant deviations in the use of some individual connectors in the learner language. Testing the results with a Log Likelihood (LL) calculator, LL values do not leave space for doubts: *moreover* (LL +56.40) and *for instance* (LL +58.82) are overused while *however* (LL -122.29) and *therefore* (LL -69.22) – underused in the French EFL learners' essays.

Though the model used by Granger and Tyson (1996) has often been used by other linguists, Bolton *et al.* (2002: 172) criticise their work mainly due to the methodology. The authors claim (*ibid.*), that the primary function of connectors to connect linguistic elements above sentence level is clearly ignored in the method based on the use of

quantitative results. As an alternative it has been suggested that calculations should focus on connector frequencies per sentence instead of frequencies per corpus. The results acquired using this methodology could be informative but almost incomparable with the research of other learner groups. This is the reason for rejecting Bolton *et al.* (2002) and opting for Granger and Tyson's (1996) methods for the dissertation.

Bolton *et al.* (2002) also expressed criticism of Granger and Tyson's (1996) research on the use of the LOCNESS corpus. The corpus was blamed for being insufficiently representative of the NSs language. The criticism of LOCNESS is a recurrent phenomenon, but bearing in mind that it was compiled following the same criteria as ICLE, with some limitations, it has remained one of the most suitable corpora for interlanguage studies. In order to get more objective and representative results this dissertation occasionally uses data from one more corpus which was compiled to represent the written discourse of British students – the BAWE corpus.

Altenberg and Tapper (1998) analysing connector usage in the Swedish subcorpus of ICLE noticed a similar tendency to the one described by Granger and Tyson (1996) – Swedish EFL students as well as French EFL students underuse connectors in their argumentative essays. The given frequencies allow for testing the statistical significance of the reported underuse which, contrary to that observed in the French subcorpus, appears to be statistically significant (LL -16.03). It can be noted that EFL learners of both mother-tongue backgrounds tend to overuse (*moreover, for instance, on the contrary*) and underuse (*hence, therefore, however, though*) the same groups of connectors. This gives some basis for talking about a set of features characteristic of advanced level EFL interlanguage. This hypothesis is supported by the results retrieved from the LICLE corpus. It can be noted that Tapper's (2005) research on the use of connectors by the Swedish EFL learners contradicts the results published in Altenberg and Tapper's (1998) study. The most likely explanation is the difference in subcorpora of LOCNESS used. While Altenberg and Tapper (1998) used the British, Tapper (2005) retrieved data for comparison from the American subcorpus of LOCNESS. This exemplifies the difference of connector usage in essays written by British and American students. Due to this reason, in this dissertation, unlike the method followed by Juknevičienė (2011), it was decided to analyse the subcorpora of LOCNESS separately.

General underuse tendencies reported by Altenberg and Tapper (1998) stand aside in the context of EFL learners' research on connector usage. The majority of the studies on the topic can be classified into two groups:

- (1) reporting general overuse in NNSs writing (Tankó 2004; Milton 2001; Yoon 2006) or
- (2) Granger and Tyson (1996) alike, reporting on the deviation of individual connector usage in EFL learners' writing (Narita *et al.* 2004, Tang and Ng 1995).

This dissertation aims to compare the use of connectors in the Lithuanian EFL learners' writing with the usage of learners with different mother-tongue backgrounds. It can be hypothesised that the more culturally or linguistically groups are related, the fewer differences in the use of connectors are observed.

In linguistic literature, numerous mentioning of the problematic usage of connectors is often ascribed to the characteristics of interlanguage (Granger 1999: 192). There are attempts to classify the problems of connector usage into connector overuse (when connectors are not obligatory and their use does not add to the quality of the text) and misuse (when the use of connectors misleads the reader) [Milton and Tsang 1993]. Since the latter is almost a universal feature of EFL learners' writing (Crewe 1990: 317), attempts have been made to clarify the reasons for such misuse. One of the prevalent reasons is inappropriate presentation of connectors in academic literature (Narita *et al.* 2004; Fei 2006). Institutionally induced reasons, native language influence and the possible role of cultural factors should also be taken into consideration (Milton 2001). This dissertation is one more attempt to shed light on the possible reasons for the problem.

No agreement has been reached on the degree of correlation of the effective use of connectors and the level of the writer (Granger and Tyson 1996: 17). While Tierney and Mosenthal (1983: 225), Crewe *et al.* (1985), Neuner (1987), Basturkmen (2002: 52) and Hartnett (1986: 144) try to ascertain that connectors cannot guarantee text coherence, other linguists note their vital importance for text understanding. According to Barnet and Stubbs (1995: 62), Richard Wagner compared opera writing to an art of transition from one note to another, from one scene to another. Similarly to transitions that add to the unity of an opera, text connectors called 'landmarks' of a discourse by Leech and

Svartvik (1994: 177) assist the reader in connecting individual parts of a text into one coherent structure. The **proper** use of connectors undoubtedly adds to greater clarity and understandability of the discourse (Altenberg and Tapper 1998: 80) and therefore deserves to be analysed in the written language of each and every EFL learner group.

## DATA AND METHODS

The data for the research was retrieved from the Lithuanian subcorpus of ICLE compiled at Vilnius University and the LOCNESS corpus. At the time, LICLE consisted of 268 essays written by 216 students, 154 992 words (Table 1) on argumentative (91.27 per cent – 141 457 words) and literary topics (8.73 per cent – 13 535 words).

Table 1. Overall figures for the LICLE corpus

	Number of essays	Number of words
<b>Third study year essays</b>	110	69 178
<b>Fourth study year essays</b>	158	85 814
<b>Total</b>	268	154 992

In order to determine the reasons for the Lithuanian EFL learners' connector usage, for the analysis of individual connectors, the corpus of first year students of English Philology at Vilnius University compiled by Juknevičienė was used. It consists of 293 essays, 118 552 words.

For the comparison to be informative it is essential to choose “corpora that are comparable in text type” (Granger and Tyson 1996: 24). This determined the choice of LOCNESS as the corpus compiled following the same guidelines as LICLE. Consequently, following Leech (1998: xv), the LOCNESS corpus, if not mentioned otherwise, is considered “as a standard of comparison, or norm, against which to measure the characteristics of the learner corpora”.

The sample corpus of LOCNESS consists of the British subcorpus LOCNESS-BR and the American subcorpus LOCNESS-US. The scope of the former is 90 essays, 95 695 words. The scope of the latter is 232 essays, 168 231 words. The third subcorpus of LOCNESS, A-level, consists of British A-level essays. The scope is 30 essays, 15 434 words.

In order to clarify tendencies and make broader generalisations, LOCNESS data was supplemented by data retrieved from the following NSs corpora:

- (1) BYU-BNC (*British National Corpus*) –96 263 399 words. Academic section consists of 15 331 668 words.
- (2) COCA (*Corpus of Contemporary American English*) – 437 785 714 words. Academic section consists of 85 791 918 words.
- (3) BAWE (*British Academic Written English*) –6 696 026 words.

