UNIVERSITY OF TARTU

DEPARTMENT OF ENGLISH STUDIES

A CORPUS-BASED STUDY OF THE POLYSEMY OF *CHEERS* IN AMERICAN

AND BRITISH ENGLISH

BA thesis

SIGNE-RENATE SAAR

SUPERVISOR: *Lect.* JANE KLAVAN

Tartu

2018

ABSTRACT

Corpus-based studies are an important field of study as it gives a good overview of a specific part of a particular language. With the present thesis, I aspire to give a glimpse into the different uses of the word *cheers* in American and British English. I chose this word because it is polysemous, i.e. it is a word with many meanings. My main interest was to discover if there are any differences in the patterns how this word is used in the two main varieties of English. For example, I was curious as to whether *cheers* would be used more in the function of thanking in British English compared to American English.

The thesis consists of seven sections. The introduction gives an overview of the topic of polysemy in general and an explanation of the corpus-based analysis that was conducted. The first part of the thesis provides a theoretical discussion of polysemy and focuses on the corpus-based and traditional approaches to polysemy as well as a glimpse into cognitive linguistics. The second half of the thesis focuses on the corpus-based study of the word *cheers* in American and British English. It explains the annotation scheme and highlights the main results.

TABLE OF CONTENTS

1. INTRODUCTION

The aim of this research paper is to find out in which ways the word "cheers" is used in American and British English. This topic sparked my interest as it seemed like a research I had no experience with previously. It seemed like a good way to challenge myself by analyzing a particular word in detail. This word in particular seemed a good one to base a research paper upon as it has various definitions and the word also belongs to different parts of speech. Furthermore, there seem to be differences between how the word "cheers" is used in American and British English. To my knowledge, no empirical study exists on this topic.

In order to fulfil the aim of the thesis, an empirical corpus-based study was conducted. The dataset was compiled by randomly selecting 1,000 sentences which use the word pattern "cheers", 500 from American English and 500 from British English. Two corpora were used: British National Corpus (http://www.natcorp.ox.ac.uk/) and Corpus of Contemporary American English (https://corpus.byu.edu/coca/). Both spoken and written subcorpora were used. . The data were extracted from the corpus interface and transported into a Google spreadsheet. The data were then manually annotated in the Google spreadsheet for various categories. The data were divided into categories based on their parts of speech and meaning. The main categories of meaning were toasting (*She took the cocktail which the waiter had just brought. "Cheers", she said raising her glass to him.*), thanking (*"Thanks then." "Cheers"*), leaving (*"Thanks for the lift, Rainbow. Cheers for now."*), other and "not applicable". The aim of the study was to look at how frequently each of the meanings is attested in the two corpora and to see if there are any differences between the types of meaning and register (spoken vs. written) and the types of meaning and variety (American vs. British).

The category "other" was divided into TV-series (*"What is your favourite TV show?" "Cheers"*), noun in plural (*So in summary, two cheers for integration*), verb form (*It's always sunny, so it cheers you up*) , communication (*Cheers Steve, bye bye*), politics (*To cheers, it was announced yesterday morning that a composite...*) and name (*The first project of New African Visions, a not-for-profit organization created by Cheers*). The label "not applicable" was given to occurrences, the meaning of which could not be determined based on the corpus sentence.

The thesis is divided into six sections. The first theoretical section discusses the notion of polysemy in a very broad sense. In this section, the definition of polysemy is given. The next theoretical section discusses how the traditional approach to polysemy is different from the corpus linguistic approach to polysemy. This topic is important because the thesis is a corpus-based study of polysemy. In addition, the first section of the thesis gives a very brief overview of usage-based linguistics, because this is the general approach to linguistics that is especially favourable to corpus-linguistic studies of meaning. The thesis then gives a short overview of the different meanings of "cheers" given in various dictionaries of British and American English. Then, a corpus-based study is carried out and the results are discussed according to the meaning categories annotated in the sample, and the distribution of the categories across register and variety. The discussion compares the meanings of the corpus-based results. The thesis ends with a conclusion.

