

**UNIVERSITY OF TARTU
DEPARTMENT OF ENGLISH STUDIES**

**THE USE OF PHRASAL VERBS BY ESTONIAN EFL
LEARNERS: A CORPUS-BASED STUDY**
BA thesis

**SANDRA-LEELE TOOM
SUPERVISOR: JANE KLAVAN, PhD**

**TARTU
2020**

ABSTRACT

Phrasal verbs are an integrate part of English grammar and often difficult for foreign learners of English (EFL) to grasp. In addition to understanding the meaning of phrasal verbs, it is important for learners to know how and when it is most appropriate to use them. The aim of this thesis is to analyze, by replicating the study by Gilquin (2015), the use of phrasal verbs by Estonian-speaking EFL learners in spoken and written language. More specifically, this paper investigates how often Estonian-speaking EFL learners use phrasal verbs in spoken and written language, which phrasal verbs are used most frequently, and how their use compares to the use of phrasal verbs by native speakers of English. To achieve this, a corpus-based analysis was carried out. The corpus used for written language analysis was the *Tartu Corpus of Estonian Learner English* (TCELE) and for spoken language analysis, the Estonian subcorpus of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI-EST) was used.

The thesis begins with an introduction, which gives an overview of the motivation behind this paper as well as a summary of the core chapters that follow. In the literature review, phrasal verbs, learner language and learner corpora are defined and previous research on relevant topics discussed. The empirical section introduces the methodology, which is followed by the results and the discussion. The thesis ends with a conclusion.

TABLE OF CONTENTS

ABSTRACT	2
LIST OF ABBREVIATIONS	4
INTRODUCTION	5
1. LITERATURE REVIEW	7
1.1 Phrasal Verbs	7
1.2 Learner Language and Corpora	12
1.3 EFL Learners and Phrasal Verbs	14
1.4 LINDSEI Project.....	18
2. EMPIRICAL ANALYSIS	19
2.1 The corpora	19
2.2 The phrasal verbs	21
2.3 Methodology	21
2.4 Data analysis and results.....	22
2.4.1 The use of phrasal verbs by Estonian learners in spoken and written English	22
2.4.2 The use of phrasal verbs by Estonian EFL learners and native speakers.....	26
2.5 Discussion.....	28
CONCLUSION	32
REFERENCES	35
APPENDICES	37
Appendix 1	37
Appendix 2.....	38
RESÜMEE.....	39

LIST OF ABBREVIATIONS

EFL – English as a Foreign Language

LINDSEI-EST – The Estonian subcorpus of the *Louvain International Database of Spoken English Interlanguage*

TCELE – *Tartu Corpus of Estonian Learner English*

INTRODUCTION

Phrasal verbs are considered as a difficult topic for learners of English as a foreign language (EFL) to grasp because they can be defined in different ways. Furthermore, learning the meaning of phrasal verbs can also be a difficult task, as it is not always apparent from the words that make up a phrasal verb. As described by De Cock (2006), EFL learners face several issues when using phrasal verbs, such as avoidance, style deficiency, semantic confusion, lack of collocational awareness, among others. Despite their challenging nature, phrasal verbs cannot be avoided when learning English, because they are an essential part of the language and add to the native-like quality of a learner's speech.

The native-like quality and importance of phrasal verbs is one of the reasons why it is necessary to research phrasal verbs, especially among EFL learners. Collecting learner data and analyzing how they use phrasal verbs provides linguists and teachers with an opportunity to discover the major difficulties learners face and develop teaching methods that target their specific needs. Although the topic is widely researched in the EFL field, there has not, unfortunately, been a lot of research done about the use of phrasal verbs in the context of Estonian EFL learner language.

There has been, however, research done about phrasal verbs on other EFL learners. One such noteworthy example is Gilquin's (2015) study which analyzed the use of phrasal verbs by French-speaking EFL learners and compared their results to those of native speakers. Following Gilquin's (2015) method, the thesis at hand intends to contribute to this area of research by studying the use of phrasal verbs by Estonian EFL learners.

The aim of this paper is to answer the following research questions: (a) how often are phrasal verbs used by Estonian EFL learners in general, as well as in written and spoken

language individually; (b) which phrasal verbs are used most frequently; (c) how does the use of phrasal verbs by Estonian EFL learners compare to the use of phrasal verbs by native speakers. To answer these questions, an empirical analysis was carried out using data from two Estonian EFL learner corpora: the *Tartu Corpus of Estonian Learner English* (TCELE) for written language data and the Estonian subcorpus of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI-EST) for spoken data.

The thesis is divided into two major chapters: 1. Literature Review and 2. Empirical Analysis. The literature review first provides an overview of phrasal verbs and phrasal verbs in Estonian. The following subsection discusses learner language and corpora as well as previous research done on Estonian EFL learners. Then, a summary of previous research on EFL learners and phrasal verbs is given, including a detailed overview of Gilquin's (2015) study. Finally, the chapter ends with a summary of the LINDSEI project.

The first subsection of the Empirical Analysis chapter gives a comprehensive description of the learner corpora used for the empirical analysis. The following section describes which phrasal verbs were used for the analysis. The third subsection introduces the methodology and is followed by the data analysis and results. The chapter ends with a discussion of the findings.

1. LITERATURE REVIEW

1.1 Phrasal Verbs

There are different ways to categorize or define phrasal verbs. Online dictionaries, such as the Oxford (n. d.), Cambridge (n. d.) and Merriam-Webster (n. d.) Learner's Dictionaries, mostly provide a similar definition for phrasal verbs: a construction consisting of a verb and a particle (preposition or adverb or both), which as a whole unit carries a different meaning from the original verb, e.g. *come up* = *appear*. However, when looking at different grammar books, the definitions begin to somewhat vary and get more complicated.

Alexander (1988: 152) points out that in a general sense, the term 'phrasal verb' may be used to describe any commonly used verb + preposition/adverb construction. Should we follow this broad definition, phrasal verbs fall into three categories. These categories are essential combinations, non-essential combinations and idiomatic combinations (Alexander 1988: 152). Essential combinations are constructions in which a verb must appear with a preposition when it has an object, e.g. *listen to music* (Alexander 1988: 152). In non-essential combinations, verbs do not need to appear with a preposition, but may, to reinforce the meaning of the verb, e.g. *Drink your milk!* vs. *Drink up your milk!* (Alexander 1988: 152). Alexander (1988: 152–153) describes idiomatic combinations as constructions that carry a completely different meaning to the original verb, e.g. *make off* = *run away*, *make up* = *invent*.

The aforementioned categories are similar to what Quirk et al. (1985: 1150), Carter and McCarthy (2006: 429) and Foley and Hall (2012: 282) describe as multi-word verbs. All of these authors (Quirk et al. 1985: 1150; Carter & McCarthy 2006: 429–430; Foley & Hall 2012: 282) begin with a similar definition to describe multi-word verbs: multi-verb words consist of a verb plus a particle, which can be a preposition or an adverb, and they fall into three categories, phrasal verbs, prepositional verbs and phrasal-prepositional verbs. Quirk et al. (1985: 1150) and

Carter and McCarthy (2006: 429) agree that for these constructions to be considered as multi-verb words, they must behave as a single unit of meaning. Foley and Hall (2012: 282) do not make such a distinction, but do point out that occasionally these combinations carry a different meaning from the original verb.

Depending on the source, the definitions for phrasal verbs vary. The one aspect that various authors (Quirk et al. 1985: 1152–1153; Carter & McCarthy 2006: 429–430; Foley & Hall 2012: 282; Alexander 1988: 154–158) seem to agree on, is the fact that phrasal verbs can be transitive or intransitive. How exactly they define phrasal verbs, or differentiate them from other multi-verb constructions, is where we begin to see differences. For example, Alexander (1988: 154) makes a distinction between phrasal and non-phrasal verbs on the basis of the association between the different parts of the construction, as explained by the following quote: “The combination of verb + preposition or particle can be described as phrasal when the /.../ parts are in common association /.../ and yield a particular meaning which may be obvious /.../ or idiomatic /.../.” However, Quirk et al. (1985: 1152 – 1155) and Carter and McCarthy (2006: 431–436) differentiate phrasal verbs from other multi-word verbs by looking at different grammatical rules (e.g. how they are used with objects) or performing different tests, which will be explained further below.