The data from the Lithuanian language and its academic register was retrieved from the following corpora:

- (1) DLKT (*Corpus of Contemporary Lithuanian*) –140 921 288 words. Its non-fiction section consists of 21 024 249 words.
- (2) CorALit (*Corpus of Academic Lithuanian*) – approximately 9 000 000 words.

**Research methods** include contrastive interlanguage analysis (CIA) and statistical data analysis. CIA (as a type of Contrastive Analysis) was chosen since it “establishes comparisons <...> between native and learner varieties of one and the same language” (Granger 1996: 43). CIA comprises both quantitative and qualitative aspects of analysis (Granger 2009: 18). The former contributes to language description and its pedagogy (Guo 2006: 221), while the latter allows for the evaluation of generalisability of interlanguage (Granger 2009: 18).

The LICLE corpus is unannotated therefore data retrieval was carried out by using a combinatory method: connectors were retrieved with the help of TextSTAT 2.5. (Hüning 2000/2007) and AntConc 3.2.1w (Anthony 2007) software and then instances of non-connectors were eliminated manually (for example, the adjectival use of *first*).

Log-likelihood (LL) calculator (<http://ucrel.lancs.ac.uk/llwizard.html>) was used to test the statistical importance of the observed frequencies ( $p < 0.01$ , critical value = 6.63). Due to the different size of the corpora under consideration normalised frequencies per 10 000 words were also used.

## RESEARCH FINDINGS

Results of quantitative overview of connectors. The study showed that from a compiled list of 129 types of connectors only slightly more than 50 per cent were actually used in the LICLE and LOCNESS subcorpora. 69 different types of connectors

were found in LICLE, 67 different types in LONCESS-BR and 66 in LOCNESS-US (Table 2).

Table 2. Connector frequencies in LICLE, LOCNESS-BR and LOCNESS-US

	LICLE	LOCNESS-BR	LOCNESS-US
Raw frequency of connectors	2712	1381	1549
Normalised frequency of connectors per 10 000 words	175	144	92
Different types of connectors	69	67	66

Though both NNSs and NSs rely on a similar number of different types of connectors in their argumentative writing, which allows hypothesising about a group of connectors characteristic to argumentative essays, the quantitative analysis of the connectors used showed statistically significant overuse in the LICLE corpus compared to the British as well as American subcorpora of LOCNESS, LL +34.36 and LL +426.44 respectively.

Following Quirk's *et al.* (1991) classification, connectors were grouped into seven semantic categories (Table 3). The quantitative analysis confirmed the same tendencies to rely greatly on the overt marking of listing, contrastive, resultive and appositive semantic relations both in the Lithuanian EFL and in NSs argumentative essays. The comparison with data reported on the distribution of connectors into semantic categories in EFL corpora of other linguistic groups shows the absence of unanimity: the four most common semantic categories observed in LICLE and LOCNESS coincide with the categories in the Swedish subcorpus of ICLE (SWICLE) [Altenberg and Tapper 1998] but not in the Hungarian EFL corpus (Tankó 2004) confirming the status of language as a cultural phenomenon.

Table 3. Raw and normalised frequencies of connectors by semantic categories

Category	LICLE	LOCNESS-BR	LOCNESS-US
1. Listing	987 (64) <sup>1</sup>	679 (40)	416 (43)
2. Contrastive	667 (43)	361 (21)	438 (46)
3. Resultive	581 (37)	322 (19)	372 (39)

<sup>1</sup> Normalised frequency per 10 000.

4.	<b>Appositive</b>	276 (18)	112 (7)	103 (11)
5.	<b>Summative</b>	160 (10)	11 (1)	6 (1)
6.	<b>Inferential</b>	29 (2)	17 (1)	16 (2)
7.	<b>Transitional</b>	12 (1)	47 (3)	30 (3)
<b>Total</b>		2712 (175)	1381 (144)	1549 (92)

The comparison of the top ten connectors in LICLE and LOCNESS as well as in other EFL corpora indicates some tendencies in connector usage (Table 4):

1. There is a set of connectors favoured by all (*however*) or nearly all the groups under consideration (*for example, thus, therefore*).
2. All the top ten connectors used by the Lithuanian EFL learners appear in at least one other corpus in Table 4.
3. All the groups rely heavily on the top ten list: it accounts for about three quarters of all the connectors used in the corpora. The range is between 77 and 73 per cent in LOCNESS-US and LOCNESS-BR respectively and 60 per cent in SWICLE. The LICLE corpus results with 67 per cent lie between the two extremes.

There have been a number of studies reporting on the differences of sentence position of connectors in EFL learners' and English NSs writing (Tankó 2004: 176, Granger and Tyson 1996: 24, Narita *et al.* 2004: 1173, Anping 2000: 46, Altenberg and Tapper 1998: 89 *etc.*). The Lithuanian EFL learners follow the same tendency for EFL learners to choose a non-preferred sentence initial (SI) position of connectors. In LICLE 61.1 per cent of all the connectors are used sentence initially while in LOCNESS-BR it attests for only 26.3 and in LOCNESS-US for 59.0 per cent of all the use. The reasons for EFL learners choosing SI position are manifold. On the one hand, this position can be considered as the safest possible ensuring at least surface coherence. On the other hand, this usage can be inspired by educational material where it is presented as the norm or standard usage. The necessity to warn students against the non-preferred positions arises from the fact that different sentence positions can lead to different argumentation strategies which can be less common to the English writing tradition.

Table 4. Top ten connectors

	LICLE	LOCNESS-BR	LOCNESS-US	SWICLE <sup>2</sup>	PICLE <sup>3</sup>	HUNGARIAN EFL <sup>4</sup>	CHINESE EFL <sup>5</sup>	TAIWANESE EFL <sup>6</sup>
1.	<i>however</i>	<i>however</i>	<i>also</i>	<i>for example</i>	<i>also</i>	<i>however</i>	<i>First</i>	<b><i>however</i></b>
2.	<i>also</i>	<i>also</i>	<i>however</i>	<i>however</i>	<i>however</i>	<i>also</i>	<i>Second</i>	<b><i>therefore</i></b>
3.	<i>for example</i>	<i>therefore</i>	<i>so</i>	<i>of course</i>	<i>therefore</i>	<i>therefore</i>	<i>However</i>	<i>for instance/ <b><i>for example</i></b></i>
4.	<i>thus</i>	<i>so</i>	<i>therefore</i>	<i>so</i>	<i>for example</i>	<i>thus</i>	<i>secondly</i>	<b><i>thus</i></b>
5.	<i>so</i>	<i>thus</i>	<i>for example</i>	<i>therefore</i>	<i>so</i>	<i>furthermore</i>	<i>for example</i>	<i>moreover</i>
6.	<i>therefore</i>	<i>for example</i>	<i>yet</i>	<i>thus</i>	<i>thus</i>	<i>moreover</i>	<i>Although</i>	<i>besides</i>
7.	<i>moreover</i>	<i>yet</i>	<i>again</i>	<i>for instance</i>	<i>moreover</i>	<i>secondly</i>	<i>Though</i>	<b><i>also</i></b>
8.	<i>first of all</i>	<i>again</i>	<i>thus</i>	<i>that is</i>	<i>on the other hand</i>	<i>though</i>	<i>Finally</i>	<i>first</i>
9.	<i>of course</i>	<i>on the other hand</i>	<i>finally</i>	<i>still</i>	<i>consequently</i>	<i>in addition</i>	<i>Firstly</i>	<i>then</i>
10.	<i>on the other hand</i>	<i>too</i>	<i>too</i>	<i>furthermore</i>	<i>nevertheless</i>	<i>first of all</i>	<i>of course</i>	<i>in addition</i>