2. USAGE-BASED LINGUISTICS

Both Cognitive and Functional linguistic research proceeds from the premise of a usage-based approach to doing linguistic research. This boils down to using empirical methods, either linguistic experiments or corpus-based methodology. The present thesis adopts a usage-based

perspective because it is important to look at how language is used in everyday situations and to collect empirical data from a corpus. It is not scientifically conductive for the researcher to rely only on his or her intuition, although intuition is clearly important when coming up with the research question and interpreting the results. An aspect which plays a crucial role in usage-based analysis (of linguistic structure) is the regularity of occurrence. The usage-based approach has over time switched its focus from functional and cognitive linguistics to the effects of frequency and organization and development of linguistic knowledge (Diessel 2017: 3). Its importance lay in the fact that the aim of usage-based linguistics is to establish a framework for the analysis of the emergence of linguistic meaning and structure (Diessel 2017: 1).

According to Glynn (2014: 8), semantics which is usage-based has to take two significant steps. Firstly, inductive research methods need to be used. Based on the data samples, generalisations are the only possible variant for hypothesis testing. Secondly, corpus-driven semantic analysis must be developed. Natural contextualised language production has to be examined. Corpora means natural language samples which are large enough to enable inductively valid claims. Statistics lends a helping hand here. To identify structure, multivariate statistics need to be used as they are an influential tool because of the complexity of the data. (Glynn 2014: 8) However, in the present thesis only frequency counts are used since there is no training provided for statistical techniques at the BA level.

3. POLYSEMY

In the following section, a short definition will be given of what is meant by polysemy in the present thesis. The definitions are based on four sources: Glynn's (2014) paper on the concept

of polysemy and synonymy in Cognitive Linguistic research, Gries' (2015) contribution on polysemy to the Handbook of Cognitive Linguistics, Chapter 9 from the textbook on practical lexicography by Atkins and Rundell (2008) and Saaed's (2003) textbook on semantics. In addition, some dictionary meanings will also be discussed.

The most often used and accepted explanation of polysemy would be "a form of ambiguity where 2+ related senses are associated with the same word" (Gries 2015: 472) This notion of polysemy was proposed by Bréal (1897) (cited in Gries 2015). In his paper, Glynn defines polysemy as "different concepts-functions of a form" and synonymy as "different forms of for a concept-function" (2014: 10). He then elaborates this idea and comes to the conclusion that polysemy is also the functional-conceptual variation of any given form and that synonymy is the functional-conceptional relation between any symbolic forms (Glynn 2014: 11).

According to the Oxford Guide to Practical Lexicography (2008), "some polysemous words have a particular relationship with others in their lexical set, in that several of their meanings seem to parallel each other. Certain specific semantic components result in sets of words behaving lexicographically in a very similar way. This is known as regular polysemy." (Atkins and Rundell 2008: 1)

One of the other lexical relations that needs to be discussed in relation to polysemy is homonymy. Saeed (2003) regards homonyms as unrelated senses of the same phonological word. (Saeed 2003: 63). When comparing homonymy and polysemy, Saeed says that "both deal with multiple senses of the same phonological word, but polysemy is invoked if the senses are judged to be related." (Saeed 2003: 64). For lexicographers it is important to make the distinction, as polysemous senses are listed under the same lexical entry, whereas homonymous senses are given separate entries. (Saeed 2003: 64).

Saeed used the Collins English Dictionary (Treffry 2000: 743) to provide the reader with some examples of polysemy. The example word was "hook". The following are the different meanings: 1) a piece of material, usually metal, curved or bent and used to suspend, catch, hold, or pull something. 2) short for fish-hook. 3) a trap or snare. 4) chiefly US something that attracts or is intended to be an attraction. 5) something resembling a hook in design or use. 6.a) a sharp bend or angle in a geological formation, esp. a river. 6.b) a sharply curved spit of land. 7) boxing a short swinging blow delivered from the side with the elbow bent. 8) cricket a shot in which the ball is hit square on the leg side with the bat held horizontally. 9) golf a shot that causes the ball to swerve sharply from right to left. 10) surfing the top of a breaking wave.

These definitions and different thoughts regarding polysemy (and homonymy) are attached to this thesis to give some insight to the topic as it is something that this thesis very closely deals with.

