

**UNIVERSITY OF TARTU
DEPARTMENT OF ENGLISH STUDIES**

**THE OVERUSE OF MAKE
IN ESTONIAN LEARNER
LANGUAGE**

BA thesis

**KADI TALPSEPP
SUPERVISOR: JANE KLAVAN, PhD**

**TARTU
2021**

ABSTRACT

High frequency verbs in learner language have received considerable attention in previous research, but little research has been done about the Estonian EFL learners and comparing the latter to native speakers' speech. This thesis investigates the use of the high frequency verb *make* by Estonian EFL speakers and native speakers of English. The motivation for writing this paper is due to little research done about the Estonian EFL learners and comparing the latter to native speakers' speech. The research questions that this paper aims to answer are a) how much is *make* used by EFL speakers compared to native speakers, b) in which situations does the overuse of *make* take place most often, c) how is the use of *make* different between EFL learners and native speakers. To find answers to these questions, a corpus study has been conducted.

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, high-frequency verbs, learner language and the learner corpora of the LINDSEI project are defined and introduced. The empirical analysis 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	8
1.1 High-frequency verbs	9
1.2 Learner Language	12
1.3 The LINDSEI Project	13
2. EMPIRICAL ANALYSIS	15
2.1 Methodology	16
2.1.1 The corpus data	17
2.1.2 MAKE categories	20
2.2 Results of the analysis	22
2.3 Discussion	27
CONCLUSION	30
REFERENCES	32
RESÜMEE	34

LIST OF ABBREVIATIONS

CECL – Centre for English Corpus Linguistics

CEFR – Common European Framework of Reference for Languages

EFL - English as a Foreign Language

ICLE – International Corpus of LearnerEnglish

LINDSEI – Louvain International Database of Spoken English Interlanguage

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

LOCNEC – Louvain Corpus of Native English Conversation

LOCNESS - LouvainCorpus of Native English Essays

L1 – First Language

L2 – Second Language

MICASE- Michigan Corpus of Academic Spoken English

TCELE – Tartu Corpus of Estonian Learner English

INTRODUCTION

Avoiding the overuse of high-frequency verbs can be a difficult task for anyone speaking a foreign language. In particular, it is a real challenge for students studying English as a foreign language (EFL). Hugon (2008) has said that high-frequency verbs can be difficult elements for L2 (second language) learners due to their wide variety of uses and meanings. One of the reasons for this issue could be that learning the meanings of high-frequency verbs is very often overlooked (Hugon 2008). At the same time, we should not assume that native speakers using English as their first language (L1) have little or no issues with high-frequency verbs.

Several studies have been conducted on high-frequency verbs; two of the most significant are by Nesselhauf (2004), who researched the use of verb constructions regarding *make*, *have*, *give* and *take* among advanced German EFL learners. Also, Gouverneur, Meunier and Granger (2008) analysed the use of *make* and *take* in EFL textbooks. Both of these studies were conducted using corpora, yet researchers consistently agree that not enough research has been done in this field. This could be the case because more comprehensive studies requiring extensive amounts of data are only possible when using computer-aided software, which is a relatively new development.

This thesis focuses on analysing the native and Estonian speakers' use and frequency of the English verb *make*. This topic has been researched several times, with numerous EFL students' use of English already analysed. However, no research has been done studying the uses of *make* by Estonian EFL learners. Studying the use of different learners of English is important because learners' mother tongue may influence their use of English. For example, Estonian EFL learners may have difficulties with aspects of English, such as using articles and phrasal verbs as they are different or non-existent in the Estonian language. The verb *make* is also used in several phrasal verbs, something future research

may potentially look at.

One of the most significant previous studies in the field of high-frequency verbs is Altenberg and Granger's (2001) article *The Grammatical and Lexical Patterning of MAKE in Native and Non-Native Student Writing*. In this study, the authors researched the overuse of the verb *make* among Swedish and French EFL learners and later compared the findings to one another. This thesis intends to follow Altenberg and Granger's (2001) methods and compose a separate study among Estonian EFL students and compare the results to native students' use of *make*.

The aim of this thesis is to answer the following research questions: a) how much is *make* used by EFL speakers compared to native speakers, b) in which situations does the overuse of *make* take place most often, c) how is the use of *make* different between EFL learners and native students. To answer these questions, a corpus study was carried out using EFL students' interviews from the LINDSEI-EST (*The Estonian subcorpus of the Louvain International Database of Spoken English Interlanguage*) and native students' interviews from LOCNEC (*Louvain Corpus of Native English Conversation*). The data of all inflectional forms of *make* were collected, after which each use was sorted into a usage category – these categories were compiled by following Altenberg and Granger's (2001) example. The methods and process of how the study was conducted are explained in more detail in the section covering the empirical analysis of the thesis.

The thesis is divided into two main sections, the first is the literature review, and the second is the empirical analysis. The first section begins with a summary and examples of international studies conducted regarding high-frequency verbs. A subsection of the word *make* specifically goes over Altenberg and Granger's (2001) study in more detail. This is followed by an overview of the LINDSEI project (*Louvain International Database of Spoken English Interlanguage*), which introduces the corpus from which the data of this

study is from. Next, the term learner language is defined and explained in the context of this thesis after which some similar studies from the Department of English Studies at the University of Tartu that have researched some other EFL aspects are mentioned.

The empirical analysis first gives an overview of the methodology that was used to conduct the study. The section following this is about the corpus study compiled by the author and a discussion about the categories of *make* and patterning in which the verb was used most often. Finally, the results of the corpus study are analysed. The chapter ends with a discussion about the findings.