<sup>2</sup> From Altenberg and Tapper (1998)

<sup>3</sup> From Leńko-Szymańska (2007).

<sup>4</sup> From Tankó (2004).

<sup>5</sup> From Chen (2006).

<sup>6</sup> From Ai and Peng (2006).



### Semantic categories of connectors.

Though listing relations are the most often overtly marked relations in the LICLE and LOCNESS-US corpora and occupy the second position in LOCNESS-BR, significant overuse was recorded in LICLE comparing to LICLE-BR (LL+44.71) and comparing to LICLE-US (LL+85.31). The overuse of this category can be ascribed to interlanguage characteristics since the same tendency was reported in writings of Korean (Yoon 2006), Chinese (Yachoen 2006, Ai and Ping 2006), Taiwanese (Chen 2006), Hungarian (Tankó 2004) EFL learners. The LICLE results most closely resemble results of the A-level corpus. This can be attested by the requirements to place special attention on formal structuring strategies of the written work.

*Moreover* is the most overused connector in the LICLE corpus, LL +96.30 and +127.09 comparing to the British and American subcorpora of LOCNESS respectively. Though having a clear equivalent in the Lithuanian language, it proves to be problematic for Lithuanian learners. LICLE data indicates that *moreover* is often used for connecting paragraphs instead of propositions within a paragraph or for simply adding new or reformulation the previously stated information. The information presented in dictionaries and grammars can attest for such misuse since only *Macmillan English Dictionary for Advance Learners* (2007) comments on the functions of *moreover*.

The contrastive category of connectors also belongs to the most overtly marked categories in LICLE and LOCNESS. The quantitative results, though, show insignificant underuse in LICLE in comparison to LOCNESS-BR (LL -1.00), while comparison to LOCNESS-US yield significant overuse (LL +119.29). The difference between the results of the NSs corpora is due to the use of *however*, the most widely used connector both in LICLE and LOCNESS-BR. It belongs to so-called ‘pet-connectors’ (Tankó 2004) – it is used in more than 75 per cent of all the essays in LICLE and LOCNESS-BR. The study indicated the correlation of the level and maturity of writers and the frequency of the connector under consideration: the higher the level, the higher the frequency of the connector in writing. This correlation is applicable to both Lithuanian EFL learners and NSs of English. Overuse of *however* signals not only the lack of knowledge of other linking

devices but argumentation problems as well when counterarguments are presented without proper explication.

The third category most common to argumentative essays – resultive connectors – duplicates the tendencies observed for the category of contrastive connectors. In LICLE, slight underuse can be noted comparing to LOCNESS-BR (LL -0.30) and significant overuse comparing to LOCNESS-US (LL +98.09). The study shows that the American students compensate the lower frequency of resultive connectors by other linguistic devices, e.g. lexeme *result\** is 1.5 times more often used in LOCNESS-US than LOCNESS-BR.

Two resultive connectors appearing in the top ten list, *therefore* and *thus*, indicate the existence of two different models of usage in NSs academic discourse: the American with the dominant use of *thus* and the British where the dominant position is occupied by *therefore*. The Lithuanian learners' data suggest the insufficient incorporation of the academic model and stronger reliance on general English norms.

Appositive connectors as a “recurrent feature of academic writing <...>, a part of the routine ways in which writers in all fields seek to make their ideas accessible and persuasive” (Hyland 2007: 270) show an overuse tendency in EFL writing (Tankó 2004: 172, Altenberg and Tapper 1998: 85, Tapper 2005: 123, Hinkel 2002: 147 *etc.*). LICLE data suggests significant overuse compared to both British (LL +20.37) and American (LL +85.64) varieties of student writing. The most significant differences were observed in the use of *for example* and *for instance*. Both of them are significantly overused in the LICLE corpus though their ratio almost corresponds to the ratio observed in the academic sections of the BYU-BNC and COCA while being significantly different from the ratio in LOCNESS-BR. The results indicate positive tendencies in the ratio of connectors under consideration comparing the Lithuanian EFL corpora of different levels of proficiency. The overuse of neither of the two connectors can be attributed to the learners' mother-tongue influence since their equivalents were found to be more scarcely used both in the DLKT and in the CorALit corpora.

Summative connectors belong to significantly overused connectors in the LICLE corpus (LICLE vs. LOCNESS-BR: LL +113.80, LICLE vs. LOCNESS-US: LL +167.92).

Research results for this semantic category of connectors indicate the greatest resemblance of quantitative data of the LICLE corpus to A-level data (LL +0.05), the normalised frequencies being 10.3 and 9.7 respectively as compared to 0.6 in LOCNESS-BR or 1.5 in BAWE. The language and maturity level thus can be seen as playing important roles in interpreting the results. The high frequencies of summative connectors cannot be ascribed to teaching material induced features (cf. Oshima and Hogue 1991:82) since the results in LICLE show almost a double decrease compared to first year students' writing. The nature of summative connectors accounts for the corresponding text position observed in LICLE and both subcorpora of LOCNESS where they tend to be used in the first sentence of the last paragraph.

The Lithuanian learners are noted for indicating summative relations with the help of *in conclusion* or *to conclude* which show the same frequencies in the LICLE corpus. Aiming at more native-like academic English, Lithuanian students should be made aware of the prevalent use of *in conclusion* in such contexts in the academic sections of BYU-BNC and COCA.

The inferential category of connectors presents no significant differences: LICLE vs. LOCNESS-BR (LL +0.13) and LICLE vs. LOCNESS-US (LL +4.23). The sentence position of inferential connectors, however, replicated the observed tendency of SI usage in learner writing. The Lithuanian learners' use of individual connectors resembles the use in the LOCNESS-BR corpus. Comparison with LOCNESS-US yielded the only significant difference in the use of *in other words* (LL +11.90). It was used in only less than six per cent of essays in the LICLE corpus therefore the observed differences can be ascribed to individual learner style rather than justifying generalisations about the Lithuanian learners' usage.