According to the Oxford Dictionary, polysemy is "the coexistence of many possible meanings for a word or phrase" (https://en.oxforddictionaries.com/definition/polysemy). Merriam-Webster Dictionary explained the word as "having multiple meanings" (https://www.merriam-webster.com/dictionary/polysemy). Thus, in the context of "cheers" we can see that it has more than 2 senses related to the same word.

In the present thesis, I am following Gries' (2015) thoughts, meaning I count polysemy as having 2 or more senses related to the same word.

3.1 LINGUISTIC APPROACHES TO POLYSEMY

Gries (2015) explains very well the treatment of polysemy in cognitive linguistics. "It involves viewing meaning/sense as categorization, recognizing the importance of context for

meaning/senses and that linguistic and encyclopaedic knowledge are hard to keep separate, and incorporating prototype theory into linguistics." (Gries 2015: 473)

The first aspect of Gries' (2015) definition means that meaning/sense is viewed as categorization. The author gives an example using sparrows - this means recognizing that a sparrow is a bird which leads to establishing that birds is a category which a sparrow is a member of.

For the second category, to be able to mark the meaning of a lexical item means to take into consideration both the context and encyclopeadic real-world knowledge of the given lexical item.

As for the third category, "cognitive linguistics has drawn on research in cognitive psychology that showed subjects/speakers do not categorize objects using necessary/sufficient features but by comparing their similarity to the prototype of the candidate category/categories." (Gries 2015: 473)

The different meanings that I have pinned upon the word *cheers* may have some cultural connection. For example, for thanking and toasting in British English, there is a high probabilitythat they derived from one another. The same goes for thanking and using *cheers* as a form of saying goodbye.

However, I agree with the third category in the sense that not so much categorization using sufficient features is done. The context and encyclopeadic real-world knowledge can be taken into consideration and can be used to mark the meaning of the word. For example, encyclopaedic knowledge was necessary to determine that a certain number of meanings in the corpus samples of *cheers* were related to the US TV-series of the same name.

The Oxford Guide to Practical Lexicography (Atkins and Rundell 2008) describes that in classical semantic theory, a discrete meaning embodies a cluster of 'criterial features', that

means the particular conjunction of sufficient and necessary conditions which identify that meaning. (Atkins and Rundell 2008: 276) Atkins and Rundell refer to Aitchison (2003) who said that "words and meanings are not so much 'precision instruments' as 'slippery customers', whose exact boundaries can rarely be drawn with any confidence. (Aitchison 2003: 41 cited in Atkins and Rundell 2008: 276)

"With an understanding of prototype theory, and of the inherent (and pervasive) fuzziness of word meaning, we are in a better position to take on the task of identifying and describing dictionary senses." (Atkins and Rundell 2008: 278)

It was brought out in The Oxford Guide to Lexicography (Atkins and Rundell 2008) that the prototype approach has two substantial advantages over the classical model: 1) it is a reflection of the way people create meanings when they communicate, therefore it goes with the grain of the language and contains fuzziness and creativity. 2) the lexicographer's task is made more manageable, as it allows us to focus on the prototype and its exploitations, instead of requiring us to predict and use our intuition for every possible instantation of a meaning. (Atkins and Rundell 2008: 280)

Glynn (2014) proposes that if we assume that the relation between a concept-function and a form becomes stable through repeated contextualized use, then the result is an identifiable source of study. With that being said, Glynn gives Langacker's (1987: 59-60) explanation to polysemy and synonymy. Polysemy is therefore an "entreched functional-conceptual variation of a schematic or non-schematic form" and synonymy is an "entreched functional-conceptual relation between schematic and non-schematic forms." Langacker's theory provides a frequency-based operationalisation of grammaticality - "the more often a form-meaning is used, the more automated its processing becomes and the more 'grammatically acceptable' it is according to the

speaker's intuition." (cited in Glynn 2014: 13). In the context of *cheers* it can be assumed that the more frequently a meaning of *cheers* is instantiated, the more entrenched it becomes.

