

The present paper deals with the analysis of spoken and written English produced by non-native speakers, namely, Lithuanian EFL (English as a foreign language) learners. The aims of the article are twofold. First, it introduces two corpora of English that represent the spoken (corpus LINDSEI-LITH) and the written language (corpus LICLE) of Lithuanian learners, both corpora being the first electronic databases of Lithuanian learners of English. Second, the article demonstrates how a contrastive study of spoken and written learner language may contribute to a better understanding of mother tongue influence in the process of foreign language learning. Specifically, the paper will present findings of a case study of learner language, focused on the analysis of participial *-ing* clauses. The linguistic findings are seen as a starting point for a further discussion of linguistic and pedagogical implications and applications of learner corpus research.

INTRODUCTION: SPOKEN AND WRITTEN LEARNER CORPORA AND EFL RESEARCH

Learner corpora are becoming a significant asset for second language acquisition (SLA), since most researchers agree that learner corpora reveal the learners' needs and can inform language teaching in ways in which native speaker corpora cannot (Granger 2002: 21). Learner corpora, which are a relatively recent phenomenon and the compilation of which started in the 1990s with the corpora of written learner language, are increasingly becoming more varied and sophisticated (Barlow 2005, Myles 2005, 2012, Pravec 2002). One of the first corpora compiled for research purposes was the International Corpus of Learner language (ICLE) (Granger 2003: 63). The ICLE project, aimed at the collection EFL learner essays that would represent university student writing with different mother tongue backgrounds, was launched in Belgium. Its second version, which came out in 2009 (Granger et al. 2009), contains

3.7 million words of written learner English representing 16 mother tongue backgrounds. Compilation of the Lithuanian component of ICLE (LICLE), which will also be included in the ICLE database, was completed in 2011. It contains advanced student essays from Vilnius University and Kaunas Vytautas Magnus University. The size of LICLE is 240,332 words (420 essays): 175,990 words (305 essays from Vilnius University) and 65,342 (115 essays) from Kaunas Vytautas Magnus University. The corpus consists of argumentative essays on a variety of topics (Grigaliūnienė et al. 2008). The pilot versions of LICLE provided material for the first corpus-driven and corpus-based analysis of written Lithuanian learner language (see below).

The success of the ICLE project prompted the launch of its spoken counterpart the Louvain International Database of Spoken English Interlanguage (LINDSEI) (Gilquin et al. 2010). The LINDSEI project was launched in 1995 by the members of the ICLE team at the Centre of English Corpus Linguistics (CECL), Université Catholique de Louvain (Belgium). A number of other LINDSEI components have been and are currently being compiled (Bulgarian, Chinese, Dutch, French, German, Greek, Italian, Japanese, Polish, Spanish, Swedish, Arabic, Basque, Brazilian Portuguese, Norwegian, Turkish). Compilation of the Lithuanian component, which will be referred to as LINDSEI-LITH, started at the Department of English Philology of Vilnius University in 2011 (Grigaliūnienė & Juknevičienė 2011). Recording of interviews took place in 2011-2012. At present, the corpus material is being processed: the transcriptions are being double-checked and revised according to the LINDSEI transcription guidelines. The final version of the corpus will consist of 82 interviews (ca. 16 hours) between a Lithuanian learner and a native speaker interlocutor¹. The non-native interviewees are all advanced learners of English, labelled 'advanced' on the basis of an external criterion – they are all third-year students of English, native speakers of Lithuanian, studying English at the Department of English Philology, University of Vilnius. The informal interviews, which last approximately 15 minutes each, were recorded with the consent of the students² (the students completed learner profile forms, giving information about their age, nationality, native language, father's and mother's mother tongue, languages spoken at home, education, years of English at school and university, medium of instruction, stay in an English-speaking country, and other languages, and signed the forms by giving permission to use the interview for research purposes). The data for LINDSEI-LITH are collected using a specific LINDSEI format: the interviews are of approximately the same length – 2000 words each and follow the same pattern:

¹ We express our gratitude to Dr Francis Whyte and lecturer Mark Fearon, who very kindly agreed to act as interviewers in the project.

² We are also grateful to the third-year students (of the year 2011 and 2012) of English Philology at the Department of English Philology who have agreed to participate in the project and gave their permission to use the interview data for research.

they start with an informal discussion of university life, hobbies, travel or future plans. Then the interviewees are asked to choose one of three topics (an experience that taught them an important lesson; a country which impressed them; and a film or a play which they particularly liked or disliked). The students are asked not to make any notes and speak without preparation – this is done for the sake of spontaneity. Each interview ends with a short picture-based story telling. The interviews were transcribed using an orthographic transcription scheme.³

Compilation of ICLE and LINDSEI has given rise to a plethora of studies in the field. Researchers are particularly interested in such features of learner English that make it sound unnatural and different from native speaker English. An exhaustive bibliography of research is currently available on the website of the CECL (<www.uclouvain.be/en-cecl-lcbiblio.html>). While certain features of non-nativeness are common to different learner groups irrespective of their mother tongue, for example, overuse of connectives in writing (Bikeliënè 2012: 59) and misuse of lexical phrases (Waibel 2007), some of the deviations from standard use can be explained by mother tongue influence and are L1-specific. So contrastive studies of learner language of the basis of the ICLE and LINDSEI corpora enable to establish specific problem areas for different learner groups and address them in greater detail.