Transitional connectors form the smallest category of text connectors. This category is significantly underused by the Lithuanian EFL learners (LL -19.07 and LL -19.42 comparing to LOCNESS-BR and LOCNESS-US respectively). The scarcity of the examples however does not allow for generalisations on their usage.

## CONCLUSIONS

The dissertation focused on the use of (text) connectors in Lithuanian EFL students' and native English speakers' written language corpora. The comparison with EFL groups of different mother-tongue backgrounds was also made. The results permit conclusions to be drawn about the Lithuanian EFL characteristics of overt marking of semantic relations with connectors, the relationship with other EFL linguistic groups in connector usage, the existence of a universal model of reasoning, reasons that might have led to the results and the ways in which EFL language could be brought closer to NSs English.

Quantitative analysis of the connectors indicates significant overuse of overt marking of semantic relations in the LICLE corpus in comparison to the British and American LOCNESS subcorpora, LL +34.36 and LL +423.97 respectively. The results show that the Lithuanian EFL students are more likely not only to use connectors but to do so in a non-preferred position of a sentence. In the LICLE corpus, the preferred position is sentence initial (SI) [61 per cent] while in both subcorpora of LOCNESS, the preferred position is sentence medial (SM): in LOCNESS-BR and LOCNESS-US, 71 and 59 per cent respectively. Although the text position of connectors appeared to cause fewer problems for the Lithuanian EFL learners than sentence position, it can be noted that some semantic categories of connectors were used in parts of the text not common to NSs writing, for example, *for instance* in the concluding paragraph. The results thus indicate not only the need to encourage Lithuanian EFL learners to choose different strategies for structuring the text instead of relying too heavily on overt marking of semantic relations with the help of connectors but also to make them aware of the most appropriate position for individual connectors in the text.

The study revealed differences between NSs language of the British and American subcorpora LOCNESS-BR and LOCNESS-US. British students overtly mark semantic relations with the help of connectors more than one and a half times more often than American students do. The data retrieved from the BAWE corpus, as a corpus representing the British academic written language, also differs from the results obtained from the LOCNESS-BR subcorpus. The comparison of the corpora data with data from the academic

sections of BNC-BYU and COCA does not allow for unambiguous interpretation of the Lithuanian EFL students' connector usage. The LICLE corpus deviations from NSs students' usage tendencies do not necessarily indicate negative trends. Sometimes, on the contrary, they indicate a positive tendency, since the Lithuanian EFL students' quantitative use of connectors is closer to the academic part of the national corpus (BYU-BNC or COCA) than to the student corpus of the English language. It can be noted that the Lithuanian students' usage of connectors occupies an intermediate position between the British and American students' usage. However, generalisations about greater similarities between LICLE and LOCNESS–BR than LICLE and LOCNESS–US can be drawn based on the quantitative results of different semantic categories of connectors, their diversity and the most widely used connectors. These results can be interpreted as indicating the Lithuanian EFL students' developmental phase, which can be characterised by the intersection of different language features.

The comparison of connector usage of the Lithuanian EFL and other EFL linguistic groups indicates that there is no universally common quantitative model for all the groups of EFL learners – the data from different EFL groups varies greatly and can be attributed to the distinctive features of that EFL learners' group. Frequencies of certain connectors can be also attributed to the distinctive features of the Lithuanian EFL learners. Although there is a group of generally overused (e.g, *moreover*, *on the contrary*) or underused (e.g, *hence*, *therefore*) connectors in NNSs writing, the LICLE corpus data contradicts the general tendencies, for example, when commonly overused *still* is underused by the Lithuanian EFL learners or commonly underused *however* in LICLE is used more often when in LOCNESS.

Despite the observed diverse tendencies of various EFL learner groups, several common features have been established. All the EFL groups under consideration appeared to overuse overt marking of relatively more simple semantic relations, such as listing, and to place connectors into sentence initial position.

Though differences of connector usage in NNSs and NSs were defined, the study suggests the existence of an almost universally accepted model of reasoning based on the marking of arguments, counterarguments and results. Connectors of these semantic

categories (listing, contrastive and resultive) most commonly are used to indicate semantic relations not only in the LICLE, LOCNESS-BR and LOCNESS-US corpora but in corpora of other EFL language groups, e.g. Hungarian or partly in SWICLE. This trend is considered to be almost universal, independent of the frequency of connectors in different corpora. It also provides a universal tendency to rely on a small group of connectors – the most commonly used connectors make up three-quarters of all the connectors used in the corpora.

With a view to revealing the possible explanations for the use of connectors in LICLE, several hypotheses were indirectly confirmed: native language transfer, the EFL language level, teaching materials and teaching influence. An analysis suggests that the use of connectors in Lithuanian EFL is only minimally affected by the Lithuanian language. It is noted that in cases where there is a clear equivalent of an English connector in the learners' mother-tongue, from several alternatives they are more likely to choose the latter.

It has been observed that the EFL level plays an important role in the use of connectors. On the one hand, the relatively easier categories (summative or the connector *moreover* belonging to the listing category) confirmed the hypothesis of an inverse correlation: the rising level of language proficiency and maturity of the writer is inversely proportional to connector frequency. The connector *however* belonging to a relatively more complex category is more scarcely used in the corpus of 1<sup>st</sup> year students of English Philology, Vilnius University. The same trend in the use of the NSs students' corpus indicates the usage as being related to the reasoning strategies. At an advanced level, writers tend to use more complex reasoning and present greater numbers of arguments. Similar tendencies can be seen in analysing the ratio of connectors belonging to “more complex” categories (e.g. *thus* and *therefore*). The obtained results justify speaking about the Lithuanian EFL connector usage as a reflection of linguistic development. The data retrieved from the LICLE corpus could be treated as a positive evaluation of the academic writing course taught at Vilnius University.

The study revealed that there is an ongoing problem of dictionaries and grammars presenting too scarce or even misleading information on the use of connectors. During the

study process therefore it is essential to provide students with clearer guidelines for the use of connectors, highlighting not only the differences in register but the most common contexts in NSs writing. For example, in the LICLE corpus, one fifth of the cases of *moreover* were in the first sentence of a paragraph while not once was this position occupied in LOCNESS.

Difficulties observed at advanced level EFL are rather subtle and unobtrusive (Nesselhauf 2004: 145), making it likely that this work will find practical application for improving academic writing courses by specifically designing them for Lithuanian EFL students. Hopefully, it will also help to determine which connectors require special attention in the context of Lithuanian EFL writing.

**Limitations of the study** are mainly twofold. The object of the research was only one aspect possibly ensuring text cohesion and coherence, namely, (text) connectors. This undoubtedly left out many other aspects such as lexical or grammatical elements, the addition of which could shed light on organisational strategies of learners' discourse. Correlation between the use of connectors and text quality could also be of great value for future studies.