1. LITERATURE REVIEW

How do EFL learners apply their knowledge of English in expressing themselves? Are high-frequency verbs widely overused amongst EFL learners? This compelling topic has been investigated by linguists for years and as a result, extensive research has already been completed by Källkvist (1999), Altenberg and Granger (2001) as well as Hasselgren (1994). This section will give an overview of the most important studies conducted that are compatible with the context of this thesis. Although there is research done focusing thoroughly and specifically on one or two language background groups and comparing the findings to one another, no broader investigation has been done that would compare EFL learners' use of English on a wider spectrum, for example, between different Indo-European language groups. As Altenberg and Granger (2001:174) have pointed out, high-frequency verbs have "high-frequency equivalents in most languages".

The studies that have been conducted in this field include Altenberg and Granger's (2001) article, from which the methodology and empirical analysis of this thesis has been followed. Also, Källkvist (1999) analysed Swedish learners' overuse of verbs in the paper *Form-Class and Task-Type Effects in Learner English. A Study of Advanced Swedish Learners*. Another example is Hasselgren's study with Norwegian EFL learners *Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary* (1994). Whilst comparing these studies, different authors have arrived at different conclusions in their research. For example, in their paper, Altenberg and Granger (2001:174) have not agreed with the statement that EFL overuse tendencies may be reinforced among Germanic mother-tongue students, while Källkvist (1999) indicated this may be the case. To contribute to this discussion, the comparison of Germanic mother-tongue students and Estonian students both speaking English could be an interesting study to pursue in further research. This thesis intends to explore the use of the

word *make* by the Estonian EFL learners and native speakers' of English. One of the aspects of interest in this study is to compare which students over- or underused the verb similarly to the other studies conducted about Swedish and Norwegian EFL learners.

1.1 High-frequency verbs

High-frequency verbs are very often a complex matter for EFL learners. They are taught very early on and are vital for having fluent conversations (Hugon 2008). Their seemingly endless variations of meanings and nuances make them harder to comprehend, even for advanced L2 speakers. Hugon (2008:5) has stressed that one of the reasons for this issue could be because “these verbs are then neglected on the assumption that they are known”. The reason for this could be due to EFL teachers who may mistake students' insufficient knowledge to be adequate enough to move on to more advanced topics during the learning process.

More often than not, even advanced learners are turning to easier and better-known verbs. Nesselhauf (2004) has noted that students tend to misuse verbs to a great extent. In a similar research as this thesis, Hasselgren (1994) studied the use of English verbs by Norwegian students. Whilst dealing with the same issue of students overusing high-frequency verbs, Hasselgren (1994) termed this to be the ‘teddy-bear effect’ when Norwegian students would clutch onto their most familiar English verb that felt the safest. Another factor is that high-frequency verbs appear very often in phrasal verbs (Hugon 2008). As a result of the ‘teddy-bear effect’, EFL students usually avoid these as well. Sinclair (1991:79) has discussed that this produces more occasions where students will make mistakes and sound clumsier because they try to use rarer words and avoid idiomatic phrases. However, this statement by Sinclair (1991) has had researchers disagree on the matter as Altenberg and Granger (2001:174) point out that Sinclair (1991) “does not back

up his statement with corpus data”.

The basic high-frequency verbs in EFL listed by Altenberg and Granger (2001:173) can be seen in Table 1; these are lexemes that disregard *be* as well as modal auxiliaries and come up the most in many corpus-based lists. Some other research, which has been conducted about the use of high-frequency verbs shown in Table 1 among EFL learners also include research by Nesselhauf (2004) and Gouverneur, Meunier, and Granger (2008). The latter researched the use of high-frequency verbs *make* and *take* in a series of EFL textbooks. They found significant inconsistency in how the phraseological patterns of *make* and *take* were treated (Gouverneur, Meunier and Granger 2008). Meanwhile, Nesselhauf (2004) conducted a study based on a subcorpus of ICLE which investigated advanced German EFL learners’ use of verb constructions, with *make* being one of them (2004:112). However, Nesselhauf (2004) also included *have*, *take*, and *give* in their study. The aim of Nesselhauf’s study was to look at support verb constructions and how to contribute to language teaching via learner corpus analysis (Nesselhauf 2004).

Table 1. The high-frequency verbs of EFL (Altenberg and Granger 2001)

<i>HAVE</i>	<i>DO</i>	<i>KNOW</i>	<i>THINK</i>	<i>GET</i>
<i>GO</i>	<i>SAY</i>	<i>SEE</i>	<i>COME</i>	<i>MAKE</i>
<i>TAKE</i>	<i>LOOK</i>	<i>GIVE</i>	<i>FIND</i>	<i>USE</i>

Hugon (2008: 8) has stated that *make* can be classified as one of the most frequent verbs in the English language. This makes the verb *make* an engaging subject to analyse in various non-native English speakers as well as among L1 speakers. This was done in Altenberg and Granger’s (2001: 173) study which aimed to answer the research questions “do learners tend to over- or underuse these verbs? Are high-frequency verbs error-prone or safe? What part does transfer play in misuse of these verbs?”. The methodology used to

investigate the topic includes two corpus samples from the ICLE database, both of these samples were around the same size of around 170,000 words, but one illustrates French EFL learners and the other shows Swedish EFL learners' use of English (Altenberg and Granger, 2001: 175). After compiling the data from both non-English speaking groups, the findings were then compared to native English speakers' samples from LOCNESS (*Louvain Corpus of Native English Essays*) in which there was the same amount of words as in the previous data (Altenberg and Granger 2001: 175).