More recently, learner corpora are increasingly used as empirical databases for the specification of proficiency levels according to the Common European Framework of Reference (CEFR, Council of Europe 2001), which has become a broadly recognized reference tool in many educational contexts across Europe. As regards the English language and ELT/EFL educational contexts, the English Profile research programme at the University of Cambridge is undoubtedly one of the largest learner language projects undertaken in this field. Ongoing research and publication of findings of the project team has become an influential reference base in Europe as it gives valuable data and information about characteristic features of EFL learners at different levels of proficiency, namely, A1-C2 (Hawkins & Butterby 2009, Thewissen et al. 2006, Thewissen *forthcoming*). Undoubtedly, mother tongue background should also be taken into account when creating a learner profile at a particular level. It is in this respect that LICLE and LINDSEI-LITH corpora have a considerable potential. Obviously, there is no other reliable way to investigate achievement and proficiency of Lithuanian learners' competence in English except for learner corpus research.

The study of the acquisition of English as a foreign language has been unduly neglected in Lithuania for decades. The earliest publications in the field were written by Aprijaskytė (1975) and Aprijaskytė and Pareigyte (1982) and dealt with

³ As the corpus compilation project was part of the students' professional practice course, they were asked to transcribe their own speech first, which the students said had been an "eye-opening, sobering, very useful and meaningful activity", although very hard and time-consuming.

lexical, or collocational, errors of the Lithuanian EFL learners. The authors pointed out specific problem areas that should be addressed more carefully in the teaching of English vocabulary to the Lithuanian learners and compiled a set of exercises for remedial purposes. It took, however, another two decades before EFL theory and practice gained more attention in Lithuania. Manuals of writing and EFL research publications were focused on different aspects of teaching academic written English (Katkuvienė 2003, Katkuvienė and Šeškauskienė 2005, Šeškauskienė 2008, Burkštaitienė, 2006, Stanevičienė, 2007). It was only with the compilation of the LICLE that the authentic learner language became an object of research. Several studies reported evidence from the analysis of connectors (Bikeliienė 2008, 2009, 2012) and lexical expressions (Juknevičienė 2008, 2011) in the Lithuanian learner writing.

The focus of learner language research until quite recently has been on written rather than spoken language. The compilation of the LINDSEI-LITH has opened the way for a contrastive approach to learner writing and speech. The first attempts were made to look at the phraseology of the Lithuanian learner speech (Grigaliūnienė and Juknevičienė 2011). The study gives an overall picture of recurrent word phrases in the Lithuanian learner speech and provides empirical evidence of the learners' lexical competence. The findings of the study suggest that the advanced Lithuanian learners of English use many formulaic sequences even though they are seldom idiomatic. While formulaic sequences are broadly understood as set phrases with a degree of idiomaticity (Wray 2002, Schmitt 2004), the majority of the items in the LINDSEI-LITH corpus are semantically transparent and their formulaic nature largely depends on the external, or contextual, factors. They thus become formulaic due their pragmatic functions, and their semantic opaqueness, or idiomaticity, seldom comes into play. As to the pragmatic functions, the predominant use of the formulaic sequences could be explained by the learners' attempts to manage their speech and gain time for processing and planning further discourse. Uncertainty and hesitation in the learner speech are also indicated by clusters of such formulaic sequences as *I think, you know, I don't know*. Another study (Grigaliūnienė and Juknevičienė *forthcoming*) presents a quantitative and qualitative analysis of the recurrent formulaic sequences in spoken and written English of the Lithuanian EFL learners. The findings of this study show that although the speech of the Lithuanian EFL learners is more formulaic than the written language, there is a considerable overlap between spoken and written language in terms of formulaicity. The learners have built a core set of formulaic sequences which recur both in speech and in writing. The most frequent formulaic sequences in writing are expressions of discourse organization while high-frequency formulaic sequences in spoken language, which often appear in clusters of several formulaic sequences, usually indicating the speaker's hesitation and uncertainty. Although it would be impossible to make sweeping generalizations at this stage of research, the

data collected and analysed provide us with empirical evidence of the learners' lexical competence. Hopefully, a study reported in the following part of the article will shed more light on factors influencing the acquisition of English as a foreign language by Lithuanian learners.

PARTICIPLE *-ING* CLAUSES IN LITHUANIAN LEARNER ENGLISH

This part of the article will present findings from another study where learner speech and writing were contrasted. The study deals with participle clauses, more specifically, *-ing* clauses which, as our experience shows, appear to be problematic to Lithuanian learners. Contrasting data from LICLE and LINDSEI-LITH might provide useful information on the use of *-ing* clauses by Lithuanian learners.