Another limitation could be the size and nature of the learner corpora used for the study. The LICLE corpus data is representative of the language of only one study program in only one educational institution. The extension of the study by adding written work of students from other educational institutions as well as a corpus representing the written language of students from other study programs could allow researchers to make broader generalisations and deepen insights into connector usage in the Lithuanian learners' written work. The majority of the LICLE corpus essays were written as exam essays when students were under time and psychological pressure. One more way thus to better understand the reasons for the choice of connectors made by the Lithuanian learners' could be the comparison of their usage in essays written under conditions involving different stress levels such as exam essays and homework assignments.

TEKSTO JUNGIMO PRIEMONIŲ VARTOSENĄ GIMTAKALBIŲ IR  
NEGIMTAKALBIŲ STUDENTŲ RAŠTO DARBUOSE ANGLŲ KALBA.  
GRETINAMOJI KONEKTORIŲ ANALIZĖ

Reziumė

Darbo objektas – teksto konektoriai, vartojami lietuvių, besimokančiųjų anglų kaip svetimosios kalbos (BASK), ir gimtakalbių rašiniuose. Šiame darbe teksto konektoriai suprantami kaip teksto jungimo priemonės, atskleidžiančios rašančiojo požiūrį į tarp atskirų lingvistinių vienetų esančius santykius, pvz., priešpriešos, priežasties ir pan. (Quirk *et al.* 1991: 631–632, Biber *et al.* 1999: 875). Tyrimui pasirinktas sąrašas, sudarytas iš 129 Quirk *et al.* (1991) bei Biber *et al.* (1999) pateiktų teksto konektorių. Pripažįstant, kad egzistuoja konektorių daugiafunkciškumo problema, darbe konektoriai priskiriami semantinėms kategorijoms pagal jų prototipinę reikšmę.

Darbo tikslai ir uždaviniai

Disertacijos **tikslai**:

1. Išanalizuoti teksto konektorių vartojimą lietuvių BASK studentų samprotaujamuosiuose rašiniuose.
2. Palyginti lietuvių BASK studentų konektorių vartojimą su vartojimu tikslinėje, t. y. gimtakalbių anglų kalboje.
3. Išsiaiškinti, kokie yra lietuvių BASK studentų konektorių vartojimo ypatumai kitų kalbinių grupių BASK kontekste, atskleisti galimas tokio vartojimo priežastis.
4. Nustatyti būdus, kurie pagerintų lietuvių BASK studentų konektorių vartojimo rezultatus.

Užsibrėžtiems tikslams siekti buvo keliami šie **uždaviniai**:

1. Sukaupti pažengusiųjų (angl. *advanced*) lygio lietuvių BASK studentų tekstyną.
2. Tirti bendrus ir atskirų semantinių kategorijų kiekybinius konektorių vartojimo parametrus LICLE ir gimtakalbių studentų tekstynuose.
3. Nustatyti konektorių poziciją sakinyje ir vietą tekste.
4. Atlikti kiekybinių rezultatų statistinio reikšmingumo analizę.



5. Palyginti lietuvių BASK konektorių vartojimo rezultatus su kitų kalbinių grupių BASK rezultatais.
6. Nagrinėti priežastis, kodėl skiriasi lietuvių BASK ir gimtakalbių studentų konektorių vartoseną.

#### Darbo naujumas ir aktualumas

Nuoseklūs anglų kaip svetimšios kalbos (ASK) tyrimai prasidėjo dvidešimtojo amžiaus šeštajame–septintajame dešimtmečiuose (Hinkel 2011: 523), tačiau rankinis informacijos apdorojimo būdas riboja tiriamosios medžiagos apimtį. Situaciją pakeitė tarptautinis besimokančiųjų anglų kalbos tekstyno (*International Corpus of Learner English* – ICLE) projektas – atsirado geresnės galimybės tirti BASK rašytinį diskursą bei mėginti nustatyti bendrus bei atskiroms kalbinėms BASK grupėms būdingus bruožus. Šiame darbe naudotas neseniai baigtas kaupti lietuvių BASK kalbą reprezentuojantis lietuviškasis ICLE potekstynis (LICLE) – naujovė Lietuvoje, kuri leis užpildyti lietuvių BASK rašytinės kalbos tyrimų spragą.

Ši disertacija – viena pirmųjų teksto konektorių, kaip vienos iš teksto koheziją užtikrinančių priemonių, vartojimo lietuvių BASK studentų rašytiniame diskurse analizė. Tikėtina, kad joje pateiktos išvalgos galės būti atskaitos tašku atliekant tolesnius tyrimus: nustatant, kaip vieno ar kito teksto konektoriaus dažnesnis ar retesnis vartojimas lietuvių BASK rašiniuose, palyginus su gimtakalbių anglų darbais, padeda kurti nuoseklesnį tekstą ar tirti koreliaciją tarp tam tikrų teksto konektorių vartojimo ir besimokančiųjų kalbos lygio.

#### Darbo vertė ir pritaikymas

Šiame darbe tirti teksto konektoriai, eksplicitiškai žymintys semantinius santykius, nėra vienintelis teksto kohezijos ir koherencijos užtikrinimo būdas, tačiau jeigu jų nepakanka, skaitytojui reikia įdėti daugiau kognityvinių pastangų, kad suprastų tekstą, be to, didėja tikimybė, jog skaitant bus iškreipta teksto prasmė (Olshtain ir Cohen 2005: 316–317). Taip pat problemiškas pergausus konektorių vartojimas, kuris gali pakenkti teksto stiliui, jį apsunkinti (Remeikytė 2002: 149). Tyrimas, be kita ko, nustatė, kad lietuvių BASK kalba skiriasi nuo kitų kalbinių grupių BASK kalbos, todėl mokymo priemonės, tinkančios kitų šalių besimokantiejiems, gali netikti lietuvių BASK studentams. Dėl šios priežasties

tikėtina, kad šio darbo rezultatai, išryškindami opiausias lietuvių BASK konektorių vartojimo problemas, bus vertingi kuriant būtent į šią tikslinę grupę orientuotą mokomąją medžiagą.

### Ginamieji teiginiai

1. Kiekybinis tyrimas atskleidžia, kad lietuviai ASK studentai turi polinkį dažniau nei gimtakalbiai semantinius teksto jungimo santykius žymėti eksplicitinėmis priemonėmis, t. y. konektoriais.
2. Skirtingų kitų kalbinių grupių svetimkalbiai, taip pat kaip gimtakalbiai, dažniausiai eksplicitiškai žymi tų pačių semantinių kategorijų – išskaičiavimo, priešpriešos, priešasties ir apozicijos ryšius.
3. Lietuvių BASK studentų konektorių vartosenos nuokrypis nuo gimtakalbių studentų vartosenos kai kuriais atvejais žymi pozityvias tendencijas – konektorių vartojimas LICLE tekstyne panašesnis į jų vartoseną gimtakalbių akademinio lygmens tekstynuose nei LOCNESS.
4. Lietuvių BASK studentų konektorių vartoseną skiriasi nuo kitų kalbinių grupių vartosenos, todėl kvestionuotinas į visas grupes orientuotos mokomosios medžiagos efektyvumas.
5. Aiškiau ir išsamiau pristatius konektorius mokomojoje literatūroje, galėtų sumažėti teksto konektorių vartojimo problemos.