As mentioned before, Altenberg and Granger (2001) studied the use of the verb *make* in three language groups. They divided the major uses of the verb into eight main categories: delexical use, causative use, produce something, earn (money), link verb uses, idiomatic, phrasal/prepositional uses or other uses (Altenberg and Granger 2001: 177). These were greatly influenced by the *WordSmith Tools*' sorting facility and by determining these categories the sorting process was a lot more straightforward and organised (Altenberg and Granger 2001: 177). Their findings of their study show that the most common category of *make* was the causative use, and the second being the delexical use of the verb (2001:177). The article was concluded by saying that even proficient L2 users struggle with overusing *make*. Main mistakes could be found in the delexical and causative structures where even if at first the use appears safe to use then it is still somewhat problematic (2001: 189-190). Another interesting finding is that, at times, the results show that different nationalities can end up making the same mistakes while speaking in English. Other findings suggest that mistakenly overusing *make* may have also been influenced by the learner's L1 peculiarities such was the case in Altenberg and Granger's (2001: 173) paper where both English and Swedish have dominant high-frequency verbs with dominant patterning (*make/göra*), however they have different, often lexically specific, alternative constructions. For example a constructions "*make it (im)possible (for sb)* and *make it*

easy/easier (for sb) to both have equivalents in Swedish and both can be replaced by causative verb constructions in English (for example, *enable sb to, prevent sb from, facilitate sth*).” However, overuse is persistent in overusing *make* in this structure (Altenberg and Granger 2001: 182).

1.2 Learner Language

In the context of Estonian EFL learners, defining learner language is unproblematic and easily understood. Specifically, Granger (2008: 260) has said that learners of a language whose language is explored by learner corpora can be classified as foreign language learners. This means that learners are acquiring a language that is not their mother tongue, nor can we call this an institutionalised additional language of their home country (Granger 2008: 260).

Since English is not considered a second language in Estonia but that of a foreign language, then, in light of this thesis, learner language is not considered a complex aspect. Like Granger (2008: 260) has noted, this definition is not problematic in most countries like Spain, Sweden, and China because English does not hold an official status in those countries. Whereas, in countries like India, Nigeria or the Philippines, English has had a historical influence and has achieved a notable status in education and administration. Thus, a nativised version of English is used even to this day (Granger 2008: 260).

It is crucial to study learner language as this will help improve the general knowledge as to what is expected from teachers, pass on valuable language skills to learners and decrease the incomprehension even advanced language learners have made in their speech and writing, especially when corpus studies are conducted. Various studies regarding learner language have been conducted at the Department of English Studies at the University of Tartu. One example is a study of the use of formulaic language by native

and non-native speakers (Piiri 2020), which studied whether or not non-native speakers used formulaic language in their writing or speech and also comparing these uses with native speakers. This study was conducted by a corpus-based analysis using data from TCELE (*Tartu Corpus of Estonian Learner English*), LINDSEI-EST and MICASE (*Michigan Corpus of Academic Spoken English*) (Piiri 2020). The study found interesting observations about formulaic clusters in spoken language such as *I think* and discussed the possible reasoning of this use in learner language (Piiri 2020).

Another interesting study that should be mentioned in the context of the present thesis is a corpus-based study of the use of phrasal verbs in Estonian EFL learners by Toom (2020). Similarly to this thesis, Toom (2020) replicated a study by another author by using their methodology and analysed the frequency of the use of phrasal verbs in Estonian EFL learners and later compared the results to native speakers' use of phrasal verbs by carrying out a corpus-based study. Toom (2020:28) found that Estonian EFL learners used phrasal verbs less frequently compared to native speakers, they noted that this may be due to the fact that in Estonian, there are no phrasal verb constructions. These studies have broadened the knowledge about EFL learners' performance in using their acquired language.

1.3 The LINDSEI Project

The data used to conduct the empirical analysis of this thesis is from the LINDSEI (*Louvain International Database of Spoken English Interlanguage*) Estonian subcorpus LINDSEI-EST which is currently being compiled at the Department of English Studies at the University of Tartu (CECL, n. d. (b)). For native speakers, however, LOCNEC (*Louvain Corpus of Native English Conversation*) was used for collecting the necessary data. LINDSEI was launched as a project in 1995 by CECL (*the Centre for English Corpus*

Linguistics). The motivation for this development was to contribute to the advanced EFL learners' oral language analogue to complement the ICLE (*International Corpus of Learner English*) (CECLn.d.(a): para. 1).

All components of the corpus follow the same structure. Each component has around 50 interviews that include three tasks: set topic, free discussion and picture description. Each interview is linked to the profile of the person that was interviewed, the interviewer and the interview itself. Comprising all of this information together makes it easier and possible to study different aspects of learner language (CECLn.d.(a): para. 2).

In addition to the non-native data, a comparable corpus is also available for interviews by native speakers of English. This is compiled into LOCNEC (*Louvain Corpus of Native English Conversation*) which we can classify as the native speaker counterpart. Due to the same structure of the interviews, it is possible to compare the use of language and among other things, this has helped identify universal L1-specific oral interlanguage (CECLn.d.(a): para. 3).

As of today, numerous international partners have made their contribution to the project, and the irrespective subcorpora have been completed. These include EFL learners who are native speakers of Finnish, French, German, Italian, Spanish and many others. Meanwhile, the compilation of the subcorpora of Arabic (Saudi Arabia), Arabic (Lebanon), Croatian, Iranian and Estonian students are still in progress (CECL, n. d. (b)).

The corpus of the LINDSEI project has helped tailgate various new studies at the Department of English Studies at the University of Tartu such as Toom (2020) and Piiri (2020), whose works are discussed in more detail in the learner language section of this thesis. Numerous further studies have used the LINDSEI data to further inspect the lexis, syntax, phraseology, discourse, and pragmatic aspects of learner English (CECL, n. d. (a): para. 4).