As can be seen from the previous discussion, there is not one single way to define phrasal verbs. In order to carry out my empirical analysis, a choice had to be made on what I consider as a phrasal verb and to do so, I chose to follow the account of phrasal verbs described in Quirk et al. (1985: 1150–1154). Following is a more detailed overview of the approach taken in Quirk et al. (1985) since this is relevant for the empirical part of this thesis.

As mentioned earlier, Quirk et al. (1985: 1150) consider phrasal verbs as part of the category of multi-word verbs. Multi-word verbs consist of a lexical verb and a particle, and function in sentences as single units (Quirk et al. 1985: 1150). For example, in the phrasal verb *make up*, *make* is the lexical verb and *up* the particle. Particles fall into two overlapping yet distinct categories: prepositions and spatial adverbs (Quirk et al. 1985: 1150). However, some particles can behave both as a preposition or a spatial adverb, depending on the multi-word verb it appears in (Quirk et al. 1985: 1150–1151). For example, particles such as *against*, *at*, *beside*, *from*, *into*, etc. only function as prepositions, whereas *aback*, *aside*, *away*, *forward*, etc. function exclusively as spatial adverbs (Quirk et al. 1985: 1151). However, particles like *about*, *above*, *by*, *in*, *out*, etc. can function as both (Quirk et al. 1985: 1151).

Quirk et al. (1985: 1152–1153) define two categories of phrasal verbs: intransitive phrasal verbs (Type I) and transitive phrasal verbs (Type II). Intransitive phrasal verbs do not take a direct object, while transitive phrasal verbs may take a direct object in a sentence (Quirk et al. 1985: 1153). Type I phrasal verbs, such as *catch on*, *get by*, and *turn up*, consist of a verb and an adverb particle (Quirk et al. 1985: 1152). These phrasal verbs are typically informal and, in most cases, the lexical verb and particle cannot be separated from one another (Quirk et al. 1985: 1152).

In addition to taking a direct object, Type II phrasal verbs also differ from Type I phrasal verbs in the way that they appear in a sentence. As mentioned before, Type I phrasal verbs typically appear together as a single unit (Quirk et al. 1985: 1152). With Type II phrasal verbs, the particle can appear either before or after the direct object (Quirk et al. 1985: 1153), for example ‘please *switch off* the radio’ vs. ‘please *switch* the radio *off*.’ This may be explained by the adverbial function of the particle (Quirk et al. 1985: 1154), in which case the SVOA

(*subject-verb-object-adverbial*) sentence structure is preferred. The particle usually appears in front of the object when the object is long or requires an end-focus (Quirk et al. 1985: 1154). The fact that Type II phrasal verbs can follow this construction makes distinguishing them from free combinations of lexical verbs and particles somewhat difficult.

As briefly discussed earlier, a distinction can be drawn between phrasal verbs and free combinations of lexical verbs and particles, as explained by Quirk et al. (1985: 1152). In the case of phrasal verbs, such as “*give in* (‘surrender’) *catch on* (‘understand’), and *blow up* (‘explode’), the meaning of the combination manifestly cannot be predicted from the meanings of verb and particle in isolation”¹ (Quirk et al. 1985: 1152). However, in free combinations, the verb behaves as a typical intransitive verb and the particle carries its own meaning, as exemplified by “He walked *past*. (= ‘past the object/place’)”² (Quirk et al. 1985: 1152).

Although making the distinction between phrasal verbs and free combinations can be difficult, several tests can be used to illustrate the separability between them. Quirk et al. (1985: 1152) explain that substitutions can be made for both the verb and the particle in free combinations. For example, in *wade across*, *wade* can be substituted with *walk*, *run*, *swim*, etc. and *across* with particles such as *in*, *through*, *over*, etc. (Quirk et al. 1985: 1152). It is also, in many cases, possible to place modifying adverbs between the verb and the particle, for example “Drink *right* up. [and] Walk *straight* in.” (Quirk et al. 1985: 1152). This cannot be done with phrasal verbs. Additionally, the adverb can be placed in front of the verb in free combinations with subject–verb inversions, which is, again, not possible with phrasal verbs; compare “*Out* came the sun. and *Out* he *passed*.” (Quirk et al. 1985: 1153). These test work for both Type I

¹ To avoid confusion due to our department style guidelines, I have replaced the square brackets, which are used in the original text to provide definitions, with normal parentheses.

² To avoid confusion due to our department style guidelines, I have replaced the square brackets, which are used in the original text to provide definitions, with normal parentheses.

and Type II phrasal verbs, with the exception of putting the particle in front of the subject method, which only works with Type I phrasal verbs (Quirk et al. 1985: 1154). The described tests were used in the empirical section of the current paper to analyze and decide which particles were used in phrasal verb constructions.

According to Veldi (2006: 590), there are two types of phrasal verbs in Estonian: particle verbs and expression verbs. “Particle verbs consist of an adverb particle + verb, for example *alla kirjutama* lit. ‘down + write’, i.e. ‘to sign’.” (Veldi 2006: 590). These phrasal verbs can be both idiomatic and non-idiomatic. “Expression verbs consist of a noun/adjective + verb, for example *aru saama* lit. ‘reason + get’, i.e. ‘to understand’.” (Veldi 2006: 590). Expression verbs are typically idiomatic. (Veldi 2006: 590)

Historically, it has been suggested that particle verbs may have entered the Estonian language as loans from German (Hasselblatt 1990, as referenced in Veldi 2006: 590). Veldi (2006: 590) points out that there are similarities between the two, for example in *üles kirjutama* and *aufschreiben* (‘write down’), which both use the particle *up*, while the English equivalent uses *down*. More recently, however, Estonian linguists have discovered thousands of phrasal verbs, which form an extensive system with unique components and interactions between them (Veldi 2006: 590).

There has not been a lot of comparative research done on Estonian and English phrasal verbs or on the use of phrasal verbs by Estonian EFL learners. One noteworthy exception is the study by Kährik (2002). Kährik (2002) analyzed the semantics of the Estonian adverbials *alla* and *maha* and compared their meaning with the English counterpart *down*. Through her cognitive linguistic approach and analysis, Kährik (2002) was able to draw three major conclusions. First, “the spatial meanings of the Estonian adverbial particles can be analysed in

terms of image schema configurations and their transformations” (Kährik 2002: 105). Second, “the figurative senses of *alla*, *maha* and *down* are motivated by the schemas underlying the spatial meanings” (Kährik 2002: 105). And third, Kährik (2002: 106) discovered that although Estonian particle-verb constructions (PVCs) are influenced by the same orientational metaphors as English PVCs, different perceptions of abstract domains were found, caused by the existence of two ‘downward verticality’ categories in Estonian, as opposed to one in English. Her research and findings provided a good start and motivation for future research on the same topic.

1.2 Learner Language and Corpora

Much like defining phrasal verbs, finding one definition for ‘learner language’ is difficult, especially when it comes to a language like English, which is very widely spoken and serves a variety of different purposes. There is, however, a definition that is generally accepted and considered mostly unproblematic. According to Granger (2008: 259), learner language is spoken by speakers who learn a language different from their mother tongue, which is also not an additional official language spoken in their country. From this definition, it becomes apparent why defining learner language is difficult in the context of English. The definition is straightforward when it comes to data collected from English learners from countries like Spain, Sweden or China, where English has no official status, but more complex in countries like India, where a nativized variety of English has developed (Granger 2008: 259). In the context of Estonian EFL learners, this definition is easily applicable and appropriate.

Since this thesis aims to analyze the use of phrasal verbs by Estonian EFL learners and relies on data from learner corpora, it is necessary to also define which type of corpora can be considered as learner corpora. Learner corpus research is a relatively new field of research, which got its start in the late 1980s. Learner corpora can be described as “electronic collections of texts produced by language learners” (Granger 2008: 259) and they have two main functions:

(a) contributing to the Second Language Acquisition theory by providing interlanguage data about second or foreign language learners and (b) giving a better understanding of the factors that influence EFL language. Additionally, learner corpus data can be used to assess the problems learners face and develop better pedagogical tools and teaching methods to more accurately target their needs. (Granger 2008: 259)

Learner corpora, and corpora in general, can be distinguished on the basis of several categories such as size and language among others. The distinction that is most relevant to the thesis at hand is the medium of the language data – spoken vs written. Granger (2008: 261) argues that although there are currently more written corpora than spoken, this is likely to change due to the increasingly easier and quicker compilation process of spoken data between learners. Such a trend is also evident in the context of Estonian EFL learner corpus research.