### Tyrimo metodai ir medžiaga

Disertacijoje aprašomas tyrimas atliktas naudojantis gretinamojoje kalbotyroje ir tekstynų lingvistikoje taikomus metodus: gretinamąją tarpukalbės analizę bei statistinę duomenų analizę. Tyrime pirmenybė teikiama lietuvių BASK ir gimtakalbių studentų konektorių vartojimo palyginimui, tačiau pasitelkiant kitas kalbines BASK grupes tyrusių lingvistų duomenis, konektorių vartoseną lyginama ir K2 vs. K2 aspektu.

Tyrimui buvo pasirinktas suderintas automatizuotos, t. y. atliekamos kompiuterinės programos, ir rankinės analizės metodas: iš *TextSTAT 2.5*. (Hüning 2000/2007) ir *AntConc 3.2.1w* (Anthony 2007) programine įranga generuotų tiriamų žodžių konkordansų rankiniu būdu buvo pašalinti konektoriais nepriklausantys žodžiai, kurie patenka atliekant

automatizuotą analizę. Galutinis konektorių sąrašas toliau buvo peržiūrimas pagal jų vartojimo vietą sakinyje. Detalesniam tyrimui pasirinktų konektorių atveju apžvelgiama ir vieta rašinyje, t. y. ar jie vartojami išanginėje, dėstymo ar apibendrinamojoje dalyje.

Šiame darbe statistinio reikšmingumo rodiklis apskaičiuojamas naudojant Rayson (2004) sukurtą logaritminės tikimybės (LL) skaičiuoklę. Dėl skirtingo tirtų tekstynų dydžio darbe taip pat buvo naudotasi dažnių normalizavimu.

Atliktam tyrimui duomenys buvo renkami iš LICLE. Tyrimo metu LICLE imtį sudarė 2006–2009 m. 216 studentų 268 rašiniai, 154 992 žodžiai, parašyti samprotaujamosiomis (91,27 procento – 141 457 žodžiai) ir literatūrinėmis temomis (8,73 procento – 13 535 žodžiai).

Lietuvių BASK studentų konektorių vartosenos ypatumų priežastims nustatyti atliekant atskirų konektorių analizę buvo naudojamas Juknevičienės sukauptas pirmo kurso VU anglų filologijos studentų rašinių tekstynas. Jį sudaro 293 rašiniai, 118 552 žodžiai.

Gimtakalbių studentų duomenys paimti iš trijų LOCNESS tekstyno potekstynių: britiškojo LOCNESS-BR (90 rašinių, t. y. 95 695 žodžiai), amerikietiškojo LOCNESS-US (232 rašiniai, t. y. 168 231 žodis) ir britų mokinių potekstynio A-level (30 rašinių, t. y. 15 434 žodžiai).

Tyrimo taip pat naudojami gimtakalbių duomenys iš britų nacionalinio tekstyno BYU-BNC (96 263 399 žodžiai), šiuolaikinės amerikiečių anglų kalbos tekstyno COCA (96 263 399 žodžiai) bei jų akademinį potekstynių (atitinkamai 15 331 668 ir 85 791 918 žodžių) ir britų akademinės rašytinės anglų kalbos tekstyno BAWE (6 696 026 žodžiai) tekstynų.

Disertacijoje taip pat naudojamosi tekstynų, reprezentuojančių lietuvių kalbą ir jos akademinį registrą, duomenimis: dabartinės lietuvių kalbos tekstyno (DLKT) negrožinės literatūros potekstynio (21 024 249 žodžiai) ir lietuvių mokslo kalbos tekstyno CorALit (apie 9 000 000 žodžių).

#### Disertacijos struktūra.

Disertaciją sudaro įvadas, teorinę apžvalgą bei medžiagą ir metodus pristatantys skyriai, tyrimų analizė, išvados, literatūros sąrašas ir priedai. Įvade aptariamas darbo

naujumas ir aktualumas, pristatomas tyrimo objektas, tikslai ir uždaviniai, aptariama taikomoji vertė ir pateikiami ginamieji teiginiai. Pirmoji dalis skirta darbe naudotų terminų problemai ir tarpukalbės bei konektorių tyrimų apžvalgai. Antroje dalyje pristatoma tyrimui naudota medžiaga ir metodai. Trečioji disertacijos dalis skirta tyrimo rezultatų aptarimui: pirmiausiai pateikiama kiekybinė konektorių apžvalga, toliau pateikiama atskirų semantinių konektorių tipų vartosenos analizė: supažindinama su bendrais tiriamo semantinio konektorių tipo rezultatais, analizuojamas (-i) vienas ar keli tiriamam tipui priklausantis (-ys) konektorius (-iai) bei pateikiami apibendrinimai apie to tipo konektorių vartoseną lietuvių BASK ir gimtakalbių studentų rašiniuose. Disertacija baigiama išvadomis, pirminių šaltinių ir cituotos literatūros sąrašu bei penkiais priedais.

### Išvados.

Disertacijoje buvo tiriama (teksto) konektorių vartoseną svetimkalbių (lietuvių) BASK studentų ir gimtakalbių anglų rašytinės kalbos tekstynuose. Taip pat buvo atliktas palyginimas su kitų K1 BASK grupių darbais. Susisteminti tyrimo rezultatai leidžia daryti išvadą apie lietuvių BASK kalbai būdingus eksplisitinio minčių žymėjimo konektoriais ypatumus, santykį su kitų K1 kalbinių grupių konektorių vartosenos bruožais, universalaus samprotavimo modelio egzistavimą, priežastis ir būdus, galėjusius lemti rezultatus, kurie gali priartinti svetimkalbių anglų kalbą prie tikslinės kalbos.

Kiekybinė konektorių analizė parodė, kad LICLE tekstyne nustatytas reikšmingai dažnesnis semantinių santykių eksplisitinis žymėjimas konektoriais nei gimtakalbių britų bei amerikiečių LOCNESS potekstyniuose, atitinkamai LL +34,36 ir LL +423,97. Tyrimo rezultatai atskleidė, kad lietuvių BASK studentai ne tik dažniau nei gimtakalbiai mintis jungia konektoriais, bet tai daro skirtingose sakinio pozicijose: LICLE tekstyne pirmenybinė yra SP (61 proc.), o LOCNESS-BR ir LOCNESS-US potekstyniuose – SV pozicija (atitinkamai 71 proc. ir 59 proc.). Nors disertacijoje tirta konektorių pozicija tekste lietuvių BASK studentams kėlė mažiau problemų nei pirmenybinės pozicijos sakinyje pasirinkimas, paaiškėjo, jog kai kurių semantinių kategorijų atveju lietuvių BASK studentai konektoriais žymėjo semantinius santykius gimtakalbiams nebūdingose teksto dalyse (pvz., *for instance* apibendrinamojoje pastraipoje). Taigi duomenys rodo, kad reikėtų ne tik skatinti lietuvių

BASK studentus semantinius santykius rečiau žymėti konektoriais, bet ir skirti pakankamai dėmesio aptarti gimtakalbių kalbai būdingas konektorių pozicijas.