The approach to categorize objects using necessary and sufficient features is known as the traditional approach to polysemy. "The knowledge of sense frequencies resulting from the behavioural, although of limited use for a traditional cognitive-linguistic analysis of word meaning, are in fact very useful for psycholinguistic findings." (Gries 2006: 91). However, in the cognitive semantics approach, a different view of word meanings is taken. According to Saeed (2003: 342), in many semantic approaches it is presumed that language is thought of as a mental faculty and that linguistic abilities are supported by and need special forms of knowledge. (Saeed 2003: 342). However, the view of cognitive linguistics is that there is no separation of linguistic knowledge from general thinking or cognition. (Saeed 2003: 342).

„Cognitive linguists often point to a division between formal and functional approaches to language. Formal approaches, such as generative grammar, are often associated with a certain view of language and cognition: that knowledge of linguistic structures and rules forms an autonomous modul (or faculty), independent of other mental processes of attention, memory and reasoning. This external view of an independent linguistic module is often combined with a view of internal modularity: that different levels of linguistic analysis, such as phonology, syntax and semantics form independent modules." (Saaed 2003: 343). In the present thesis, I view polysemy and *cheers* as it is viewed in cognitive linguistics (as this is a corpus-based analysis), meaning intuition is used but context and encyclopaedic knowledge is more trusted and they stand behind the intuitive analysis.

As time went by and research developed, research on polysemy went two different ways. New theoretical approaches were developed and polysemy research started using more diverse data, which included corpus and psycholinguistic experimentation. (Gries 2015: 477)

One of the new and most notable theoretical approaches taken was *Principled Polysemy approach* by Tyler and Evans (cf. Tyler and Evans 2001; Evans 2005). They introduced criteria by which to determine when two usages constitute different senses by distinguishing polysemy from vagueness. They also "proposed criteria to identify the prototype, or sanctioning sense, of a polysemous category." (Gries 2015: 477) For (some) usage to be counted as a distinct sense of *x*, it must also consist of an additional meaning which is not apparent with *x.* It will also feature highly distinctive syntagmatic and/or collocational patterns and similarly distinctive structural dependencies. (Gries 2015: 477)

The principled polysemy approach is a promising approach as it was among the first of such to introduce more proper „decision principles" for deciding where to draw the line between different meanings. Many aspects of polysemy, however, are gradable and they may not converge. Still, these criteria assist to make decisions more replicable because more linguists' decisions with more empirical evidence are gathered. (Gries 2015: 478) According to Glynn, it is precisely this research tradition that freed the study of semantic relations from the notions of discrete senses and context independent semantics. Radial network studies were the first and essential step towards this realisation – both theoretically and analytically. Glynn (2014: 26) continues, "Secondly, such studies are an essential step in empirical research. They represent hypothetical models of language structure, based on careful and systematic introspection-based analysis of language."

Linguistic research has to follow a sample-based methodology if it has no constructs such as ideal competence or langue. This goes for both Cognitive and Functional linguistic research. Different techinques which are thought of as experimental have been established for the analysis of semantics, however corpus methods have little representation in the field. (Glynn 2014: 7,8)

The theory of Cognitive Linguistics does not distinguish internal language module, such as semantics, lexis, pragmatics or syntax. Therefore, corpus-driven methods needs to account for integrating these factors of language structure and do that in a conceptually and functionally plausible manner. (Glynn 2014: 8)

3.2 A CORPUS LINGUISTIC STUDY OF POLYSEMY: THE CASE OF *RUN*

Following is a short overview of one of the key studies of polysemy that is considered a classical study within corpus-based cognitive semantics and which has been an inspiration for the present study. The present thesis follows a similar methodology and hence it is pertinent to give a short overview of Gries' study on the polysemy of *run* (2006).

According to Gries' 2006 study on the polysemy of *run,* the investigation of polysemy of lexemes and constructions has been one of the central areas of cognitive linguistics. Usually, the thought of a polysemous word requires that the particular lexeme which is being investigated has more than one distinct sense and that the senses are related (or else the lexeme would be thought of as homonymous) (Gries 2006: 58). For his study, Gries (2006) looked at the word *run* and its different senses. 815 instances were chosen from two different corpora: the British section was taken from the International Corpus of English and the Brown Corpus of American English was used for the American portion. The sentences were identified manually based on the match of the citation to senses listed in dictionaries. As Gries (2006) conducted a very in-depth research, he came across a surprising 55 different senses in which *run* was used. Gries (2006) brought out "motion" as the main category. However, this was subcategorized into abstract motion, to cause motion, fast motion and fast pedestrian motion.