In recent years, data has been collected on Estonian EFL learners and compiled into different corpora – the *Tartu Corpus of Estonian Learner English* (TCELE) contains data on written language and the Estonian subcorpus of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI-EST) is comprised of spoken language data. Both corpora are described in greater detail in the second chapter of the paper (see section 2.1). There have been several theses written about Estonian EFL learners, which relied on these corpora for data. For example, using data from TCELE, Tammiste (2016) analyzed the use of different collocations in Estonian EFL learners' writing. Undo (2018) worked with TCELE to calculate the error percentage of an automated part-of-speech tagger. Kirsimäe (2017) compiled a mini-corpus of spoken ELF (English as a lingua franca) to analyze the lexico-grammatical features of Estonian EFL learners. Similarly, Rahusaar (2019) in her thesis described the process of compiling the

LINDSEI-EST subcorpus and analyzed how Estonian EFL learners use the word *well* as a pragmatic marker.

1.3 EFL Learners and Phrasal Verbs

Phrasal verbs are known as difficult constructions for EFL learners to master and for this reason, the use of phrasal verbs by EFL learners is a widely researched topic. Because of the difficulty of these constructions, as well as the different definitions they have, learners often use them incorrectly or avoid using them altogether. Depending on the learners' mother tongue, the types of mistakes they most often make differ, which may be influenced by whether or not their mother tongue uses phrasal verbs. De Cock (2006) has highlighted some of the most common errors EFL learners make while using phrasal verbs. Based on data collected from different learner corpora, De Cock (2006) underlines the following main problems: avoidance, style deficiency, semantic confusion, lack of collocational awareness, using idiosyncratic phrasal verbs, and syntactic errors.

Avoidance is likely caused by the lack of phrasal verbs in the learner's mother tongue and has been noted as an issue by different authors (e.g. Dagut & Laufer 1985). Style deficiency is described as the incorrect use of language in informal speech and formal writing. This means that learners use formal or 'bookish' language in their speech, but informal constructions in their formal writing. Semantic confusion refers to learners' incorrect understanding of the meaning of the phrasal verbs, which is likely also connected to their poor collocational awareness of the particular phrasal verb. Furthermore, learners use idiosyncratic phrasal verbs, which are rarely used or do not actually exist in English. Syntactic errors are made with transitive and intransitive phrasal verbs. (De Cock 2006)

Although all of the aforementioned problems are important, the most relevant for the thesis at hand is avoidance, as the paper aims to look at the general use (and not the mistakes)

of phrasal verbs by Estonian EFL learners. It is necessary, therefore, to discuss some of the possible reasons EFL learners might avoid using phrasal verbs.

Dagut and Laufer (1985) found that when compared to native speakers, Hebrew-speaking EFL learners avoid phrasal verbs to a large extent. In their study, Dagut and Laufer (1985: 74) found that 67% of native speakers preferred using phrasal verbs over their single-word equivalent in 15 out of 20 sentences (each sentence containing a different phrasal verb–single-word pair). The authors then carried out three different tests on Hebrew-speaking EFL learners and found that they use phrasal verbs considerably less, even when phrasal verbs were presented in the tests as possible answers (Dagut & Laufer 1985). The authors concluded that Hebrew-speaking EFL learners avoid using phrasal verbs, because they are not present in their mother tongue (Dagut & Laufer 1985: 77). Dagut and Laufer (1985: 77) also point out that phrasal verbs are characteristic of Germanic languages, a claim which is supported by both De Cock (2006: para. 7) and Jacobsen (2012). According to Jacobsen (2012: 18), the frequency of phrasal verb use by Swedish EFL learners, who do have phrasal verbs in their mother tongue, is similar to that of native speakers.

Interestingly, despite the absence of phrasal verbs in Korean, Ryoo (2013) found in her corpus-based study that the use of phrasal verbs by Korean EFL learners does not dramatically differ from that of native speakers. This suggests that there are other factors that cause avoidance of phrasal verbs. Furthermore, Hulstijn and Marchena (1989) found in their study with Dutch EFL learners, who do have phrasal verbs in their mother tongue, that they avoid using phrasal verbs as well, challenging the claims of Dagut and Laufer (1985). They (Hulstijn & Marchena 1989: 250–251) suggest that avoidance may be due to semantics and not necessarily only because a learner's native tongue does not have phrasal verbs. Hulstijn and Marchena (1989:

250) believe this to be true, because they noticed more avoidance among intermediate learners, who may “feel tempted to adopt a play-it-safe strategy by using multi-purpose one-word verbs with general meanings, rather than restricted-purpose phrasal verbs with specific, sometimes idiomatic, meanings.” Proficient learners avoided phrasal verbs less (Hulstijn & Marchena 1989: 250), suggesting that as learners become more knowledgeable in English, they are better able to differentiate phrasal verbs and use them appropriately.

Another possible reason EFL learners avoid using phrasal verbs may be the fact that students are, especially in writing, discouraged from using them due to their informal nature. As referenced in Alangari et al. (2020: 1), several sources (Bailey 2003, Coxhead & Byrd 2007, Swales & Feak 2004) suggest replacing phrasal verbs with one word equivalents. If students are taught to avoid phrasal verbs and receive lower grades if they do happen to use them in their assignments, it can be considered as one of the reasons why EFL learners tend to underuse these constructions.

The thesis at hand replicates the study of Gilquin (2015), who analyzed the use of phrasal verbs by French-speaking EFL learners. Her article (Gilquin 2015) is one of the most recent and thorough studies on phrasal verbs in the EFL field. For this reason, it is necessary to provide an overview of her work, method and major findings, which can be found in the following paragraphs.

The article begins with an introduction, in which the author describes the motivation for the study and provides an overview of the main sections. Following the introduction, Gilquin provides background information about phrasal verbs in the form of a literature overview, focusing mainly on the different ways phrasal verb constructions can be defined. The following section introduces Gilquin’s methodology. The learner population whose use of phrasal verbs

was analyzed is comprised of native speakers of French who are upper-intermediate or advanced EFL learners. The data for learner English is from two corpora: the French subcorpus of the *International Corpus of Learner English* (ICLE-FR) for written language and the French subcorpus of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI-FR) for spoken language. The word count for the corpora is 190,544 and 91,440 words respectively. The data for native English is from the *Louvain Corpus of Native English Essays* (LOCNESS) and the *Louvain Corpus of Native English Conversation* (LOCNEC), with respective word counts of 155,167 and 118,398 words. (Gilquin 2015: 51–57)

Gilquin extracted the phrasal verbs from the four corpora in a concordance program using 24 different particles for the search. Because these particles may also have other grammatical functions, each concordance line had to be analyzed separately. However, the exact criteria for what is considered a phrasal verb construction is not mentioned in the article. (Gilquin 2015: 58).

Gilquin's methodology is both constructional and collocationist. It is constructional because it aims to analyze the use of phrasal verbs at three levels: phrasal verbs in general, the structure of phrasal verbs, and specific phrasal verbs. To achieve this, "the verb functioning with the particle, the structure, (V OBJ Prt), (V Prt OBJ) or (V Prt)³, and the nature of the object, if any (noun or pronoun)" (Gilquin 2015: 58) were manually determined. The analysis was carried out in an Excel spreadsheet for both the learner and native corpora, and compared to one another by means of log-likelihood statistics. The method is collocationist because it investigates the interaction between words and constructions. (Gilquin 2015: 58)

³ To avoid confusion due to our department style guidelines, I have replaced the square brackets, which are used in the original text to provide definitions, with normal parentheses.

At the higher level of analysis, Gilquin (2015: 74) found that French-speaking EFL learners severely underuse phrasal verbs and, unlike native speakers, use them more often in written than in spoken language. At the intermediate level of analysis, which aimed to examine the use of different phrasal verb constructions, more similarities were found between native and non-native speakers. Across media, the rank of different phrasal verb constructions follow the same order for both French-speaking EFL learners and native speakers (Gilquin 2015: 74). The lower level of analysis aimed to highlight the differences in the use of specific verbs and particles between native and non-native speakers. Gilquin (2015: 75) discovered several idiosyncrasies in the EFL language, such as overusing *together* and *back*, as well as a tendency to use the same common verbs in both spoken and written language.