Atliekant tyrimą išryškėjo skirtumai tarp gimtakalbių britų ir amerikiečių LOCNESS potekstynių: britų studentai konektoriais semantinius ryšius žymi daugiau kaip pusantro karto dažniau nei tai daro amerikiečių studentai. BAWE tekstyno, turinčio atspindėti britų akademinę rašytinę kalbą, duomenys taip pat skiriasi nuo LOCNESS-BR potekstynio. Šių tekstynų duomenų palyginimas su BYU-BNC akademinio potekstynio bei LOCNESS-US su COCA akademinio potekstynio duomenimis neleidžia vienareikšmiškai vertinti lietuvių BASK studentų konektorių vartosenos: LICLE tekстыne nustatyti nukrypimai nuo gimtakalbių studentų rezultatų kai kuriais atvejais (pvz., išskaičiavimo kategorijos) žymi pozityvias tendencijas, kai lietuvių BASK studentų konektorių kiekybinis vartojimas yra panašesnis į gimtakalbių nacionalinių tekstynų akademinės dalies rezultatus.

LICLE tekstyno analizė atskleidė, kad eksplicitinis žymėjimas konektoriais lietuvių BASK studentų darbuose užima tarpinę poziciją tarp britų ir amerikiečių studentų darbų. Nepaisant to, skirtingų semantinių kategorijų konektorių kiekybinis tyrimas, konektorių įvairovės ir dažniausiai vartojamų konektorių analizės rezultatai leidžia kalbėti apie didesnius LICLE ir LOCNESS-BR vartojimo panašumus. Tokius rezultatus galima interpretuoti kaip žyminčius lietuvių BASK studentų kalbinių ieškojimų bei raidos tarpsnį, kuriam yra būdinga skirtingų kalbos bruožų sankirta.

Palyginus lietuvių ir kitų kalbinių grupių BASK konektorių vartojimo rezultatus, reikia pabrėžti, kad nėra visuotinai apibendrintino bendro kiekybinio konektorių vartojimo modelio BASK kalboje – skirtingų kalbinių grupių duomenys smarkiai varijuoja ir gali būti priskiriami tų kalbinių grupių BASK savitumui. Lietuvių BASK kalbos savitumui taip pat priskirtini kai kurių konkrečių konektorių vartojimo dažniai. Nors yra nustatyta grupė konektorių, kuriuos skirtingų kalbinių grupių BASK su tam tikromis išlygomis vartoja dažniau (pvz., *moreover, on the contrary*) ar rečiau (pvz., *hence, therefore*) nei gimtakalbiai, LICLE duomenys pateikia ir priešingų rezultatų, kai kitų kalbinių grupių BASK darbuose itin dažnai sutinkami konektoriai LICLE tekстыne vartojami rečiau nei gimtakalbių (pvz.,

*still*) ar, priešingai, BASK kalboje rečiau sutinkami konektoriai lietuvių BASK studentų rašiniuose yra vartojami dažniau nei gimtakalbių studentų darbuose (pvz., *however*).

Nepaisant kiekvienos kalbinės grupės BASK kalbos savitumo, buvo nustatyta keletas bendrų vartojimo ypatumų: BASK yra linkę dažniau nei gimtakalbiai žymėti lengviausiai suprantamus semantinius ryšius, tokius kaip išskaičiavimo ar konektorius vartoti gimtakalbiams nebūdingoje sakinio pradžios pozicijoje.

Nors skirtingų kalbinių grupių BASK ir gimtakalbiai studentai konektorius vartoja nevienodai, tyrimo metu išanalizuoti jų vartojimo rezultatai leidžia teigti apie egzistuojanti beveik visuotinai priimtina samprotavimo modelį. Ir svetimkalbiai, ir gimtakalbiai studentai samprotavimą dažniausiai grindžia argumentų išvardijimu ir kontrargumentų bei priešasčių nurodymu. Būtent šiuos (išskaičiavimo, priešpriešos ir priešasties) santykius eksplicitiškai žyminčių kategorijų konektoriai dažniausiai vartojami ne tik LICLE, LOCNESS-BR bei LOCNESS-US, bet ir kituose BASK tekstynuose, pvz., vengrų ASK, o dvi paskutinės ir SWICLE tekстыne. Ši tendencija laikytina beveik universalia, nepriklausoma nuo konektorių dažnio atskiruose tekstynuose. Taip pat nustatyta visuotinė tendencija samprotaujamosiuose rašiniuose pasikliauti tam tikra nedidele konektorių grupe – dažniausiai vartojamų konektorių dešimtukas sudaro apie tris ketvirtadalius visų konektorių vartojimo atvejų.

Siekiant atskleisti, kas galėjo vaidinti svarbų vaidmenį, lėmusį disertacijoje nustatytą lietuvių ASK studentų konektorių vartojimą, netiesiogiai pasitvirtino kelios hipotezės: K1 kalbos perkėlimo, BASK kalbos lygio bei mokomosios medžiagos ir mokymo proceso įtakos. Atliktas tyrimas leidžia manyti, kad konektorių vartojimas lietuvių BASK kalboje tik minimaliai yra veikiamas BASK gimtosios kalbos. Pastebėta, kad tais atvejais, kai K1 kalboje esama aiškaus angliškojo konektoriaus atitikmens, yra didesnė tikimybė, kad BASK iš kelių alternatyvų pasirinks būtent pastarąjį (pvz., *for example*).

Nemažas vaidmuo tenka BASK kalbos lygiui. Viena vertus, sąlyginai lengvesnių kategorijų atveju (apibendrinimo kategorijos ar išskaičiavimo kategorijos konektoriaus *moreover*) pasitvirtino atvirkštinės priklausomybės hipotezė: kai kylant kalbos lygiui ir rašančiojo brandai, mažėja konektoriaus dažnis. Sąlyginai sudėtingesnės, priešpriešos,

kategorijos konektorius *however* – priešingai, rečiau vartojamas būtent VU anglų filologijos pirmakursių studentų darbuose. Kadangi tokia pati vartojimo tendencija nustatyta ir gimtakalbių britų studentų ir mokinių tekstynuose, galima teigti, jog tai sietina su tobulėjant argumentavimo strategijai naudojamu sudėtingesniu argumentavimu ir didesniu pateikiamų argumentų skaičiumi. Analogiškos vartosenos tendencijos matomos ir tiriant „sudėtingesnėms“ kategorijoms priskirtinų konektorių santykį (pvz., *thus* ir *therefore*). Gauti rezultatai leidžia kalbėti apie lietuvių BASK konektorių vartojimo ypatumus kaip apie kalbinės raidos etapo atspindį, o LICLE tekstyne nustatyta konektorių vartojimo raida traktuotina kaip teigiamas VU universitete vykdomo akademinio rašymo dėstymo kurso įvertinimas.