Participle *-ing* clauses in English have nominal, adverbial, and adnominal (post-modifying) uses (cf. Quirk et al. 1985, Biber et al. 1999). In nominal uses, the clauses with *-ing* participles can function as subjects, subject complements, objects, and prepositional complements, whereas in adverbial and adnominal (or adjectival) uses they occur as adjuncts, disjuncts, complements, and postmodifiers. The following examples from the British National Corpus illustrate some of the instances (*-ing* clauses in bold):

Nominal uses:

*I started **thinking about Christmas** then, didn't I?* (object)

***Thinking about Christmas** was almost as much fun as celebrating it.* (subject)

Adverbial uses:

*(...) there are still just as many options open to them **when looking for a loan**.* (adjunct)

***Frankly speaking**, I don't want an architecture centre in our Institute headquarters in London.* (disjunct)

Adnominal (postmodifying) uses:

*In fact, in the mature Italian dried pasta market it is the only sector **showing a substantial increase**.* (postmodifier)

*My dad took me to watch Newcastle when I was about 10 but I couldn't see because of all the people **standing in front of me**.* (postmodifier)

The nominal uses are traditionally interpreted as instances of the English gerund which is comparable in terms of meaning and function to the noun, even if this distinction, as suggested by De Smet (2010), is not necessarily clear-cut. In his study, the linguist provides contradictory evidence which makes the distinction into nominal

vs. adverbial and adjective uses rather ambiguous because language users' behaviour and linguistic choices do not demonstrate a consistent pattern. Yet for the purposes of teaching English as a foreign language, the main difference between the nominal uses on the one hand and adverbial and adnominal uses on the other is primarily related to the existing alternative ways of expressing meaning conveyed by *-ing* clauses. As it appears, the nominal uses are indeed distinctive in that they have no other systematic finite alternatives whereas *-ing* clauses used adverbially or adnominally might be rephrased by finite clauses. As a consequence, these two types of *-ing* clauses always involve a choice for the learner while their use in learner language might reveal specific problem areas in the process of language acquisition. Hence our analysis will be focused on adverbial and adnominal (postmodifying) uses of *-ing* clauses.

Several earlier studies have revealed that participle clauses cause a considerable difficulty to non-native users of English. In comparison to native speakers, foreign learners tend to underuse and misuse them (Granger 1997, Springer 2012), which may be caused by mother tongue influence (Cosme 2008). In general, participle clauses are a characteristic feature of written language, and, as a consequence, they are usually mastered at more advanced levels of proficiency, namely, B2 and higher (Hawkins & Filipovic 2012). Furthermore, participle clauses acquire a number of functions, which are sometimes ambiguous to interpret, which, arguably, makes the understanding and use of such clauses problematic for non-native speakers (Koskinen et al. 1987).

Among Indo-European languages, Lithuanian is unique for having 13 different forms: Participles, Half-participles, Gerunds (Ambrazas 1997, Klimas 1987). The data from the morphologically annotated corpus of Lithuanian shows that 49.3% of verb forms are participle forms (Rimkutė 2006). Participles are frequent both in speech and writing: they can be used in attributive (pre-/postmodifying clauses) as well as adverbial clauses (Ambrazas 1997). Therefore, it would be interesting to find out whether the misuse of participle clauses in the language of Lithuanian learners is influenced by the mother tongue or is teaching-induced or is caused by some developmental problems.

DATA AND METHODS

The data used for the research came from the LICLE corpus for written English and the LINDSEI-LITH corpus for spoken English. Both corpora were searched for *-ing* clauses, then the WordSmith Tools (version 5, Scott 2008) program was used for concordancing. Since the two corpora contain no part-of-speech tags, extracting participles was semi-manual. The procedure was simple but quite reliable. First, with the help of the “replace” function in “MS Word” every instance of “ing” in our corpora

was replaced with the tag “ing PARTICIPLE”. Then a concordance of the tag was generated in “WordSmith Tools” (version 5). Finally, a careful manual revision of the concordance lines was undertaken to eliminate irrelevant lines, for example:

... during PARTICIPLE the lectures ...

... I was asking PARTICIPLE about ...

according PARTICIPLE to ...

... is looking PARTICIPLE at ...

where the tag appeared after such instances of “ing” that were not instances of the participle. After primary revision, concordance lines were analysed again to establish uses of *-ing* clauses.

As expected, the quantitative data revealed an overuse of both adverbial and adnominal (postmodifying) clauses in written learner English in relation to speech (see Table 1), with both types being more than four times more frequent in writing than in speech, which is evident from the normalised frequencies of the clauses. As indicated in the literature, participle clauses are more characteristic of written rather than spoken register as a condensed packaging of information (Biber et al. 1999:754, Downing & Locke 2006: 452, Granger 1997: 11), which, consequently, is good news for our learners – their written language has a feature typical of authentic written English. Frequencies of adverbial and adnominal *-ing* clauses in argumentative essay writing of native speakers, however, is higher than in our data. Granger (1997) reports 680 instances (normalized per 100,000 words) in the LOCNESS corpus, which represents native speaker English. So no matter how frequently the clauses are used by Lithuanian learners, their written language contains fewer *-ing* clauses than native speakers’ use.