1.4 LINDSEI Project

The data used to analyze spoken language in the empirical part of this paper is from the Estonian subcorpus of the LINDSEI (*Louvain International Database of Spoken English Interlanguage*) project, currently being compiled at the English Department at the University of Tartu. The LINDSEI project was launched by CECL (Center of English Corpus Linguistics) in 1995 as the spoken counterpart to the *International Corpus of Learner English* (ICLE) (CECL n.d.(a): para. 1). The spoken data in the LINDSEI corpus is produced by advanced learners of English from many different mother tongue backgrounds (CECL n.d.(a): para. 1). Currently, 14 international partners, including, for example, Bulgarian, Chinese, Dutch, French, Italian, and others, have completed their projects, while nine international partners, including Norwegian, Estonian, Finnish, Croatian and others, are in the process of completing their components (CECL n.d.(b)).

All components follow the same structure. They each have about 50 interviews, each one of which contains three tasks: set topic, free discussion, and picture description. The

interviews are transcribed based on set markers and linked to a profile, which contains background information about the speaker, interviewer and the interview itself. This information provides the opportunity for studying the influence of certain factors on learner language. (CECL n.d.(a): para. 2).

It should be noted, that there is a similar native speaker counterpart (LOCNEC) to both the written and spoken learner corpora, which makes it possible to compare learner English to native English (CECL n.d.(a): para. 3). Studies that have been carried out using data from LINDSEI illustrate that the database can be utilized to investigate different aspects of learner English including lexis, syntax, phraseology, etc. (CECL n.d.(a): para. 4).

The process of compiling data for the LINDSEI-EST subcorpus is discussed in greater detail in the methodology chapter of the current paper (see section 2.1).

2. EMPIRICAL ANALYSIS

The current chapter begins with an overview of the corpora from which the data for the empirical analysis was extracted from. The following subsection introduces the particles and the grammatical criteria that were used to identify the phrasal verbs in the data, followed by an explanation of the methodology. The chapter ends with the results of the empirical analysis and a discussion of the findings.

2.1 The corpora

To analyze the use of phrasal verbs by Estonian EFL speakers in written language, the *Tartu Corpus of Estonian Learner English* (TCELE) was used. The corpus is comprised of 127 entrance essays written in 2014 by Estonian learners as a part of the entrance examination to the English Language and Literature BA program. The participants were given the task of writing a 200-word essay based on an academic article about the future of the English language. The main features of the TCELE written corpus are as follows: the total word count of the essays is

24,610; the average length of an essay is 193 words, the length varying from 60 to 320 words; all participants have an Estonian citizenship, their mother tongue, however, is not specified; the average age of the participants is 19, ranging from 18 to 35; 88 participants out of 127 are female and 39 are male; the participants had no reference tools available to them; mistakes and illegible words were left in during the process of typing up the corpus. (Tammiste 2016)

To examine the use of phrasal verbs in spoken learner language, 12 transcribed⁴ TCELE interviews from the Estonian Learner subcorpus of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI-EST) were analyzed. The interviewing process was as follows. In the first part of the interview, the interviewees chose a topic out of three options given to them and spoke about it spontaneously for about three to five minutes without prior practice or preparation. The interviewer then asked further questions about what had been discussed as well as general questions about their hobbies and university life. In the second part, the interviewee had to create a story based on four pictures by interpreting the situation depicted on the photos. The interviewer then asked clarifying follow up questions. The interviews were audio recorded and manually transcribed. (Rahusaar 2019)

The main features of the LINDSEI-EST corpus are the following. The 12 interviews amount to 168 minutes of speech and the length of the interviews ranges from 10.5 – 17.5 minutes, with an average length of 14 minutes. The total word count is 21,362 and contains the texts spoken by both the interviewers as well as the interviewees. There were 12 participants, 3 of them male and 9 of them female with an average age of 24 years, ranging from 21 to 37. 7 of the interviewees were 3rd year bachelor's students at the University of Tartu, 3 were 1st year and 2 were 2nd year master's students. All participants were native speakers of Estonian and

⁴ Currently, 17 interviews have been transcribed, however during the empirical analysis of this study, only 12 transcriptions were available.

gave a general overview of their English language background in a learner profile prior to the interview, which included information about how long they have learned English, what other foreign languages they spoke, etc.

2.2 The phrasal verbs

The phrasal verbs were extracted from the TCELE written corpus and LINDSEI-EST spoken corpus by a lexical search carried out with a concordance program (*AntConc version 3.5.8*). Unlike Gilquin (2015: 58), who analyzed 24 of the 25 particles, for this analysis, all 25 particles taken from Huddleston and Pullum (2002: 281 as referenced in Gilquin 2015) were analyzed. The 25 particles are the following: *aboard, about, across, ahead, along, apart, around, aside, away, back, by, down, forth, forward, home, in, off, on, out, over, round, through, together, under, up*.

It is important to note that the words mentioned above can behave as particles in phrasal verbs, but may also have other functions. Therefore, each concordance line had to be analyzed individually to determine the function of the particle. Because Gilquin (2015) did not specify in her article how she determined what is a phrasal verb for her analysis, a definition of phrasal verbs had to be chosen to carry out the empirical analysis of the study at hand. For the purpose of this paper, the tests described by Quirk et al. (1985: 1152–1154), which were discussed in greater detail in the literature review section of this paper, were used to distinguish the particles used in phrasal verbs from particles used in free associations and other grammatical functions.

2.3 Methodology

The aim of this paper is to identify (a) how often phrasal verbs are used by Estonian EFL learners in general, as well as in written and spoken language separately, (b) which phrasal verbs are used most often and (c) how does the use of phrasal verbs by Estonian EFL learners compare to the use of phrasal verbs by native speakers. To achieve this, the text files of the

entrance essays from TCELE and the interviews from LINDSEI-EST were opened in the concordance program and each of the aforementioned particles were individually entered into the program's search bar. The resulting concordance lines were then copied into an Excel spreadsheet and analyzed according to the following criteria: did the word in the concordance line function as a particle in a phrasal verb or did it carry another grammatical function; did the word appear in written or spoken language; in the case of a phrasal verb, what phrasal verb was it.

The concordance lines that contained phrasal verbs were copied into another spreadsheet, excluding the other concordance lines, which were, however, kept in the original file to later calculate frequencies. The identified phrasal verbs were written out and the following frequencies were found: how many different phrasal verbs occurred with each particle; how many times were each of the identified phrasal verbs used; how many times were they used in written and spoken language. For example, in case of the particle *back*, 8 phrasal verbs were identified: *come back*, *get back*, *go back*, *head back* and others. *Come back* was used a total of 4 times, all 4 times in spoken language and with no usage in written language. This analysis was carried out on each of the identified phrasal verbs and the entire process repeated for all 25 particles.

2.4 Data analysis and results

2.4.1 The use of phrasal verbs by Estonian learners in spoken and written English

The query of 25 particles in the two Estonian EFL corpora produced a total of 1,533 concordance lines; 235 (15.33%) of these lines contained phrasal verbs. Of the 25 analyzed particles, 7 were not used in any phrasal verb constructions. These particles were the following: *aboard*, *across*, *ahead*, *forth*, *forward*, *home*, and *round*. The remaining 18 particles all had concordance lines which contained phrasal verbs, the number of concordance lines ranging from

1 to 66. The particle with the most results (66) was *up* and the particle with the least amount of results was *apart* (1). These results are summarized in Table 1.

Table 1. Summarized Concordance Lines Results

Particle	Nr of concordance lines with phrasal verbs	Total nr of concordance lines
Total	235	1533
about	1	171
along	4	9
apart	1	1
around	2	37
aside	1	2
away	13	19
back	17	20
by	4	72
down	7	9
in	15	729
off	3	4
on	30	222
out	44	58
over	15	56
through	4	8
together	6	13
under	2	5
up	66	75

The usage of phrasal verbs was divided fairly evenly between spoken and written language, with 50.64% (119 phrasal verbs) of the results in spoken and 49.36% (116 phrasal verbs) in written language. To analyze each particle individually, relative frequencies had to be calculated, due to the different total word counts of the TCELE and LINDSEI-EST corpora. Like in Gilquin's study (2015), relative frequencies were calculated per 10,000 words and the formula used is the following: $\frac{\text{absolute frequency } (N) \times 10,000}{\text{corpus total word count}}$.