The author (Gries 2006) highlighted the fact that the corpus sample he used was not nearly large enough, which I do agree with in the context of the present thesis. A sample of 500 occurrences per variety of English may not be sufficient for a detailed analysis of the distribution between the various meanings of the word "cheers". With corpus-based researches and studies the fear of not collecting enough data to begin with is very prevailing, as I can now say from my own experience as well. Another important consideration is which corpora to use, since there are very many good corpora available for different registers and varieties.

Gries (2006: 87) claimed that cognitively-oriented analyses of polysemy gain from a corpus-based perspective. In addition, a thorough empirical approach to polysemy allows for the analysis of results using state-of-the art statistical techniques. Gries (2006: 88) states that "[t]he main multifactorial technique employed above has been the hierarchical agglomerative clustering technique; its main emphasis has been on determining degrees of similarity between (groups of) senses." Thanks to the development of corpus-based semantic studies researchers can use more objective techniques to tease apart the different meanings of one word. Since there are no statistical techniques taught at the BA level, the present thesis only looks at proportions and no advanced statistics has been used. This remains a possible research topic for future research.

Two crucial issues were brought out by Gries (2006). One of them being that it would be rewarding to discard traditional word senses on behalf of meaning components. The second issue was that "a more explicit cognitive analysis of *to run* could provide more evidence of the frequencies of mechanisms which figure in extensions of words (i.e. metaphorical, metonymical or image-schematic mappings, profile shifts, frame additions etc.)" (Gries 2006: 89).

3.3 POLYSEMY OF "CHEERS"

Cheers was a good example to base this present thesis on as there is a variety of different senses and ways of use, starting with toasting and ending with the verb sense. In addition, it seems there may be a difference in how this word is used in American and British English. Based on an intuitive observation, it can be assumed that in British English the use of this word as a synonym for thank you is much more frequent than in American English. However, in order to validate this intuition-based hunch, a corpus-based study needs to be conducted.

Following are the uses which can be found in a few different online dictionaries. The entries below served as input when putting together the meaning categories to be annotated in the corpus sample studied for the present thesis.

The Oxford Living Dictionaries (https://en.oxforddictionaries.com/definition/cheers) gives the definitions of 1) *expressing good wishes before drinking* (informal) *"Cheers,' she said, raising her glass"* 2) *expressing good wishes on parting or ending a conversation* (British) *"Cheers, Jack, see you later."* and as an extension of that 2.1) *expressing gratitude or acknowledgement for something* '*Billy tossed him the key. Cheers, pal."*

The Cambridge Dictionary (https://dictionary.cambridge.org/dictionary/english/cheers) was the third online dictionary I used. The given definitions were 1) *a friendly expression said just before you drink an alcoholic drink, "Cheers! Your good health."* 2) *used to mean "thank you"* (UK informal) *"I've bought you a drink." "Cheers, mate."*, and 3) *used to mean "goodbye"* (UK informal) *"Bye." "Cheers, see you next week."*

Merriam-Webster (https://www.merriam-webster.com/dictionary/cheers) gives two definitions to the word, first of them being *a toast,* the second *to express thanks* and as an extension of that *a way to bid another farewell*.

According to the Online Etymology Dictionary (https://www.etymonline.com/word/cheers), the word *cheers* originates from the British form the year 1919. It was used as a plural form of *cheer*. In the year 1720, it has been recorded as a shout of encouragement or support.

4. CORPUS-BASED STUDY OF *CHEERS*

4.1 CORPUS SAMPLE

The data was collected by randomly selecting 1000 sentences which use the word *cheers* as the keyword for the corpus search, 500 from American English and 500 from British English. I used two corpora: British National Corpus (http://www.natcorp.ox.ac.uk/) and Corpus of Contemporary American English (https://corpus.byu.edu/coca/). Both spoken and written subcorpora were used. The data were manually annotated in a Google spreadsheet with 1000 rows for each of the sentences and various columns for the various variables, for example, meaning and part of speech.