Table 1. Quantitative data (normalised frequencies per 100,000 words)

	LICLE (189,756 words) ¹		LINDSEI-LITH (85,060 words)	
	Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency
Adverbial <i>-ing</i> clauses	851	447	101	119
Adnominal <i>-ing</i> clauses	279	147	29	33
Total	1130	595	130	149

¹ As the issue of the influence of the mother tongue was addressed in our research, for this particular study we chose the data from the learners who indicated Lithuanian as their mother tongue in the questionnaires.

Contrasting our learners with other learner groups, interestingly, reveals an overuse of *-ing* clauses in writing. For example, in her analysis of ICLE data which involved essays written by French, Swedish and Dutch students, Granger (1997) found the

following frequencies of *-ing* clauses (the frequencies are normalized per 100,000 words to allow comparison with our data): 288 in the French subcorpus, 335 in the Swedish subcorpus and 330 in the Dutch subcorpus. Compared with 595 instances in LICLE, the numbers are significantly smaller. Undoubtedly, a more thorough quantitative analysis should be undertaken to justify statistical overuse of *-ing* clauses in Lithuanian learner language, particularly for speech, yet even preliminary statistics seems to indicate a characteristic feature of Lithuanian-English interlanguage. A more detailed qualitative analysis highlighted several distinct tendencies in the use of participle clauses, which, in our opinion, might be explained by the mother tongue influence of the learners.

ADVERBIAL CLAUSES

Subordinate adverbial clauses are connected to the main clause of the sentence by two types of subordinators, namely, a conjunction and a preposition; alternatively, they may have a zero subordinator in which case they are also termed *supplementive* clauses (Biber et al. 1999: 820). Our data contains all types of subordination (Table 2) and, interestingly enough, the most frequent subordinators are the same both for speech and written language.

Table 2. Adverbial *-ing* clauses in learner speech and writing.

	LICLE		LINDSEI-LITH	
	Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency
Conjunction + <i>-ing</i> clause	141	74	21	25
Preposition + <i>-ing</i> clause	358	188	29	34
Supplementive clause	352	185	51	60
TOTAL	851	447	101	119

The most frequent conjunction used by Lithuanian learners to introduce a participle clause is *while*, for example:

- (1) *I mean er .. well of course I can say **while choosing these studies** I had a little bit different view of if/what I'm going to do here (LINDSEI-LITH-LT004)*
- (2) *it's it's very nice because you get a good experience with commu= by **while communicating** with with other people so I think it's it's very nice er experience (LINDSEI-LITH-LT016)*
- (3) *But actually, the main thing is that **while reading** a person gets a dose of pleasure. (ICLE-LT-VI-0015.1)*

- (4) *Consequently, students will be more careful and responsible **while choosing the study programmes**, because they will actually invest their money to their future.* (ICLE-LT-VI-0160.3)

The data shows that the conjunction *while* is used by Lithuanian learners to express time relationship (see examples 1-4). Actually, all participle clauses with *while* in our data are clauses of time. According to Biber et al. (1999: 849), this subordinator is indeed used to mark relationships of time in conversation, yet “80% of occurrences in academic prose mark concession/contrast”. As it appears, Lithuanian learners seem to be using this subordinator both in speech and writing for the expression of a single function, i. e. time relationship, which could be seen as a teaching/learning-induced consequence. On the other hand, it could also be interpreted in the context of the mother tongue influence because *while* is generally understood as an equivalent of the Lithuanian adverb *kol* ‘in the meantime, while, till’ used exclusively for marking a time relationship.

The other conjunctions that appear in the two corpora are *when*, *but* (in LINDSEI-LITH) and *when*, *thus*, *but*, *although* (in LICLE), and in both corpora time subordinators are among the most frequent. In other words, we found that the majority of adverbial *-ing* clauses with subordinators are clauses of time (74% in speech and 35% in written language) whereas other types of relationship (manner, contrast, concession) are less numerous and less varied in our data.