In both spoken and written language, the most frequently used particles were *up*, *out* and *on*. These particles were, however, used more frequently in spoken language, with relative

frequencies of 16.85 words per 10,000 (*up*), 11.23 per 10,000 (*out*), and 7.02 per 10,000 (*on*). In written language, these particles had relative frequencies of 12.19 words per 10,000 (*up*), 8.13 per 10,000 (*out*), and 6.10 per 10,000 (*on*). The relative frequencies of the remaining particles are summarized in Figure 1 and the absolute frequencies highlighted in Appendix 1.

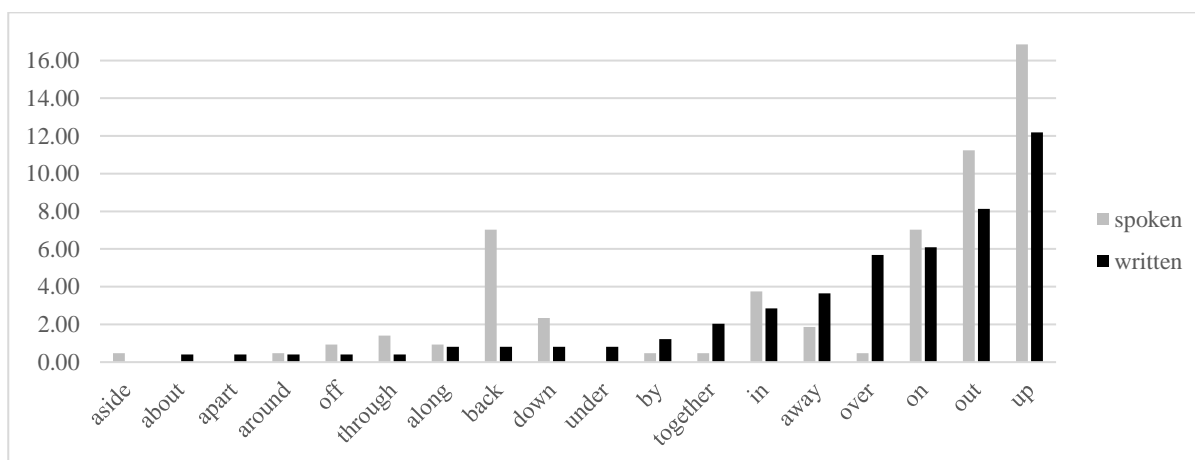


Figure 1. Relative frequencies (per 10,000 words) of particles used in spoken and written language

In total, 109 different phrasal verbs were identified. The particles with the largest variety of phrasal verbs were *out* and *up*, having respectively 21 and 29 different phrasal verbs formed with each particle. The average amount of different phrasal verbs per particle was 6. The most frequently used phrasal verbs were *go on* (15 occurrences), *take over* (13), and *sum up* (12). The most frequently used phrasal verbs in spoken language were *go on* (11 occurrences) and *make up* (7), in written language, *take over* (13) and *sum up* (12). Each particle and the phrasal verbs formed with it, as well as the absolute frequency (N) for each phrasal verb, are summarized in Table 2.

Table 2. Absolute frequency (N) of phrasal verbs formed with each particle

Particle	Phrasal verbs and their respective number of occurrences
about	bring (1)
along	come (1); get (1); go (2)
apart	set (1)
around	be (1); run (1)

aside	put (1)
away	fade (9); get (1); go (1); move (1); stay (1)
back	come (4); get (3); go (5); head (1); hold (1); move (1); send (1); think (1)
by	get (3); run (1)
down	burn (1); dumb (1); go (3); lay (1); write (1)
in	bring (1); come (4); dabble (2); factor (1); go (2); kick (1); result (1); set (2)
off	kick (1); take (2)
on	base (1); build (1); get (1); go (15); hold (4); pass (2); reflect (1); rely (1); take (1); work (1)
out	bring (1); come (5); die (7); fade (1); figure (6); fill (1); find (2); get (2); go (1); hang (1); leave (4); make (1); point (2); print (1); rent (1); step (2); take (1); turn (1); wait (1); wipe (1); work (2)
over	go (1); ice (1); take (13)
through	fall (1); get (1); go (2)
together	bring (3); come (1); put (2)
under	bring (1); come (1)
up	be (1); bring (3); brush (1); build (1); come (3); end (3); fill (1); freeze (1); get (1); give (2); go (1); grow (5); keep (3); light (1); live (2); make (8); mix (2); open (1); pick (1); put (2); rise (1); save (2); set (1); show (1); speak (2); step (1); sum (12); take (2); touch (1)

A total of 63 different lexical verbs were used to construct the phrasal verbs. The most frequently used lexical verb was *go* with 33 occurrences, followed by *take* (20) and *come* (19). All of the lexical verbs used and their absolute frequencies can be found in Figure 2.

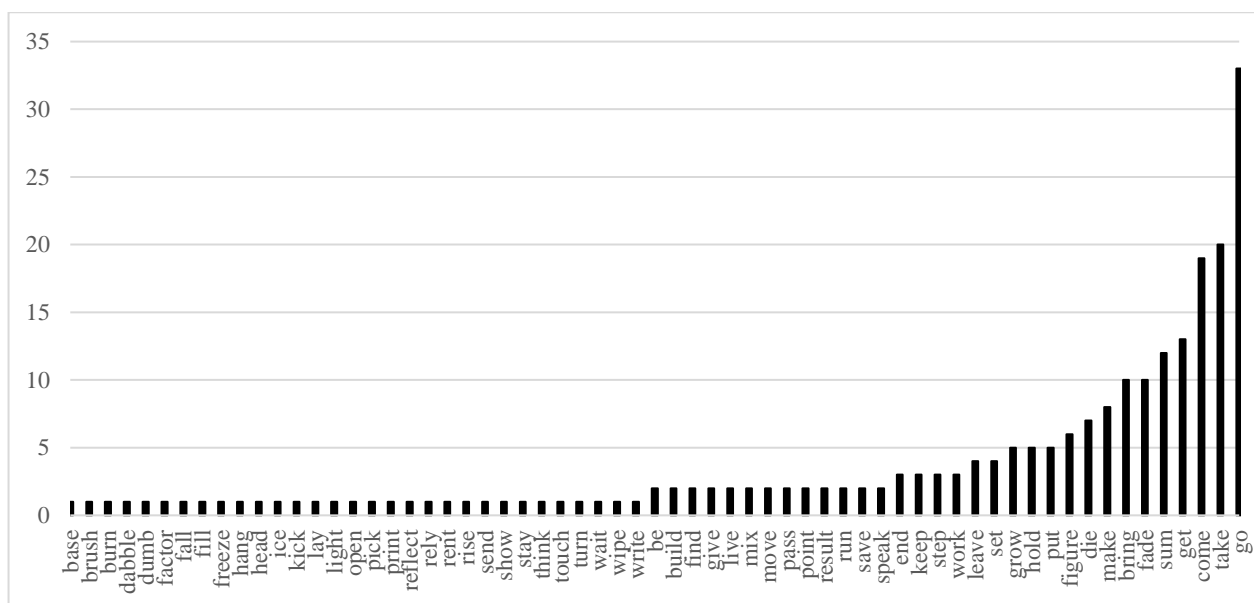


Figure 2. Lexical verbs used in phrasal verb constructions by Estonian EFL learners and their absolute frequencies

2.4.2 The use of phrasal verbs by Estonian EFL learners and native speakers

To compare the usage of phrasal verbs by Estonian EFL learners to Gilquin's (2015) findings about native speakers and French-speaking EFL learners, relative frequencies are used. The formula used for analysis in the previous section was applied here as well.

Estonian EFL learners (NNL Est), like native speakers (NL) (Gilquin 2015: 60), use phrasal verbs more frequently in spoken language than in written language. The relative frequency of phrasal verbs used in spoken language was 56 per 10,000 words (N = 119) and in written language, 47 per 10,000 (N = 116). These results are illustrated in Figure 3.