Tyrimas atskleidė tebeegzistuojančią žodynuose, žinynuose ir gramatikose pateikiamos informacijos apie konektorius problemą: ji nėra išsami, o kartais net klaidinanti. Todėl mokymo proceso metu itin svarbu studentams pateikti aiškesnes konektorių vartojimo gaires, išryškinant ne tik registro skirtumus (pvz., *moreover* vs. *furthermore*), bet ir nurodyti gimtakalbiams labiausiai ar mažiausiai įprastus kontekstus. Pavyzdžiui, LICLE tektyne pirmajame pastraipos sakinyje net penktadalyje atvejų vartojamas konektorius *moreover* prieštarauja gimtakalbių studentų tekstyno duomenims, kur šioje pozicijoje *moreover* nėra pavartotas nei karto.

Pažengusiųjų lygio BASK kalboje nustatomi sunkumai paprastai yra gana subtilūs ir sunkiau pastebimi (Nesselhauf 2004: 145), todėl tikėtina, kad šis darbas bus naudingas pritaikant akademinio rašymo kursą būtent lietuvių BASK studentams. Jis taip pat padės nuspręsti, į kuriuos konektorius, į kurių semantinių santykių žymėjimą reikėtų atkreipti ypatingą dėmesį, kad būtų išvengta dažniausių lietuvių BASK rašiniuose pasitaikančių konektorių vartojimo problemų, pvz., dažnesnio nei tikslinėje kalboje (išskaičiavimo, apozicijos ar apibendrinimo) ar netinkamo (pvz., *moreover*) konektorių vartojimo, taip pat sužinoti, kurių konektorių vartoseną yra artima gimtakalbių studentų (pvz., *thus* pozicija pastraipoje) vartosenai.

#### LIST OF PUBLICATIONS ON THE SUBJECT OF THE DISSERTATION

1. Bikelienė L. 2008. Connector Usage in Advanced Lithuanian Learners' English Writing. Lewandowska-Tomaszczyk, B. (ed.). *Corpus linguistics, computer tools, and applications - state of art*. Lodz Studies in Language Vol. 17. Frankfurt am Main : Peter Lang. 741–755.
2. Bikelienė L. 2008. Resultive Connectors in Advanced Lithuanian Learners' English Writing. *Kalbotyra* 59(3): 30–37.
3. Grigaliūnienė J., Bikelienė L., Juknevičienė R. 2008. The Lithuanian Component of the International Corpus of Learner English (LICLE): a resource for English language learning, teaching and research at Lithuanian institutions of higher education. *Žmogus ir žodis* 10(3): 62–66.
4. Bikelienė L. 2009. Priešpriešos konektorių vartojimas besimokančių anglų kalbos ir anglakalbių studentų rašto darbuose. *Kalbotyra* 61(3): 21–35.
5. Bikelienė L. 2009. Insights from the Lithuanian Learner Corpus of English: pilot study on the use of resultive connectors. Priimtas spausdinti: *Language Forum: An International Journal of Language and Linguistics* 35(2): 113–126.
6. Bikelienė L. 2010. Apibendrinimo žymėjimas konektoriais besimokančių anglų kalbos ir anglakalbių studentų rašto darbuose. *Žmogus kalbos erdvėje: mokslinių straipsnių rinkinys* 6: 546–552.

#### CONFERENCE PRESENTATIONS ON THE SUBJECT OF THE DISSERTATION

1. *Connector Usage in Advanced Lithuanian Learners' English Writing*, PALC conference, University of Lodz (Poland), 19-22 April, 2007.
2. *Connector Usage in Advanced Lithuanian Learners' English Writing*, International Scientific Conference “Tekstas: lingvistika ir poetika 15”, Šiauliai University (Lithuania), 15 November, 2007.
3. *Resultive Connectors in Advanced Lithuanian Learners' English Writing*, International Conference on English and German Studies, Vilnius University (Lithuania), 18–20 September, 2008.
4. *Insights from the Lithuanian Learner Corpus of English* (a co-authored poster presentation with Rita Juknevičienė), International Seminar on Corpus Linguistics in honour of John Sinclair, Granada (Spain), 22–24 September, 2008.
5. *Kontrasto konektorių vartojimas besimokančiųjų ir anglakalbių darbuose*, International conference on language teaching, Institute of Foreign languages, Vilnius University (Lithuania), 8–9 October, 2009.
6. *Apibendrinimo žymėjimas konektoriais besimokančių anglų kalbos ir anglakalbių studentų rašto darbuose*, International Scientific Conference „Žmogus kalbos erdvėje“, Vilnius University Kaunas Faculty of Humanities (Lithuania), 14-15 May, 2010
7. *The use of listing connectors in advanced Lithuanian learners' EFL writing*, The Fifth Inter-Varietal Applied Corpus Studies (IVACS) group International Conference, University of Edinburgh, UK, 18-19 June, 2010



8. *The use of connectors in advanced Lithuanian learners' English writing*, 43rd Annual Meeting of the Societas Linguistica Europaea, Vilnius University (Lithuania), 2-5 September 2010

Lina Bikelienė (g. 1975) 1999 m. baigė Anglų kalbotyros studijas Vilniaus universiteto Filologijos fakultete. Nuo 1999 m. dirba Vilniaus universitete. 2006 m. įstojo į Vilniaus universiteto Filologijos mokslo krypties doktorantūrą. Doktorantūros studijų metais dalyvavo tarptautinėse mokslinėse konferencijose, vykdė BASK kalbos tekstyno rengimo projektą (ICLE), paskelbė šešis mokslinius straipsnius. 2009 m. gavusi VU doktorantų mobilumo fondo paramą buvo išvykusi trumpalaikiai stažuotei į Liuveno (Louvain-la-Neuve) katalikiškojo universiteto (Belgija) Anglų tekstynų lingvistikos centrą.

Lina Bikelienė (b. 1975) in 1999 graduated from Vilnius University and received an MA degree in English Linguistics. Since 1999 she has been working at Vilnius University. In 2006 she started her PhD studies in Vilnius University. During the studies she took part in a number of international conferences in Lithuania and abroad, published six research articles, was engaged in the compilation of the Lithuanian component of the ICLE corpus. In 2009 she was on a short library visit in the Catholic University of Louvain-la-Neuve in Belgium.