Adverbial clauses of time also dominate among clauses introduced by the second type of subordinator, i.e. preposition, which is particularly evident in the spoken corpus, where adverbial clauses have the following prepositional subordinators: *after*, *before*, *by*, *without* and *for*. Similarly, in writing our learners use *after*, *without*, *before*, *instead of* etc. The most frequent preposition in writing, however, is *by* which introduces an adverbial clause of manner and accounts for 60% of all *-ing* clauses introduced by prepositional subordinators, for example:

- (5) *We could change our life **by changing our thinking**.* (ICLE-LT-VY-0080.4)
 (6) *We also get knowledge **by hearing some lectures or by discussing particular issues with other people**.* (ICLE-LT-VI-0015.1)

Finally, the third type of adverbial *-ing* clauses is the so-called supplementive clause which has a zero subordinator. The absence of subordinator leaves much ambiguity in the interpretation of the relationship between the main clause and the subordinate *-ing* clause, for example:

- (7) ***Sitting at the TV** people try to forget their problems and run away from their current reality.* (ICLE-LT-VY-0104.5)

The adverbial clause in (7) could be interpreted as a clause of time, manner or condition. Such unspecified semantic role of the supplementive clauses, as argued in Biber et al. (1999: 832) is mostly characteristic of fiction and narrative texts and very rare in spoken English. Interestingly, our data shows that supplementive clauses are more frequent in Lithuanian learner speech than adverbial clauses with an explicit subordinator (see Table 1), while in written English, clauses with a subordinator are more frequent than supplementive clauses, which, possibly, indicates learners' attempt to explicitly mark semantic relation between the main and subordinate clause.

One of the reasons behind it could be the learners' mother tongue. As mentioned above, Lithuanian has many participle forms and also participle clauses (Lith. *dalyvinės žodžių grupės*, cf. Ambrazas 2006). Participle clauses in Lithuanian usually have no subordinator or other explicit markers of their meaning, while their relation to the main clause is to be inferred from the lexical meaning of the predicate in the main clause or the meaning of the participle, or the context (Ambrazas 2006: 629). Evidently, the learners do not feel it important or necessary to specify the meaning of the English participle clause just as they do not do it in their mother tongue. The following examples from our data would be translated into Lithuanian with participle clauses:

- (8) *hm .. well in this case I guess it's the first option (eh) because . **looking at her character** I doubt she's worth (eh) (em) <starts laughing> a nice picture <stops laughing>* (LINDSEI-LITH-LT054)
- (9) *so yeah I guess all in all despite the: the difficulties <starts laughing> I faced **not knowing French language** I really enjoyed my Erasmus Program experience* (LINDSEI-LITH-LT049)
- (10) *I spend a lot of time with my friends and my boyfriend I don't know I like to spend time **walking out in the nature** (erm) I don't know but most of my time is taken by my studies* (LINDSEI-LITH-LT052)

Several studies have convincingly shown that the use of supplementive clauses might lead to the so-called dangling modifiers, i. e. *-ing* clauses with an implicit subject that differs from that of the main clause. Several studies of learner English reported the problem of dangling modifiers (Granger 1997, Springer 2012), which are viewed as unacceptable in grammars (cf. Swan 1984: 455, Biber et al. 1999: 829) as they lead to absurd interpretations. Such cases of misuse have also been established in our study, for example:

- (11) ***By increasing costs for studies**, professors' salaries will become possible to be increased.*
<ICLE-LT-VI-0161.3>

Clearly, the implied agent of the *-ing* clause in (8) is the government or the authorities while the agent of the main clause is *salaries*. The two agents make the

interpretation of the whole clause problematic. As mentioned above, such cases of misuse of *-ing* clauses are present in our data yet they are not very numerous.

Whereas dangling modifiers are problematic to learners of different mother tongue backgrounds, our study revealed yet another tendency in the use of adverbial *-ing* clauses which is observable in Lithuanian learner speech and writing and which deals with positioning of adverbial *-ing* clauses in relation to the main clause. Data from the BNC suggests that “final position is the unmarked choice for non-finite adverbial clauses in all registers” (Biber et al. 1999: 831) while initial position, particularly in case of supplementive clauses, is used to contribute to a narrative in fiction by setting a frame of activity or showing a sequence of events. Our study showed that initial position is the most preferred choice of Lithuanian learners, which seems to indicate a particular strategy in writing and speaking. Let us consider the following examples:

- (12) *Secondly, after **having passed through the stage of picture drawing**, a child is ready for the second phase of writing adventure, i.e., *helshe is ready to start writing letters (...)* <ICLE-LT-VI-0198.4>*
- (13) *By **charging language** one makes it sound vivid and attractive.* <ICLE-LT-VI-0249.6>
- (14) ***after watching it** as well as **after reading the book** you just can't stop thinking about it* <LT037>
- (15) ***Talking about studies in general**.* <breathes> *I think they . eh the studies different from em my life at school.* <LT017>

Lithuanian learners seem to prefer the initial position both in writing and in speech: nearly half of all adverbial *-ing* clauses in speech and nearly 70% of adverbial clauses in writing are placed before the main clause. In contrast, positioning of adverbial clauses has not found to be a problem area for Dutch learners of English (Springer 2012: 115). Our observation seems to indicate a conscious attempt on the part the learner to structure the message of the utterance or sentence more cautiously, and this is often done by constructing sentences in such a way which helps to preserve the chronological sequence of the events. So in examples 12-15, the action expressed in the participial clause precedes in time the action of the main clause, i.e. *passing through the stage* is prior to *being ready* (12), *charging* precedes *making it sound vivid and attractive* (13), *watching* and *reading* take place prior to *thinking* (14), and *talking about studies* in (15) is prior to *thinking*. In other words, *-ing* clauses set the background against which the idea of the main clause should be interpreted. The evidence, however, is not conclusive and more data would be needed to see whether the tendency to front adverbial *-ing* clauses is not a developmental feature that disappears at the higher level of proficiency, just as Springer's study found in the case of Dutch learners at the level of C1/C2.