Some noteworthy studies that have been conducted using the LINDSEI corpus data include Aijmer's (2011) research about the use of frequent pragmatic markers such as *well* in advanced EFL learners. Aijmer (2011) studied the use of English by Swedish EFL learners and, similarly to this study, compared their findings to native speakers. The conclusion of the study was that that EFL learners overuse *well* more often than native speakers did. As this study also aims to find out which circumstances most often lead to the overuse of *make* in learner language, then Aijmer's (2011) study is more relevant in the empirical analysis section of this thesis.

Another useful example is a study from the LINDSEI-CZ subcorpus by Gráf (2017) who researched inaccuracies in the use of verbal categories in advanced Czech EFL learners' spoken English. It was found that even advanced EFL learners make occasional or regular slips in the usage of verbs and tenses (Gráf 2017:145). Gráf notes that speculation as to why this was the case may be induced by teaching style and intensely repetitive language practice (2017: 146). This conclusion may point towards a more general explanation as to why EFL learners make mistakes even at advanced levels.

2. EMPIRICAL ANALYSIS

The empirical analysis aims to answer the research questions of this thesis by analysing corpus data compiled from the LINDSEI-EST subcorpus which represents Estonian EFL learner data and LOCNEC for native English speakers' learner data. The section begins by introducing the methodology and how the corpus data for this thesis was compiled. This is followed by an overview of the corpus data and the most noteworthy findings between EFL and native speakers' interview data. As well as demonstrating the most frequent categories into which the uses of *make* were distributed, the final step of the analysis goes over the observations about grammatical patterns and draws conclusions

about the research questions.

2.1 Methodology

The aim of the empirical analysis is to answer the research questions of the thesis: a) how much is *make* used by EFL speakers compared to native speakers, b) in which situations does the overuse of *make* take place most often, c) how is the use of *make* different between EFL learners and native students. To answer these questions, a corpus study was carried out by following the example of Altenberg and Granger's (2001) study. This was done because their findings and methodology provide an excellent base for further studies. While Altenberg and Granger's (2001) study focused on French and Swedish students then this thesis focuses on Estonian and native students' use of English.

The present study was conducted by using the Estonian students' interviews from the LINDSEI-EST subcorpus and native speakers' interviews from LOCNEC. All phrases and sentences that contained the use of *make* and its inflected forms were extracted from the two corpora and compiled into a spreadsheet. Firstly, this made it possible to calculate the frequencies of *make* and determine the frequency of each inflectional form. Secondly, having all the phrases that contained some form of *make* on one page made it easier to begin with the grammatical analysis. This was done by taking the categories used in Altenberg and Granger's (2001) study which can be seen in Table 4 (column 1) and matching the phrases with the most suitable category.

The analysis of the corpus in this study was done manually and the grammatical patterning, examples of which can be seen in Table 4 (column 3), were determined by the author. However, while Altenberg and Granger's (2001) study was also conducted manually, then in the fifth section of their article, their findings were compared to computer-aided software called *WordSmith Tools'* Collocate Display (2001: 185). In the

case of this thesis, the use of computer-aided software is not used, since *WordSmith Tools* is a programme that one needs to subscribe to.

2.1.1 The corpus data

The Estonian subcorpus LINDSEI-EST of the Louvain International Database of Spoken English Interlanguage (*LINDSEI*) is currently being compiled at the English Department of the University of Tartu. The aim of the ICLE is to investigate the interlanguage in EFL learners. The LINDSEI project allows researchers to determine the over- and underuse of linguistic patterns and find out whether these can be linked to language specific nuances or if they are universal. Additionally, these comparisons can also determine if these uses are influenced by factors in the learner's educational background or cultural implications (Pravec 2002: 83).

The corpus data for this study was compiled using data from the LINDSEI project. At its current stage, the LINDSEI-EST corpus consists of 25 interviews (33,842 words of transcribed text) recorded in 2018. All of the interviewees (18 female, 7 male; average age 23 years) were native speakers of Estonian. They were third or fourth year students of English language and literature at the University of Tartu. The LOCNEC corpus consists of 50 interviews (118,159 words of transcribed text) with native speakers of English, all of them undergraduate and graduate students at Lancaster University in the United Kingdom

Table 2. The frequency of *make* in LINDSEI-EST and LOCNEC

	EFL speakers (LINDSEI-EST)	Native speakers (LOCNEC)
Number of interviews	25	50
Number of words of transcribed text	33,842	118,159
Number of uses of <i>make</i> , <i>made</i> , <i>making</i> , <i>makes</i>	86 (254 uses per 100,000 words)	174 (147 uses per 100,000 words)

According to the LINDSEI guidelines, after obtaining informed consent regarding the metadata, the participants' profiles were composed. This included their age, gender, nationality, native language, education background, information about stays in an English-speaking country as well as any other foreign languages they speak.

In the interviews, the topics for both language backgrounds were similar and structured around three tasks. The first was a conversation that could be based on three topics: a visit to a country that has left a memorable impression, describing a meaningful film or play that the interviewee has seen or a valuable experience. This was followed by a free discussion and finally, a picture description task. Looking at the photos, the interviewees were asked to tell a story based on four pictures in which they could see an artist who was painting a woman. Without the possibility of taking notes, the interviewees described what they thought of the pictures and, in a fairly informal way, expressed their opinions and experiences as well. The interview lasted for around 15 minutes.