"The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century, both spoken and written." (http://www.natcorp.ox.ac.uk/corpus/index.xml)

The written (90%) part of the BNC includes amongst other things extracts from national as well as regional newspapers, specialist periodicals and journals, academic books and popular

fiction, letters, essays etc. The spoken (10%) part consists of orthographic transcriptions of informal unscripted conversations and spoken language collected in different contexts, ranging from formal business or government meetings to radio shows and phone-ins. (http://www.natcorp.ox.ac.uk/corpus/index.xml)

The Corpus of Contemporary American English (COCA) is the largest freely-available corpus of English, and the only large and balanced corpus of American English. COCA is one of the most widely-used corpus of English.

The corpus contains more than 560 million words of text and it is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. (https://corpus.byu.edu/coca/)

4.2 ANNOTATION OF THE DATA

The data were annotated for various variables with different category levels. The main categories for meaning, which were determined based on the various dictionary entries discussed above, were the following:

1) toasting: *She took the cocktail which the waiter had just brought. "Cheers", she said raising her glass to him.*
2) thanking: *"Thanks then." "Cheers"*
3) leaving: *"Thanks for the lift, Rainbow. Cheers for now."*
4) other
5) not applicable (NA)

The category "other" was divided into TV-series (*"What is your favourite TV show?" "Cheers"*), noun in plural (*So in summary, two cheers for integration*), verb form (*It's always sunny, so it cheers you up*) , communication (*Cheers Steve, bye bye*), politics (*To cheers, it was announced yesterday morning that a composite...*) and name (*The first project of New African Visions, a not-for-profit organization created by Cheers*). The label "not applicable" was given to occurrences, the meaning of which could not be determined based on the corpus sentence. The aim of the study was to look at how frequently each of the meanings is attested in the two corpora.

4.3 RESULT OF THE CORPUS-BASED STUDY

For the purpose of generalising the results, various tables with frequency counts are given. The results are presented according to the main variables annotated in the data sample. The tables were created in Excel spreadsheets with the PivotTable function. In addition, proportions of each category level are also discussed in order to generalise over the absolute numbers of frequency counts. As for the overall frequency of the word pattern cheers in the two corpora, the pattern has 6.62 occurrences per million words in the BNC and 16.49 occurrences per million words in COCA. Using the "Chart" functions from the BYU interface for the two corpora, it can be seen how frequent this pattern is in the different registers in the different corpora (see Table 1 for BNC and Table 2 for COCA). It can be seen from Table 1 that the word pattern cheers is especially frequent in spoken and newspaper registers in the British English. As for American English, Table 2 demonstrates that the pattern is especially frequent in spoken and newspaper registers as well. If we were to compare the data in Table 1 and Table 2, it is possible to make the prediction that the word pattern cheers is more frequent in the spoken and newspaper registers in British English than in American English.

Table 1. Overall frequency of *cheers* in BNC

Section	All	Spoken	Fiction	Magazine	Newspaper	Non-Acad	Academic	Misc
Frequency	662	162	132	48	170	39	7	104
Per million	6.62	16.26	8.3	6.61	16.24	2.36	0.46	4.99

Table 2. Overall frequency of *cheers* in COCA

Section	All	Spoken	Fiction	Magazine	Newspaper	Academic
Frequency	7,666	1,121	898	676	1,035	103
Per million	16.49	9.60	8.03	5.76	9.16	0.92

One of the outputs of the "Chart" function for COCA allows one to inspect how the word pattern has changed in time. Table 3 demonstrates that the use of *cheers* is on the rise, at least for American English, although compared to other time slots, it was also fairly frequent in the period of 1990-1994 (a time period that incidentally corresponds to the time period when the BNC was compiled).