ADNOMINAL CLAUSES

It has been claimed in the literature (Granger 1997: 6) that adnominal clauses are significantly underused and misused by learners. The choice between *-ing* clauses and finite clauses deserves a careful examination. Biber et al. suggest that *-ing* clauses, which are seen as equivalents of a full relative clause, are preferred over finite clauses for structural reasons: “many of the common *-ing* verbs occurring in postmodifying clauses are stative in meaning” (1999: 831-832), for example:

(16) *a matter **concerning the public interest***

(17) *a society **consisting of educated people***

Our data shows that Lithuanian learners tend to opt for a postmodifying *-ing* clause when a full relative clause would be a better choice. On the one hand, the learners seem to be unaware of semantic properties of the verbs they use in the participial form, so the subtle difference between *attending* and *who attend* (see the examples below), or *living* and *who lives* remains unnoticed:

(18) *Though the majority of people **attending the Sunday mass** is of older age they have a huge influence on younger generations. (ICLE-LT-VI-0205.5)*

(19) *I was talking to Puipa and we discussed that a director **living with his: er: creation with his film for a long time. a year or even more. er** he starts seeing things which maybe hm really do not exist. (LINDSEI-LITH-LT009)*

On the other hand, this structural distinction is absent in Lithuanian, where any verb can form a postmodifying participle clause. The only restriction in the choice of participle clause over a relative clause might be seen as stylistic preference because *-ing* clauses are perceived as a more compact, condensed way of expression, which characterizes a more formal, written style (Nauckūnaitė 2003, Župerka 1995). So it is possible to suggest that Lithuanian learners simply transfer a syntactic pattern from their mother tongue by using an English participle to construct a postmodifying construction. In some cases, such clauses are perhaps seen as more suitable for written language (examples 20 and 21 below) while in speech they should be easily understood by the hearer no matter how far the clause is removed from the noun it modifies (example 22):

(20) *Generally, the majority of linguists and other scholars **working in the fields focusing on language investigation** agree upon the idea that the language a person uses reflects one's world-perception and attitudes towards various issues. (ICLE-LT-VI-0070.2)*

(21) *I think children have a wrong perception of people **committing acts of violence** that are shown on television. <ICLE-LT-VY-0030.2>*

- (22) *I really like watching movies for example .. (...) well sometimes I just wanna . watch some very simple .. **not making me think of serious things** like . I don't know comedies or something* (LINDSEI-LITH-LT004)

The fact that this kind of misuse of adnominal clauses was established both in learner speech and writing implies that it can result from mother tongue influence.

CONCLUDING REMARKS

The role of learner corpus research can hardly be overestimated. It provides empirical evidence about the process of foreign language learning and teaching. Moreover, studies of learner language also allow us to investigate mother tongue influence, thus their findings could inform the process of language teaching and assessment. In Lithuania, judgments about proficiency of Lithuanian learners who learn English as a foreign language are often based on teachers' experience and intuition while corpus research in the field remains very limited. The compilation of the first learner language corpora that contain written and spoken English produced by native speakers of Lithuanian will hopefully encourage more studies in the field. Furthermore, corpus research should also be seen as an important and reliable tool that enhances the understanding of the CEFR levels which are increasingly used yet differently understood among the broad community of English teachers in Lithuania. The case study presented in the article, demonstrates that Lithuanian EFL learners have specific problem areas which could be directly related to their native language. So, clearly, the common reference points, as they are presented in the Common European Framework of Reference, need to be specified for every specific group of learners.

The study of the use of *-ing* clauses by Lithuanian learners shows that *-ing* clauses are problematic to Lithuanian learners. Looking for possible reasons of the abuse and misuse of *-ing* clauses we addressed the issue of the influence of the mother tongue. Although the influence of the first language is not easy to prove, our corpus-based analysis of the spoken and written English of Lithuanian learners can help us ease the difficulty in establishing transfer. At this stage of research we can only draw some very tentative conclusions regarding the mother tongue impact. In foreign language learning in general a whole array of factors are at work so disentangling them is not an easy task, but contrasting data from spoken and written learner language might indicate a way forward.