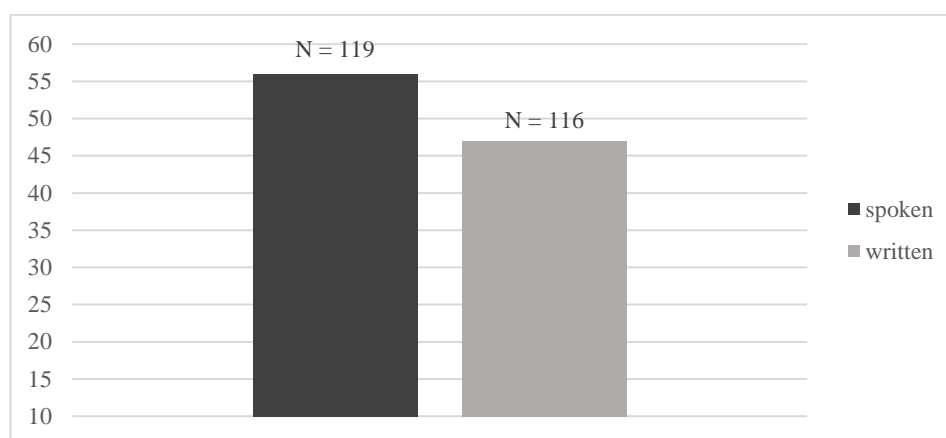


Figure 3. Absolute (N) and relative frequency (per 10,000 words) of phrasal verbs in Estonian EFL spoken and written language

Although the difference between spoken and written language is small, it nevertheless resembles the use of phrasal verbs by native speakers more closely than the use of phrasal verbs by, for example, French-speaking EFL learners, who used phrasal verbs more often in written than spoken language (Gilquin 2015: 60). The non-native speakers of both languages nonetheless underuse phrasal verbs in spoken language to a great degree. According to the findings of Gilquin (2015: 60), the native speaker relative frequency of phrasal verbs in spoken language is nearly 100 per 10,000 words. Estonian EFL learners use phrasal verbs in spoken language about half as much and French-speaking EFL learners even less, with a relative

frequency of about 20 per 10,000 words (Gilquin 2015: 60). In written language, the use of phrasal verbs is more similar between native and non-native speakers. Native speakers showed a relative frequency of about 50 per 10,000 words and French-speaking EFL learners of about 40 per 10,000 (Gilquin 2015: 60). These results are not too dissimilar from the Estonian EFL learner results, 47 per 10,000 words.

Additionally, the relative frequency of all 25 analyzed particles were calculated using the aforementioned formula, the only difference being the word count. For each particle, the combined total word count (45,972) of both the TCELE written and LINDSEI-EST spoken corpora was used. The most commonly used particles by Estonian EFL learners were *up*, *out* and *on*. For the most part, the use of particles in phrasal verbs, particularly the top four most frequently used particles, matched the ones used by native speakers. According to Gilquin (2015: 68), the most commonly used particles by native speakers were *up*, *out* and *back*. The relative frequencies of particles used by Estonian EFL learners and native speakers, taken from Gilquin (2015: 67–68), are summarized in Figure 4 and are shown in further detail in Appendix 2.

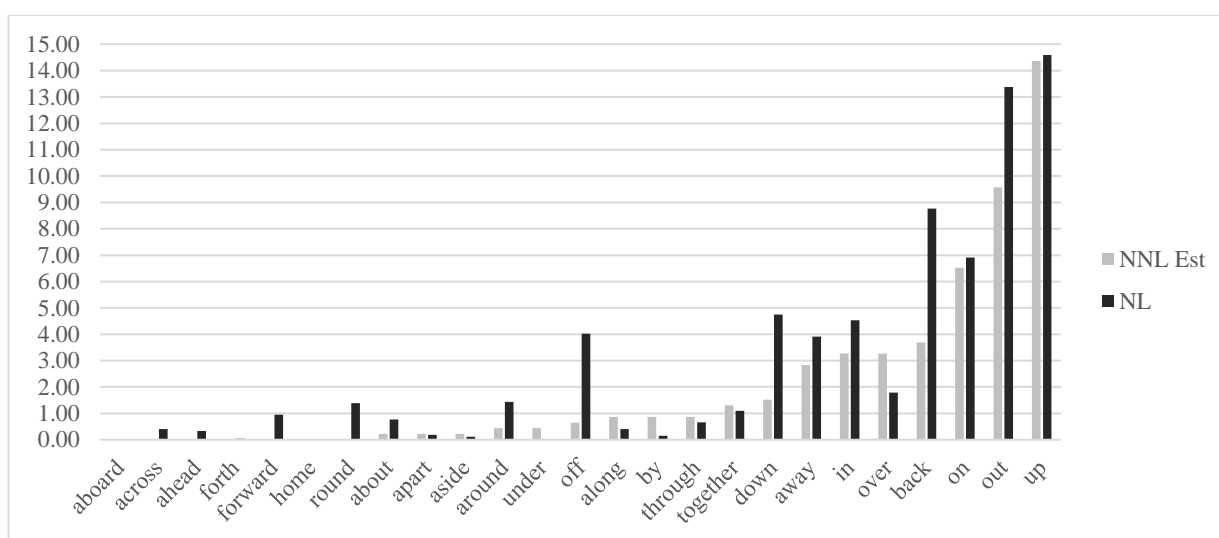


Figure 4. Relative frequency of particles (per 10,000 words) by Estonian EFL learners (NNL Est) and native speakers (NL) (Gilquin 2015: 67–68)

2.5 Discussion

According to the findings of the current paper and the results of Gilquin's (2015) research, in comparison to native speakers, Estonian EFL learners tend to underuse phrasal verbs in spoken language. As described earlier, native speakers use phrasal verbs twice as often as Estonian EFL learners, who do, however, use them more often than French-speaking EFL learners. This may be due to the fact that, unlike French, Estonian does have phrasal verb constructions. As argued by Gilquin (2015), Dagut and Laufer (1985), and De Cock (2006), non-native speakers whose mother tongue does not have phrasal verbs, may avoid using phrasal verbs, and although other researchers have found the opposite to be true as well (Ryoo 2013; Hulstijn & Marchena 1989), this can explain the difference in use between French-speaking and Estonian EFL learners.

Because of the lack of comparative research between Estonian EFL learners' use and native speakers' use of phrasal verbs, finding an explanation for the differences between Estonian learners and native speakers is more difficult. It could be due to the fact that, based on my own personal experience, there does not seem to be as big of an emphasis on phrasal verbs in Estonian language classes, so while Estonian learners do use them in their mother tongue, they may not be as aware. This claim does, however, need further research to be proven correct and valid. The differences between native and non-native use of phrasal verbs could also be due to how phrasal verbs are taught. As described by Alangari et al. (2020), learners are often discouraged from using phrasal verbs in formal settings, which could inadvertently cause learners to avoid them in spoken language as well. This claim also needs to be researched further, especially in the context of Estonian English classrooms and textbooks.

While the use of phrasal verbs was more similar between native and Estonian EFL learners in written language, the validity of these results is questionable. The source data differs greatly from the data described in Gilquin (2015) and may, in general, be inadequate for such analysis. First, the data in TCELE is from entrance essays, meaning there is no set language level (intermediate, proficient, etc.), which can, as described in Hulstijn and Marchena (1989), influence how learners use phrasal verbs. Furthermore, because applicants to the BA program who had scored at least 95 points out of 100 in the National Examination in English or had a certificate proving their proficiency in English did not have to write the entrance essay (Daniel 2015: 23), there is, undoubtedly, data missing that represents proficient learners whose English is most native-like.

Additionally, TCELE is small (24,457 words vs 155,167 in LOCNESS for native speaker data) and the essay lengths vary largely, giving some writers more opportunity to use phrasal verbs and less for others. There is also a noticeable influence of the essay task on the Estonian EFL learner results. The most frequently used phrasal verbs in written language were *take over* and *sum up*, which can be directly linked to the topic of the source text (the future of the English language) and the fact that applicants were writing essays with specific instructions. To resolve this issue, an Estonian EFL learner written corpus would have to be compiled, which more closely resembles the LOCNESS and follows the same criteria. One major difference to be taken into consideration is, for example, the use of a source text in TCELE. LOCNESS texts were literary and argumentative essays written without a source text (LOCNESS n.d.). Additionally, LOCNESS texts had to meet the requirement of at least 500 words (LOCNESS n.d.), while TCELE texts were considerably shorter (average length 193 words).

The reliability of the spoken language material is high, because the data used follows the same guidelines as the data in Gilquin's (2015) study – both studies use the LINDSEI design. There is, nonetheless, an aspect that can cause differences in the analysis of phrasal verbs used in spoken language despite the fact that they come from the same corpora. Gilquin (2015) did not explicitly describe the criteria she used to define phrasal verbs, which means different constructions may have been included or excluded from the analysis. These possible differences make it difficult to draw concrete conclusions about Estonian EFL learners' use of phrasal verbs in comparison to native speakers.