To see the overall proportion of how much the verb *make* was used amongst EFL speakers and native speakers, we can look at Table 2. In total, there were 25 EFL and 50 native interviews (see Table 2). The data collected from these interviews include 86 EFL and 174 native students' uses of *make*. These uses were first collected and compiled into a spreadsheet from where the inflectional form (*make, makes, made, making*) was determined. An overview of the inflectional forms used in the two corpora can be seen in Table 3. The EFL and native students' interviews were gathered in separate spreadsheets which enabled a more individual grasp of both groups. The first spreadsheet was for EFL speakers with data from LINDSEI-EST and the second spreadsheet was for native speakers with data from LOCNEC. In addition, by keeping the spreadsheets separate, the author also analysed whether any language-based patterns arise depending on whether the student was

native or an EFL speaker. The findings of this aspect will be discussed in the further sections of the empirical analysis.

In Table 2, we can see the total number of words of both spreadsheets and the total uses of all inflectional forms of *make*. As can be seen, when we look at the number of cases in which *make* was used, we can see that 174 occurrences of *make* in total were used in native speakers' interviews (normalized frequency of 147 uses per 100,000 words). However, when we look at EFL speakers' interviews, we can see that there were 86 occurrences in total (normalized frequency of 254 uses per 100,000 words). This means that EFL speakers used *make* more often than native speakers with an increase of about 100 more uses per 100,000 words. EFL speakers used *make* almost twice as often compared to native speakers. The results from Table 2 apply to the first research question: how much is *make* used by EFL speakers compared to native speakers? From the data in Table 2 we can establish that when comparing native and Estonian EFL speakers from LINDSEI-EST and LOCNEC, native speakers used *make* far less often compared to EFL speakers. To see exactly which inflectional forms were used more or less often we can look at Table 3.

Table 3. Overview of the uses of *make*, *makes*, *making*, *made* by the EFL and native speakers.

	LINDSEI-EST	LOCNEC
<i>make</i>	35 (42%)	64 (37%)
<i>makes</i>	8 (9%)	31 (18%)
<i>making</i>	10 (11%)	25 (14%)
<i>made</i>	33 (38%)	54 (31%)
Total uses	86 (100%)	174 (100%)

In Table 3, we can see the proportions in which *make*, *makes*, *making* and *made* were used for both language groups. We can see that EFL speakers used the verbs *make*

5% and *made* 7% more often than the native speakers. However, the verbs *makes* and *making* were used more by native speakers. While using *making* was mostly uncommon for both language groups with only a 3% increase of use in native speakers then *makes* was used twice as often by natives as it was used by EFL speakers. However, when we analyse the proportions within the groups separately, then we can still see some similar trends with the inflectional form *make* being the most common usage for both EFL and native speakers. Next comes *made* which also ranks the same place for both groups with around 31-38% of uses. Whereas with the last two inflections, the ranking is different. If we were to combine the uses of *makes* and *making* then they would make up 20% of uses among EFL speakers. However, this is much lower compared to the same calculation with native speakers which would result in 32%. From this we can claim that while EFL speakers' use of *make* is more frequent as we saw in Table 2, then Table 3 suggests that though native speakers use *make* less often, they cover a wider range on inflectional forms, thus being more diverse with their usage of *make*.

2.1.2 MAKE categories

The word *make* is used in numerous situations and can thus be categorised in many grammatically distinct ways. This is why *make* can be classified as a high-frequency verb (Svartvik and Ekedahl 1995). As the meanings and context of each use of *make* may differ, the uses of the verb need to be analysed and common categories or groups defined. For this reason, the uses of *make* have been divided into eight most commonly used categories to better analyse the meaning and use of the verb. These categories are delexical use, causative use, produce something, earn (money), link verb uses, idiomatic, phrasal/prepositional uses or other uses (Altenberg and Granger 2001: 177). Examples of some common uses of the categories can be seen in Table 4. This includes the example

given by Altenberg and Granger (2001) (column 2) as well as an example from the data used for this thesis (column 3).

Table 4. Grammatical categories and examples (Altenberg and Granger 2001:177).

Category	Example from Altenberg and Granger (2001:177)	Example from LINDSEI-EST and LOCNEC
1. Produce something (a result of creation)	<i>Make furniture, make a hole, make a law</i>	Make breakfast
2. Delexical use	<i>Make a distinction/ a decision/ a reform</i>	Make a decision
3. Causative uses	<i>Make somebody believe something, make something possible</i>	Make you have complexes
4. Earn (money)	<i>Make a fortune</i>	Making my own money
5. Link verb uses	<i>She will make a good teacher</i>	He was made older
6. Idiomatic	<i>If we run, we should make it</i>	It's difficult to make it
7. Phrasal/ prepositional uses	<i>Make out, make up, make out of</i>	Makes up a false impression
8. Other	<i>Make good, make one's way</i>	Match making

As can be seen from Table 4, the eight categories cover a lot of different meanings. The first category – produce something – is used when the term *make* is used to describe physically making or producing a product, usually something that can be seen, touched or read. The delexical use is used when additional information can be added to the action being performed by using adjectives and nouns instead of adverbs. Usually delexical uses take place with talking, sounds and making plans. The causative uses are used when someone causes somebody or something to do something. In this case, all causative uses were active causatives and not passive causatives. To earn (money) was used to describe

cases where *make* was used to describe earning money. However, this category had such a low amount of uses that the nuances of this category will not be discussed further. Linking verb uses were present when there is no action involved. These uses are linking or joining between the subject of the sentence and the subject complement. Idiomatic uses are colloquial uses oftentimes comfortable to use for the native speaker. Phrasal and prepositional uses include phrases where the use of prepositions with *make* creates a new meaning. The last category listed as “other” mostly includes cases where the interviewees stammer in their speech or misuse the word *make*.