Jonė Grigaliūnienė, Rita Juknevičienė

TEKSTYNAIS PAGRĮSTI MOKINIŲ KALBOS TYRIMAI: SAKYTINĖS IR RAŠY TINĖS KALBOS
GRETINIMAS

Santrauka

Straipsnyje nagrinėjami lietuvių gimtosios kalbos vartotojų, besimokančių anglų kaip svetimšios, sakytinės ir rašytinės kalbos skirtumai. Lietuvių studentų vartojama anglų kalba, ypač sakytinė, iki šiol mažai tyrinėta. Straipsnyje pristatomi Vilniaus universiteto Filologijos fakulteto Anglų filologijos katedroje sukaupti mokinio kalbos tekstynai: rašytinės kalbos tekstynas (LICLE) ir sakytinės anglų kalbos tekstynas (LINDSEI-LITH). Abu tekstynai – tarptautinio projekto dalis: LICLE – papildys tarptautinį besimokančiųjų rašytinės anglų kalbos tekstyną, kaupiamą Louvaino (Belgija) universitete, o LINDSEI-LITH – sakytinės anglų kalbos tekstyną, kaupiamą tame pačiame universitete. Straipsnyje nagrinėjami *-ing* predikacinių dėmenų (angl. *-ing clauses*) vartojimo ypatumai lietuvių gimtosios kalbos vartotojų, besimokančiųjų anglų kalbos kaip svetimšios, sakytinėje ir rašytinėje kalboje. Lingvistinėje literatūroje jau buvo rašyta apie tokių kalbinių vienetų vartojimą mokinių kalboje, buvo užsiminta, kad *-ing* dėmenys negimtakalbių dažnai yra neteisingai vartojami. Kai kurie autoriai šiuos sunkumus aiškina gimtosios kalbos įtaka, kiti nurodo mokymo metodų ir būdų poveikį, dar kiti teigia, jog tokie dėmenys paprastai siejami su aukštesniu – B2, C1 – užsienio kalbos mokėjimo lygiu. Lietuvių kalba yra unikali veiksmažodinių formų (dalyvių, padalyvių, pusdalyvių) įvairove bei skaičiumi ir plačiu jų vartojimu tiek rašytinėje, tiek sakytinėje kalboje. Straipsnyje keliamas klausimas apie galimą gimtosios kalbos įtaką studentų kalbai. Nors ir yra įrodymų apie tokį gimtosios kalbos poveikį, straipsnyje teigiama, kad lietuvių studentų, besimokančių anglų kalbos kaip svetimšios, vartosenos ypatumai turi būti siejami su gimtosios kalbos, mokymo metodų ir studentų kalbos mokėjimo įtakomis. Tam nustatyti reikia išsamesnių *-ing* predikacinių dėmenų sintaksinių, semantinių ir stilistinių tyrimų.

REFERENCES

- Ambrasas, V. (ed.). 1997. *Lithuanian Grammar*. Baltos lankos: Vilnius.
- Ambrasas, V. (ed.). 2006. *A Grammar of Modern Lithuanian*. 4th edition. Vilnius: Mokslo ir enciklopedijų leidybos institutas.
- Aprijaskytė, R. 1975. *Beware of Pitfalls*. Vilnius: V. Kapsuko universitetas.
- Aprijaskytė, R. & Pareigyte, E. 1982. *Some Lexical Difficulties for the Lithuanian Learner of English*. Vilnius: VU leidykla.
- Barlow, M. 2005. Computer-based analyses of learner language. In R. Ellis & G. Barkhuizen (eds.) *Analyzing Learner Language*. Oxford: Oxford University Press. 335-357.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. 1999. *The Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Bikelienė, L. 2008. Resultative connectors in advanced Lithuanian learners' English writing. *Kalbotyra* 59(3), 30-37.
- Bikelienė, L. 2009. Priešpriešos konektorių vartojimas besimokančių anglų kalbos ir anglakalbių studentų rašto darbuose. *Kalbotyra* 61(3), 21-35.
- Bikelienė, L. 2012. *Teksto jungimo priemonių vartoseną gimtakalbių ir negimtakalbių studentų rašto darbuose anglų kalba. Gretinamoji konektorių analizė*. Daktaro disertacija. Vilnius.
- Burkštaitienė, N. 2006. Learning foreign languages at university level in the context of constructivist paradigm. *Acta Paedagogica Vilnensia* 16, 18-26.
- Cosme, Ch. 2008. Participle clauses in learner English: the role of transfer. In G. Gilquin, Sz. Papp and M. B. Diez-Bedmar (eds.). *Linking up contrastive and learner corpus research*. Amsterdam: Rodopi, 177-198.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- De Smed, H. 2010. English *-ing*-clauses and their problems: the structure of grammatical categories.