Table 3. Frequency of *cheers* in COCA according to different years

Section	All	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014	2015-2017
Frequency	7,666	790	660	567	530	712	574
Per million	16.49	7.60	6.38	5.51	5.19	6.92	9.21

In order to take a closer look at how the word pattern *cheers* is used in contemporary British and American English, a sample of 500 random sentences was selected from the entire list of 622 occurrences from BNC and 7,666 occurrences from COCA. Following is a detailed analysis of 1,000 sentences according to the different meaning categories annotated and according to the register and variety.

General results

In Table 4, the main categorizations of the dataset can be seen. The most frequent usage of *cheers* was labelled as "other" which was used 719 times out of 1000 which makes 71.9% of the data less relevant for the present thesis. The main aim was to compare the use of *cheers* as used for leaving, thanking and toasting across the two main varieties of English. Second in frequency were the words which were labelled as "not applicable". These words made up 14.1% as there were 141 of them. Again, this category is not of central importance to the study, since it does not contribute to the discussion of the three main functions of *cheers* that are of interest. "Thanking" was used 72 times which makes up 7.2%. Toasting, to my surprise, was used only 45 times or 4.5%. Lastly, the usage of "leaving" was used 23 times which makes 2.3%.

Table 4. Meaning categories for *cheers*

Meaning	Number of uses
leaving	23
NA	141
other	719
thanking	72
toasting	45
Grand Total	**1000**

Meaning of *cheers* according to the variety of English

In Table 5, the varieties and meanings can be seen. In the first row, American English is brought out and in the second the British English occurrences can be viewed. As shown in Table 5, *cheers* as an expression used to say when leaving was more frequent in the British English sentences, occurring 21 times compared to 2 times in American English cases. Considering the cultural differences, this was to be expected as the people of Britain are thought of as more polite than Americans. Another culturally visible difference that can be seen in the table above, is that thanking is much more frequent amongst the British English sentences, as in British English it appeared 63 times and only 9 times in the American English sentences. *Cheers* as an expression for toasting occurred in total 45 times, 19 of them being in British English and 26 of them being in American English. This brings out another cultural difference, however, making an inference that Americans enjoy drinking more and therefore they also toast more a rather far-fetched one. The categories NA and other did not have such substantial differences across the two varieties.

Table 5. Variety and meaning.

Row Labels	American	British	Grand Total
leaving	2	21	23
NA	53	88	141
other	410	309	719
thanking	9	63	72
toasting	26	19	45
Grand Total	**500**	**500**	**1000**

Table 6 demonstrates the same categories as the previous but in percentages. All the columns in Table 6 make a total of 100%. As can be seen from Table 6, , within the American English sentences, the use of cheers as other was by far the most frequent, occurring in 82% of the 500 sentences. The same category was also the most frequent among British English sentences, appearing in 61,8% of the cases out of the 500 sentences. Leaving was the least frequent among the American English sentences as it occurred in only 0,4% of the sentences. In the British English section, toasting was the least frequent meaning of cheers used.

Table 6. Variety and meaning (proportions)

Meaning	American	British	Grand Total
leaving	0.40%	4.20%	2.30%
NA	10.60%	17.60%	14.10%
other	82.00%	61.80%	71.90%
thanking	1.80%	12.60%	7.20%
toasting	5.20%	3.80%	4.50%
Grand Total	**100.00%**	**100.00%**	**100.00%**

Meaning of *cheers* according to the different registers

The second largest categorization for the annotated sentences was whether the sentences occurred in spoken or written language. As it is shown in Table 4, written sentences were more frequent, as there were 742 out of 1000 sentences which makes 74.2%. There were 258 spoken sentences which makes 25.8% out of all the sentences. The most frequent occurrence for spoken sentences were conversations. These divided into speeches, classroom discussions, radio shows, meetings etc.

As in the BNC 90% is written data and only 10% is spoken, and COCA claims to be equally divided among spoken and written data, it is not possible for the spoken and written portions of my collected data to be 50% and 50% when it comes to register. However, the data does allow for the inspection how the different meaning categories identified in the corpus

samples are distributed among the two registers. The distribution of meaning according to register is given in Table 8.