Although analyzing mistakes made by learners while using phrasal verbs was not one of the main aims of this thesis, there were a few occasions where incorrect uses were noticed. These mistakes were mostly spelling errors or using wrong lexical verbs, which closely resembled the correct verb (e.g. *go by* when *get by*, as in 'survive', was meant). There were no major grammar mistakes noticed (tense, etc.) and because mistake analysis was not a priority, semantic mistakes cannot be commented on, as the use of phrasal verbs was not analyzed in such detail.

As briefly mentioned earlier, the criteria which were used to define phrasal verbs certainly had an effect on the empirical analysis of this study. Despite following Quirk et al.'s (1985) tests, defining and choosing which multi-word verbs to consider phrasal verbs proved to be difficult. Phrasal verbs can be defined in many different ways and had another person done the analysis using the same tests, the result might also differ. It would be interesting to carry out a similar analysis, where instead of doing the lexical search using different particles, a set of phrasal verbs most frequently used by native speakers were used. This may even be more

productive, because as can be seen in both this thesis and Gilquin's (2015: 67–68) study, there were particles that were rarely used by both native and non-native speakers.

Further research is necessary to resolve the shortcomings of this thesis and provide a better overview of the use of phrasal verbs by Estonian EFL learners. For example, because the validity of the source material for written language was found to be problematic, the analysis should also be carried out using more reliable data. Additionally, it would be interesting to study if and to what extent Estonian EFL learners avoid phrasal verbs (like in Dagut & Laufer 1985; Hulstijn & Marchena 1989). Depending on the outcome, the reason why Estonian EFL learners might avoid using phrasal verbs could then be analyzed, furthermore, consider whether learner language level has an effect on this. To support the claims made earlier about the possible reasons Estonian EFL learners use phrasal verbs less often than native speakers, the actual methods and materials used to teach English phrasal verbs to Estonian students must also be studied.

Despite the shortcomings, the thesis at hand provides a good starting point for further analysis on the use of phrasal verbs by Estonian EFL learners. To be able to conduct such semantic research, as would be necessary to resolve the issues described above, general analysis and information must be found first. The findings of this paper fulfil this need, providing an overview of the most frequently used phrasal verbs as well as material for future research topics and questions.

CONCLUSION

Phrasal verbs, although seemingly simple, are a notoriously difficult topic for EFL learners to grasp due to the different ways they can be defined and the fact that their meanings are not always obvious from the words the phrasal verbs are made of. Previous research on phrasal verbs in the EFL field has identified some of the most common problems EFL learners face when using phrasal verb constructions. Although there are several issues, the most relevant to the thesis at hand, as discussed in the literature review, is avoidance, defined by De Cock (2006). It has been found by several studies (e.g. Hulstijn & Marchena 1989; Dagut & Laufer 1985; Gilquin 2015) that EFL learners indeed underuse phrasal verbs. Different explanations for the avoidance have been offered, but no definitive conclusions can be made yet.

Despite the fact that phrasal verbs have been widely researched in the EFL field, there is not a lot of research available in the context of Estonian EFL learner language. The thesis at hand intended to contribute to this field by analyzing the use of phrasal verbs by Estonian EFL learners and hopefully provides motivation for future research on this topic.

To achieve this, three research questions were proposed: (a) how often are phrasal verbs used by Estonian EFL learners in general, as well as in written and spoken language individually; (b) which phrasal verbs are used most frequently by Estonian EFL learners; (c) how does the use of phrasal verbs by Estonian EFL learners compare to the use of phrasal verbs by native speakers. These questions were answered by conducting a corpus-based analysis.

Data from two learner corpora (TCELE and LINDSEI-EST) was entered into a concordance program and a lexical search was done with 25 different phrasal verb particles. The resulting concordance lines were analyzed and the identified phrasal verbs were collected into an Excel spreadsheet.

By finding the absolute and relative frequencies of the identified phrasal verbs, I was able to answer questions (a) and (c). Out of the 1,533 concordance lines analyzed, approximately 15% (235) contained phrasal verbs. The use of phrasal verbs by Estonian EFL learners was divided fairly evenly between written and spoken language, with a slightly higher frequency found in spoken language (116 vs 119 phrasal verbs in written and spoken language respectively). These results lead us to conclusion 1: according to the findings of this thesis, Estonian EFL learners tend to underuse phrasal verbs in both written and spoken language.

To be able to compare the use of phrasal verbs by Estonian EFL learners to the use of phrasal verbs by native speakers, relative frequencies were found. The relative frequency of phrasal verbs used by Estonian EFL learners in spoken language was 56 per 10,000 words and in written language, 47 per 10,000. Native speakers, as described in Gilquin's (2015: 60) study, used phrasal verbs more often, with a relative frequency of approximately 100 per 10,000 words in spoken language and 50 per 10,000 words in written language. Thus, we are able to answer question (c) and draw conclusion 2: compared to native speakers, Estonian EFL learners underuse phrasal verbs to a great degree, especially in spoken language, in which case the relative frequency of phrasal verbs used was twice as high among native speakers. The relative frequencies were more similar in written language, but due to the differences in the corpora of learner and native language, no concrete conclusions can be made.

The empirical analysis also provided an answer to question (b). The data analysis showed that of the 109 different phrasal verbs identified, the most frequently used phrasal verbs were *go on* (15 occurrences), *take over* (13), and *sum up* (12). The most frequently used phrasal verbs in spoken language were *go on* (11 occurrences) and *make up* (7), in written language,

take over (13) and *sum up* (12). The most frequently used particles were *up* and *out*, which were the same for native speakers (Gilquin 2015: 68).

Overall, this thesis serves as a starting point for future research on the topic of phrasal verbs in the context of Estonian EFL learners. Several possible research questions and topics can be drawn from the findings and shortcomings of the paper, which would provide further insight into the topic. These and future results can highlight the issues Estonian EFL learners face when using phrasal verbs and aid in developing teaching methods to target their specific needs to improve their understanding and use of phrasal verb constructions.

REFERENCES

- Alangari, Manal, Sylvia Jaworska and Jacqueline Laws. 2020. Who's afraid of phrasal verbs? The use of phrasal verbs in expert academic writing in the discipline of linguistics. *Journal of English for Academic Purposes*, 43, 1-13.
- Alexander, Louis George. 1988. *Longman English Grammar*. Harlow and New York: Longman.
- Bailey, Stephen. 2003. *Academic writing: a practical guide for students*. London and New York: Routledge.
- Cambridge Online Learner's Dictionary. Available at <https://dictionary.cambridge.org/dictionary/learner-english/>, accessed March 22, 2020.
- Carter, Ronald and Michael McCarthy. 2006. *Cambridge grammar of English: a comprehensive guide; spoken and written English grammar and usage*. Cambridge: CUP.
- CECL. n.d. (a). LINDSEI. Available at <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei.html>, accessed May 4, 2020.
- CECL. n.d. (b). LINDSEI partners. Available at <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei-partners.html>, accessed May 4, 2020.
- Coxhead, Averil and Pat Byrd. 2007. Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16: 3, 129-147.
- Dagut, Menachem and Batia Laufer. 1985. Avoidance of phrasal verbs: A case for contrastive analysis. *Studies in Second Language Acquisition*, 7: 1, 73-79.
- Daniel, Anna. 2015. *The Use of Adjectives and Adverbs in Estonian and British Student Writing: A Corpus Comparison*. MA thesis. Department of English, University of Tartu, Tartu, Estonia.
- De Cock, Sylvie. 2006. Learners and phrasal verbs. *MED Magazine*, 35: January–February.
- Foley, Mark and Diane Hall. 2012. *My Grammar Lab. Advanced C1/C2*. Harlow: Pearson.
- Gilquin, Gaëtanelle. 2015. The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory*, 11: 1, 51-88.
- Granger Sylviane. (2008). Learner corpora. In Anke Lüdeling and Merja Kytö (eds.) *Corpus Linguistics. An International Handbook. Volume 1*, 259-275. Berlin and New York: Walter de Gruyter.
- Hasselblatt, Cornelius. (1990), *Das estnische Partikelverb als Lehnübersetzung aus dem Deutschen*. Wiesbaden, Otto Harrassowitz.
- Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hulstijn, Jan and Elaine Marchena. 1989. Avoidance: grammatical or semantic causes? *Studies in Second Language Acquisition*, 11: 3, 241-255.
- Jacobsen, Heidi. 2012. *Figuring out phrasal verbs: An analysis of student use and fluency in Swedish upper secondary schools*. Advanced essay. Linnaeus University, Sweden.
- Kirsimäe, Merli. 2017. *The compilation and lexicogrammatical analysis of an Estonian spoken mini-corpus of English as a lingua franca*. MA thesis. Department of English, University of Tartu, Tartu, Estonia.