- Linguistics* 48(6): 1153-1193.
- Downing, A. & P. Locke. 2006. *English Grammar. A University Course*. 2nd edition. London: Routledge.
- Gilquin, G., De Cock, S. & Granger, S. (eds.). 2010. *LINDSEI: Louvain International Database of Spoken English Interlanguage*. Louvain-la-Neuve: Presses universitaires de Louvain.
- Granger, S. 1997. On identifying the syntactic and discourse features of participle clauses in academic English: native and non-native writers compared. In J. Aarts, I. de Monnick & H. Wekker (eds) *Studies in English Language and Teaching – In honour of Flor Aarts*. Amsterdam: Rodopi. 185-198.
- Granger, S. 2002. A bird's-eye view of learner corpus research. In S. Granger, J. Huang & S. Petch-Tyson (eds.). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins. 3-33.
- Granger, S. 2003. International Corpus of Learner English: a new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly* 37(3): 538-546.
- Granger, S., Dagneaux, E., Meunier, F. & Paquot, M. (eds.). 2009. *International Corpus of Learner English*. Version 2. Louvain-la-Neuve: Presses universitaires de Louvain.
- Grigaliūnienė, J., Bikilienė, 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 (III): 62-66.
- Grigaliūnienė, J. & R. Juknevičienė. 2011. Formulaic language, learner speech and the spoken corpus of learner English LINDSEI-LITH. *Žmogus ir žodis* 13 (III): 12-18.
- Grigaliūnienė, J. & R. Juknevičienė (*forthcoming*). Recurrent formulaic sequences in the speech and writing of the Lithuanian learners of English.
- Hawkins, J. A. & Buttery, P. 2009. Using learner language from corpora to profile levels of proficiency: Insights from the English Profile Programme. In L. Taylor & C. J. Weir (eds) *Language Testing Matters: Investigating the Wider Social and Educational Impact of Assessment*. Cambridge: Cambridge University Press, 158-175.
- Hawkins, J. A. & L. Filipovic. 2012. Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework. *English Profile* (vol. 1). Cambridge: UCLES / Cambridge University Press.
- Juknevičienė, R. 2008. Collocations with high frequency verbs in learner English: Lithuanian learners vs. native speakers. *Kalbotyra* 59(3): 119-127.
- Juknevičienė, R. 2011. *Lexical bundles in non-native speaker and native speaker written English*. Summary of doctoral dissertation. Vilnius: Vilnius University Press.
- Katkuvienė, L. E. 2003. *Writing Matters*. Vilnius: Vilnius University Press.
- Katkuvienė, L. & I. Šeškauskienė. 2005. *Research Matters*. Vilnius: Vilnius University Press.
- Klimas, A. 1987. The Lithuanian participles: their system and functions. *The Lithuanian Quarterly Journal of Arts and Sciences*. 33(1).
- Kortmann, B. 1995. Adverbial participle clauses in English. In M. Haspelmath and E. König (eds.). *Converbs in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter. 189-237.
- Koskinen, L., A. Mauranen & A. Virkkunen. 1987. Teaching English abbreviated clause for Finnish readers. *Reading in a Foreign Language*. 4(1): 9-19.
- Myles, F. 2005. Interlanguage corpora and second language acquisition research. *Second Language Research* 21(4): 373-391.
- Myles, F. 2012. Learner corpora and SLA theory. In *Learner corpus research 2012: "20 years of learner corpus research: looking back, moving ahead"*. Louvain-la-Neuve: Centre for English Corpus Linguistics. 10.
- Nauckūnaitė, Z. 2003. Loginiai ir lingvistiniai sakytinės ir rašytinės raiškos skirtumai. *Žmogus ir žodis* (I): 78-83.
- Pravec, N. 2002. Survey of learner corpora. *ICAME Journal* 26: 81-114.
- Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. 1985. *A Comprehensive Grammar of The English Language*. Harlow: Longman Group Limited.
- Rimkutė, E. 2006. The usage of grammatical forms of contemporary Lithuanian language in the morphologically annotated corpus. *Lituanistica* 66/2: 34-55.
- Schmitt, N. 2004. *Formulaic Sequences: Acquisition, Processing and Use*. Amsterdam: John Benjamins Publishing.
- Scott, M. 2008. *WordSmith Tools*. Version 5. Oxford: Oxford University Press.
- Springer, P. E. 2012. *Advanced learner writing. A corpus based study of the discourse competence of Dutch writers of English in the light of the C1/C2 levels of the CEFR*. Uitgeverij: Boxpress Oisterwijk.
- Stanevičienė, B. Z. 2007. Apie formaliosios akademinės santraukos rašymo mokymo turinį. *Acta Paedagogica Vilnensia* (18): 151-157.
- Šeškauskienė, I. 2008. Hedging in ESL: A case study of Lithuanian learners. *Kalbu studijos* 13: 71-76.
- Thewissen, J. (*forthcoming*) Capturing L2 accuracy developmental patterns: Insights from an error-

tagged EFL learner corpus. Special issue of *Modern Language Journal* on qualitative learner corpus research.

Thewissen, J., Y. Bestgen & S. Granger. 2006. Using error-tagged learner corpora to create English-specific CEF descriptors. Conference proceedings. Third Annual Conference of EALTA, Krakow, 19-21 May.

Waibel, B. 2007. Phrasal verbs in learner English: a corpus-based study of German and Italian students. PhD dissertation. Freiburg: Albert Ludwigs-Universität Freiburg.

Wray, A. 2002. *Formulaic Language and the Lexicons*. Cambridge: Cambridge University Press.

Župerka, K. 1983. *Lietuvių kalbos stilistika*. Vilnius: Mokslo.