2.2 Results of the analysis

The data from LINDSEI and LOCNEC used in this thesis was carefully sorted into each category. During the sorting, the author made notes of some patterns that emerged within some of the categories. Without using any computer sorting facilities, this process was time consuming because by making notes of the patterns, the list was renewed many times whilst going over the data. An example of this could be the sorting of the phrase “make a mistake” quickly into the delexical use. However, a phrase like “she did not realize she had made this complete mistake” took longer to be distributed into the correct category. The complete overview of the frequency each category was used by EFL and native speakers can be seen in Table 5. Due to there being twice as many interviews from native speakers, then the uses of *make* were converted into percentages which make the table easier to read and the real frequency of each category is displayed.

In Table 5, we can see the total uses of *make* in each category. At first glance, we see that Categories 4 and 5 (money and linking uses) both had very few uses in both groups of students, for this reason they will not be included in further discussions. With those categories being disregarded, we can see that native speakers’ use of other categories

is fairly well distributed with Categories 1, 2, 3 and 7 (produce something, delexical uses, causative uses and phrasal uses) with the first three in close range within one another.

Table 5. Uses of MAKE by EFL and native speakers

Category	LINDSEI-EST	LOCNEC
1. Produce	20 (24%)	34 (21%)
2. Delexical	22 (26%)	48 (28%)
3. Causative	16 (19%)	35 (21%)
4. Money	–	7 (4%)
5. Link uses	2 (2%)	2 (1%)
6. Idiomatic	3 (3%)	11 (6%)
7. Phrasal	14 (16%)	24 (12%)
8. Other	8 (10%)	13 (7%)
Total	86 (100%)	174 (100%)

For EFL speaker's the table is also focused primarily on the top three categories. Though there are some differences in the frequency of usages in Category 6 and 7 (idiomatic uses and phrasal uses). The author has decided to compare and discuss the top three most common categories for both EFL and native speakers which were to produce something, delexical uses and causative uses. For this, tables 6-8 show each collocation used which are listed in a decreasing order.

Table 6. Collocates of produce something MAKE in LINDSEI and LOCNEC

LINDSEI-EST	24%	LOCNEC	21%
Movie(s)	7	Friends	13
Things	4	Notes	5
Film(s)	2	Film(s)	3
Breakfast	1	Fries	1

Sentences	1	Smile	1
Bomb	1	Drafts	1
Net	1	Pictures	1
References	1	Colours	1
Room	1	Tea	1
Something	1	Classics	1
Sculptures	1	Comedy	1
Story	1	Matches	1
		Pharmaceuticals	1
		Bottles	1
		Speech	1
		Cigar	1

Category 1 – produce something collocates are shown in Table 6. The most frequent uses for EFL speakers from the LINDSEI data was to “make (a) movie(s)”, “make things” and “make (a) films(s)”. When looking at the native speakers’ data from LOCNEC, the most frequent uses were to “make notes” and “make films”. The latter is fairly similar to the EFL speakers’ use of make for “films” and “movies”, where the interviewees described a meaningful film that they had seen. Shown below are some typical instances where the author would suggest replacing the verb *make* with a different verb or rephrase the phrase entirely to avoid clumsy expressions or errors.

1. I try to *make* the sentences (LINDSEI-EST)
(correct form: *write*)
2. *Make* all kinds of things with iron (LINDSEI-EST)
(correct form: *create*)
3. They *made* the story afterwards. The book version (LINDSEI-EST)
(correct form: *wrote*)
4. I had to *make* notes when I went to see a play (LOCNEC)

(correct form: *take*)**Table 7.** Collocates of delexical MAKE in LINDSEI and LOCNEC

LINDSEI-EST	(26%)	LOCNEC	(28%)
Sb/it look	5	Sb/it look	12
Sb/sth pretty	2	Sb/sth pretty	5
Point	2	Sb beautiful	4
Changes	2	Connection	3
Mistake	2	Difference	3
Comments	1	Effort	2
Rendition	1	Decision	2
Sense	1	Sb flattering	2
Comparison	1	Realize	1
Idolized	1	Relevant	1
Claim	1	Noise	1
Adjustments	1	Different	1
Easier	1	Plans	1
Decision	1	Career	1
Something interesting	1	Professional	1
Connection	1	Alternations	1
		Easier	1
		Joke	1
		Mistake	1
		Sb perfect	1
		Sth wonderful	1
		Sb ugly	1
		Judgement	1
		Changes	1
		Clear	1
		Exception	1
		Choice	1
		Racket	1

Category 2 – the delexical uses, had a fairly even amount of uses according to Table 5, with EFL speakers using delexical uses 26% and natives 28% of the time. This difference of only 2% is seemingly small, however, Table 7 demonstrates how native speakers used *make* with delexical uses with a much wider vocabulary. The most frequent uses for both EFL and native speakers was using phrases like “make sb look” and “make her pretty/beautiful”. This was mainly for the task when interviewees were asked to describe a picture. A reason for quite few findings of mistakes or situations where a different verb would be more appropriate to use may be due to tendencies of underusing the delexical use of the verb which may be linked to Hasselgren (1994) theory of the teddy-bear effect.