- Kährik, Kaja. 2002. *Meaning of alla, maha and down in spatial and metaphorical particle-verb constructions*. MA thesis. Department of Germanic and Romance Language and Literatures, University of Tartu, Tartu, Estonia.
- LOCNESS. n.d. Available at <https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html>, accessed May 25, 2020.
- Merriam-Webster Online Learner's Dictionary*. Available at <https://learnersdictionary.com/>, accessed March 22, 2020.
- Oxford Online Advanced Learner's Dictionary*. Available at <https://www.oxfordlearnersdictionaries.com/>, accessed March 22, 2020.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London, New York: Longman.
- Rahusaar, Anne. 2019. *The compilation of the spoken sub-corpus for the Tartu Corpus of Estonian Learner English*. MA thesis. Department of English, University of Tartu, Tartu, Estonia.
- Ryoo, Mi-Lim. 2013. A Corpus-based Study of the Use of Phrasal Verbs in Korean EFL Students' Writing. *The Journal of Asia TEFL*, 10: 2, 63-89.
- Swales, John and Christine Feak. 2004. *Academic writing for graduate students: Essential tasks and skills*. Ann Arbor, MI: University of Michigan Press.
- Tammiste, Lenne. 2016. *The use of adjective-noun, verb-noun, and phrasal-verb-noun collocations in Estonian Learner Corpus of English*. MA thesis. Department of English, University of Tartu, Tartu, Estonia.
- Undo, Aare. 2018. *Calculating the error percentage of an automated part-of-speech tagger when analyzing Estonian learner English: an empirical analysis*. MA thesis. Department of English, University of Tartu, Tartu, Estonia.
- Veldi, Enn. 2006. Particle Verbs in Bilingual Dictionaries: The Case of Estonian and English. *Euralex*, 73, 589-594.

APPENDICES

Appendix 1

Absolute frequencies of particles used in spoken and written language by Estonian EFL learners

Particle	Spoken	Written
about	0	1
along	2	2
apart	0	1
around	1	1
aside	1	0
away	4	9
back	15	2
by	1	3
down	5	2
in	8	7
off	2	1
on	15	15
out	24	20
over	1	14
through	3	1
together	1	5
under	0	2
up	36	30

Appendix 2

Absolute (N) and relative frequency (per 10,000 words) of particles in NNL Est and NL (Gilquin 2015)

Particle	NNL Est	NL
aboard	0.00 (N = 0)	0.00 (N = 0)
across	0.00 (N = 0)	0.40 (N = 11)
ahead	0.00 (N = 0)	0.33 (N = 9)
forth	0.00 (N = 0)	0.04 (N = 1)
forward	0.00 (N = 0)	0.95 (N = 26)
home	0.00 (N = 0)	n/a
round	0.00 (N = 0)	1.39 (N = 38)
about	0.22 (N = 1)	0.77 (N = 21)
apart	0.22 (N = 1)	0.18 (N = 5)
aside	0.22 (N = 1)	0.11 (N = 3)
around	0.44 (N = 2)	1.43 (N = 39)
under	0.44 (N = 2)	0.00 (N = 0)
off	0.65 (N = 3)	4.02 (N = 110)
along	0.87 (N = 4)	0.40 (N = 11)
by	0.87 (N = 4)	0.15 (N = 4)
through	0.87 (N = 4)	0.66 (N = 18)
together	1.31 (N = 6)	1.10 (N = 30)
down	1.52 (N = 7)	4.75 (N = 130)
away	2.83 (N = 13)	3.91 (N = 107)
in	3.26 (N = 15)	4.53 (N = 124)
over	3.26 (N = 15)	1.79 (N = 49)
back	3.70 (N = 17)	8.77 (N = 240)
on	6.53 (N = 30)	6.91 (N = 189)
out	9.57 (N = 44)	13.38 (N = 366)
up	14.36 (N = 66)	14.59 (N = 399)

RESÜMEE

TARTU ÜLIKOOL
ANGLISTIKA OSAKOND

Sandra-Leele Toom

The Use of Phrasal Verbs by Estonian EFL Learners: a Corpus-based Study. Eesti inglise keelt võõrkeelena (EFL) õppijate fraasiverbide kasutus: korpusuuring.

bakalaureusetöö

2020

Lehekülgede arv: 39

Annotatsioon:

Käesolev bakalaureusetöö uurib Eesti inglise keelt võõrkeelena (EFL) õppijate fraasiverbide kasutust korpusuuringu meetodil. Töö eesmärk on kindlaks teha, kui palju kasutavad Eesti EFL õppijad fraasiverbe nii üldiselt kokkuvõttes kui ka kõnekeeles ja kirjalikult eraldi. Lisaks püütakse kindlaks teha, millised on nende enim kasutatud fraasiverbid ja kuidas kasutavad eestlased fraasiverbe võrreldes inglise keelt emakeelena kõnelejatega. Töö replikeerib Gilquini (2015) uuringu meetodit, mille järgi otsitakse õppijakorpustest fraasiverbe ning arvutatakse absoluut- ja suhtelised sagedused. Leitud tulemusi võrreldi Gilquini (2015) uuringus esitatud tulemustega.

Töö jaguneb kahte peatükki: kirjanduse ülevaade ja empiiriline analüüs. Kirjanduse ülevaates defineeritakse fraasiverbid, samuti eesti keele fraasiverbid. Lisaks selgitatakse õppijakeelt ja -korpuseid ning tehakse ülevaade varasematest Eesti EFL õppijate kohta tehtud uuringutest. Peatüki lõpus tutvustatakse, kuidas Eesti EFL õppijad fraasiverbe tavaliselt kasutavad, varem tehtud samateemalisi uuringuid ning LINDSEI projekti. Empiirilise analüüsi peatükk tutvustab lähemalt analüüsis kasutatud korpuseid (TCELE ja LINDSEI-EST), metodoloogiat ning tulemusi. Peatükk lõpeb tulemuste aruteluga.

Töös analüüsitud 25 partiklit andsid 1533 vastust, millest 235-s esines fraasiverb. Kõige rohkem kasutasid eesti EFL õppijad partiklit *up* (üles) ja kõige vähem *apart* (lahus/eraldi). Fraasiverbe kasutati kirja- ja kõnekeeles pea võrdselt, kuid mõnevõrra rohkem kõnes (119 vs 116 fraasiverbi). Võrreldes inglise keelt emakeelena kõnelejatega, kasutasid Eesti EFL õppijad fraasiverbe vähem, eriti kõnekeeles. Inglise keelt emakeelena kõnelejate fraasiverbide kasutamise suhtelised sagedused olid kirja- ja kõnekeeles vastavalt umbes 50 ja 100 sõna 10 000 kohta, ning Eesti EFL õppijate sagedused 47 ja 56 10 000 sõna kohta. Siit ilmneb, et üldiselt on Eesti EFL õppijatel tendents fraasiverbe kasutada liiga vähe.

Märksõnad:

Inglise keel ja keeleteadus, grammatika, õppijakeel, fraasiverbid, korpusuuring.

Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Sandra-Leele Toom,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

The Use of Phrasal Verbs by Estonian EFL Learners: a Corpus-based Study,

mille juhendaja on Jane Klavan,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Digiallkirjastatud
Sandra-Leele Toom

Tartus, 26.05.2020

Autorsuse kinnitus

Kinnitan, et olen koostanud käesoleva bakalaureusetöö ise ning toonud korrektselt välja teiste autorite panuse. Töö on koostatud lähtudes Tartu Ülikooli maailma keelte ja kultuuride kolledži anglistika osakonna bakalaureusetöö nõuetest ning on kooskõlas heade akadeemiliste tavadega.

Digiallkirjastatud
Sandra-Leele Toom

Tartus, 26.05.2020

Lõputöö on lubatud kaitsmisele.

Jane Klavan

Tartus, 26.05.2020