5. They do try and *make* it as professional as possible... (LOCNEC)

(correct form: *stay*)

6. what is the reality of life but *make* your but *make* it idolized (LINDSEI-EST)

(correct form: *keep*)

Table 8. Collocates of causative MAKE in LINDSEI and LOCNEC

LINDSEI-EST	19%	LOCNEC	21%
Sb look pretty	5	Sb/smith look/be smth	9
Sb feel	2	Sb do smth	6
Sb think	1	Sb feel	4
Sb/sth Look crazy	1	Sb hair curl	2
Sb more open	1	Sb cry	2
Sb look like smth	1	Sb happy	2
Sth realistic	1	Sb paint	2
(agent) look better	1	Sb determined	2
Sb do smth	1	Sb believe	1
Sb sad	1	Sb have complexes	1
		Sb laugh	1

Category 3 – the causative uses was used by native speakers in 21% of cases (Table 5) and EFL speakers in 19% of the time. Similarly to the delexical uses, the causative use was used by natives 2% more often. The most notable uses for native speakers were different variations of “to *make* somebody look like something” or “somebody doing/feeling something”. With causative uses, the patterning of uses came apparent the most. With EFL speakers, the most common uses included “somebody look pretty” and “somebody feel”. These may be very similar to the delexical uses, however the collocates of the causative uses included an agent which allowed/forced/made the action take place. Some typical errors or suggestions for better wording are as follows:

7. wants him to *make* the portrait pert= prettier (LINDSEI-EST)

(correct form: *paint the portrait*)

Along with Category 1, the causative uses had numerous instances when we could see the overuse of the verb *make*. This answers the second research question which aimed to find out which circumstances were present in the event of overusing *make*. EFL speakers clearly misused or overused the verb in Category 1.

2.3 Discussion

The current paper aimed to investigate the overuse of *make* in learner language, the findings of which can be seen in the empirical analysis where a corpus-based study was conducted. The first research question aimed to find out: how much is *make* used by EFL speakers compared to native speakers? This could be seen in Table 2, which showed the frequency of *make* by calculating the uses of *make* from both data sheets and comparing them to the number of words of transcribed text within the interviews. The results of this thesis confirm the results of previous studies that EFL speakers used *make* more often or had issues with using *make*. This outcome is in accordance with Altenberg and Granger’s (2001: 189) study and their results which suggested that advanced proficiency level EFL

learners have difficulties with high-frequency verbs. Among Swedish and French EFL students Altenberg and Granger (2001: 178-179) noticed that learners underused and misused delexical structures. However, regarding the causative uses of *make*, they noticed that Swedish EFL students overused these structures whereas French EFL students underused the causative *make* (Altenberg and Granger 2001: 181).

The second research question aimed to find out: in which situations does the overuse of *make* take place most often? This could be seen in Table 5, where for both EFL and native speakers, the most common uses of *make* were within the first three categories. However, after further analyzing the categories produce something, delexical uses and causative uses, the most frequent times we could see examples of overuse along with suggestions or corrections were in Category 1 and 3 – produce something and causative uses. In terms of the categorization of *make* this question is partly answered as the categories in which overuse took place were partly identified. However, further studies as to whether or not the use of *make* in Category 1 or 3 could be influenced by the speakers' first language are needed.

The final research question: how is the use of *make* different between EFL learners and native students? This thesis cannot provide a clear answer to this question. It should be considered that for native speakers, the use of high-frequency verbs will most likely stem from implicit knowledge of grammar. While for EFL speakers, the use and meanings of high-frequency verbs are learnt later in life and regardless of their CEFR (*Common European Framework of Reference for Languages*) language level, which is in this case around C1-C2, the use is still explicit knowledge and can be influenced by the speaker's native language which was also the case in the misuses seen in this thesis. Thus, the answer to the question as to which properties specifically contribute to the different usage patterns between the use of *make* between EFL learners and native students is an opportunity to

look into while doing further research on the matter.

CONCLUSION

In everyday conversations, the use of *make* will often go unnoticed. Learners will not spend much time from the day analysing which use of *make* was necessary and which was overuse, nor think about what category their use of the verb *make* could fall under. However, the issue of the overuse of *make* is present as was made clear by the points discussed in the literature review of this thesis.

Even though the use of *make* has been investigated in numerous different studies, including Altenberg and Granger's (2001) study along with other high-frequency words like Aijmer (2011) studied the use of *like* then this issue had still not been investigated among Estonian EFL speakers. The thesis at hand intended to contribute to this wide field of research by conducting a corpus-based study that would investigate the overuse of *make* in learner language and then compare the findings to native speakers' use of *make* as well.

In order to achieve this, the author formed the research questions for the thesis, which were a) how much is *make* used by EFL speakers compared to native speakers, b) in which situations does the overuse of *make* take place most often, c) how is the use of *make* different between EFL learners and native students. By extracting data from 25 EFL interviews and 50 native interviews, the author compiled a corpus. To properly carry out the study, the author followed the example of Altenberg and Granger's (2001) methodology. By doing this, the same eight categories from their study were used to determine the use of *make* in this paper. The categories into which the use of *make* were divided were as follows: produce something, delexical use, causative use, earn money, link verb uses, idiomatic uses, phrasal/prepositional uses and other. This selection was done manually and carefully by the author. However, the study by Altenberg and Granger (2001) that was followed eventually used a computer-based software which was not the case for this paper. By using a computer, the pattern-base sorting was quicker and more

precise than that of only manual work.

By completing the corpus-based study and going over the grammatical patterning of uses and dividing uses into the most suitable category, the author was able to answer research question (a) and partly, (b). For the first research question, *make* was used by EFL speakers more often than with native speakers as was demonstrated in Table 2 where 86 of uses of the verb *make* were identified (254 uses per 100,000 words) in the LINDSEI-EST data. This number was more than 100 uses per 100,000 words higher than that of the native counterpart, whose frequency of *make* was 174 (147 uses per 100,000 words). There is a considerable difference between the relative frequency counts of the verb *make* in the two data sets. For the second research question the overuse of *make* took place most in Category 1 and 3, however the author believes the answer to this question is not complete as there may be other factors at play like the influence on one's L1 and L2 on their use of *make* and not just which category was used. This is a topic that can be studied for further research.

Overall, it is clear that this thesis serves as a starting point for future research on the topic of Estonian EFL speakers' use of the verb *make*. Several questions and further studies can be done on the basis of this paper, including the shortcomings as well as the findings of the research questions that were answered along with the question that did not get an answer.

REFERENCES

Primary sources

- Altenberg, Bengt and Granger, Sylviane. 2001. *The grammatical and lexical patterning of MAKE in native and non-native student writing*. *Applied Linguistics*, 22(2), 173-195.
- CECL, n. d. (a). LINDSEI, available at <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei.html>, accessed March 25, 2021
- CECL, n. d. (b). LINDSEI, available at <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei-partners.html>, accessed March 25, 2021

Secondary sources

- Aijmer, Karin. 2011. *Well I'm not sure I think... The use of well by non-native speakers*. *International Journal of Corpus Linguistics* 16.2: 231-254.
- Gráf, Tomáš. 2017. *Verb Errors in Advanced Spoken English*. *Acta Universitatis Carolinae Philologica* 1 131-149.
- Gouverneur, Céline, Meunier F., and Granger, Sylviane. 2008. *The phraseological patterns of high-frequency verbs in advanced English for general purposes*. *Phraseology in foreign language learning and teaching*, 223-246.
- 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.
- Hasselgren, Angela. 1994. *Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary*. *International journal of applied linguistics*, 4(2), 237-258.
- Hugon, Claire. 2008. *High-frequency verbs: Starting block or stumbling block for advanced L2 communication? Insights from native and*

- learnercorpora. *Taal aandenlijve: Hetgebruik van corpora in taalkundigonderzoekentaalonderwijs*, 69-98
- Källkvist, Marie. 1999. *Form-class and task-typeeffects in learnerEnglish: A study of advanced Swedish learners* (Vol. 95). Lund University Press.
- Nesselhauf, Nadja. 2004. *Howlearnercorpusanalysiscancontributetolanguageteaching: A study of support verb constructions*. *Corpora and languagelearners*, 17, 109.
- Piiri, Andreas. 2020. *A CorpusBasedStudy Of FormulaicLanguageUse By Native And Non-NativeSpeakers*. BA thesis. University of Tartu, Department of EnglishStudies.
- Pravec, Norma. A. 2002. *Survey of learnercorpora*. *ICAMEjournal* 26. 8-14.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford University Press.
- Svartvik, Jan, and OlofEkedahl. 1995. *Verbs in private and publicspeeking. The Verb in ContemporaryEnglish*. *Theory and Description* 273-289.
- Toom, Sandra-Leele. 2020. *The use of phrasalverbs by Estonian EFLlearners: a corpus-basedstudy*. BA thesis. Department of English, University of Tartu, Tartu, Estonia.

RESÜMEE

TARTU ÜLIKOOL
ANGLISTIKAOSAKOND

Kadi Talpsepp

The Overuse of MAKE in Estonian Learner Language. MAKE

Ülekasutus Eesti Õppijakeeles

bakalaureusetöö

2021

Lehekülgedearv: 34

Annotatsioon:

Käesoleva bakalaureusetöö eesmärk on uurida Eesti inglise keelt võõrkeelena (EFL) õppija sõna *make* kasutust korpusuuringu meetodil. Töö eesmärk oli uurida kui palju kasutavad Eesti EFL õppijad sõna *make* võrreldes inglise keelt emakeelena kõnelevate õppijatega. Sealjuures uuritakse milliseid grammatilisi kaegooriaid enim kasutatakse. Töö replikeerib Altenberg ja Grangeri (2001) uurimuse meetodit, kus uuriti *make* kasutust Prantsuse ja Rootsi EFL õppijate seas. Käesoleva töö tulemusi võrreldi Altenberg ja Grangeri tööga.

Töö jaguneb kahte sektsiooni: kirjanduse ülevaade ning empiiriline analüüs. Esimeses sektsioonis tutvustatakse varasemaid sarnaseid uuringuid EFL õppijakeele kohta, millised on sagedased verbid, õppijakeelt ning seejärel LINDSEI projekti. Empiirilise analüüsis tutvustatakse töös kasutatavat metodoloogiat ning tehakse ülevaade autori koostatud korpuse kohta LINDSEI-EST ning LOCNEC andmebaasidest. Sellele järgneb diskusioon, kus arutletakse tulemuste üle ning vastatakse uurimisküsimustele.

Töös analüüsiti 25 Eesti EFL intervjuud ning 50 inglise keelt emakeelena kõnelevate intervjuud. Kokku esines Eesti EFL õppijate seas sõna *make* 254 korda iga 100 000 sõna kohta ning emakeelt kõnelevate seas 147 korda iga 100 000 sõna kohta. Tulemuseks oli Eesti EFL kõneleja sõna sagedus rohkem kui 100 iga 100 000 sõna kohta kõrgem. Nende sagedasest kasutusest ilmneseid ka sõna *make* ülekasutused.

Märksõnad:

Inglise keel ja keeleteadus, korpusuuring, õppijakeel, grammatika.

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kadi Talpsepp,

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

The Overuse of MAKE in Estonian Learner Language,

mille juhendaja on Jane Klavan,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NCND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost äri eesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olenteadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kadi Talpsepp

25.05.2021

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.

[Autori allkiri]
Kadi Talpsepp

25.05.2021

Lõputöö on lubatud kaitsmisele.

[Juhendaja allkiri]
Jane Klavan

25.05.2021