Table 7. Register

Register	Number of uses
spoken	258
written	742
Grand Total	**1000**

Table 8. Register and meaning

	Spoken	Written	Grand Total
leaving	18	5	23
NA	75	66	141
other	100	619	719
thanking	53	19	72
toasting	12	33	45
Grand Total	**258**	**742**	**1000**

Table 8 demonstrates how all the meaning categories identified for each of the occurrences were divided between the spoken and written registers. As it was revealed above in Table 5, the category of other was the most frequent, both in the written and spoken categories. The category of other appeared 100 times in spoken cases and 619 times in written cases. NA

was the second most frequent category, occurring 75 times in the spoken cases and 66 in the written cases. The third category was thanking, appearing 53 times in the spoken section and 19 times in the written section. Toasting was not as frequent, to my surprise, as it occurred 12 times in the spoken sentences and 33 in the written sentences. Lastly, c*heers* as a way of saying goodbye was used 18 times in spoken English and only 5 times in written English.

Table 9 presents the meanings and registers in a form of proportions. As it has been discussed before, the category of other was the most frequent in both written and spoken language. Out of all the sentences from the spoken language, 38.76% of the cases belonged to the category labelled "other", 83.42% of all the sentences from the written language belonged to the category of "other".NA was second in frequency, comprising 29.07% of the spoken cases end 9.89% of written cases. The third category was thanking, being used in 20.54% of the spoken sentences and 2.56% of the written sentences. Toasting spread out quite equally, occurring in 4.65% of the spoken cases and 4.45% of the written cases, coming to a total of 4.5% out of all the sentences. Lastly, leaving was the least frequent, appearing in 6.98% out of the spoken sentences and 0.67% of the written sentences. Intuitively logical was the result of thanking appearing more in spoken language as it is an expression that is used orally rather than written down. Similar to thanking, was the result of leaving as it is something people say when leaving. (E-mails and text messages that use *cheers* as a way to say goodbye or ending a conversation was labelled under the category "other").

Table 9. Meaning and register (proportions)

Meaning	Spoken	Written	Grand Total
leaving	6.98%	0.67%	2.30%
NA	29.07%	8.89%	14.10%
other	38.76%	83.42%	71.90%
thanking	20.54%	2.56%	7.20%
toasting	4.65%	4.45%	4.50%
Grand Total	**100.00%**	**100.00%**	**100.00%**

Other uses of *cheers*

  The usage of cheers which was labelled as "other" is given in Table 10. Since it comprises a relatively large part of the data sample, it merits a more detailed discussion. As was expected, the most frequent was the usage of the word as a noun in plural, which occurred in 466 times out of 718. For example, "Three cheers for Taureg!" From this example sentence, it is easy to see that cheers as a noun is associated mostly with sports events and it represents the happy shouting which occurs in competition situations.

Table 10. Specifications for "other".

Specification of other uses of cheers	Number of uses
communication	21
name	9
noun	466
politics	2
TV-series	120
verb	100
Grand Total	**718**

Second in frequency was the usage of *cheers* as the name of a TV-show, which was used 120 times out of 718. This, in hindsight, was more popular in the American corpus sample as the TV-show originates from the United States. The occurrence of this use was most usual in magazines, advertisements and news stories. For example, "Whatever happens, it's time to raise one last glass to Cheers and the end of a decade of vintage comedy."

The usage of *cheers* as a verb was next in frequency, being used 100 times out of 718. This also occurred mostly in sentences, which touched on the subject of sports events or which talked about wanting to make someone feel better (to cheer somebody up). For example, "And as soon as your child cheers up, you'll feel better too."

The next category was communication due to lack of a better title. In this section, I compiled all the sentences which dealt with someone saying goodbye, but not quite leaving. Here I collected the sentences which were either e-mails or phone calls etc. This category encompasses

the notion of writing "*Hope we all meet up again soon. Cheers, Gav*" at the end of an e-mail or saying goodbye to someone before ending the phone call.

5. DISCUSSION

While comparing the corpus samples of the pattern cheers in American English and British English, quite a few similar elements emerged. For example, the occurrence of thanking and leaving. These categories were both more frequent in the British English variety. However, toasting was used more in the American English variety.

One of the findings pertains the use of the word pattern to refer to the American TV show "Cheers". "Cheers" as a TV show was mentioned a lot more in the American sample 'than in the UK examples. This is presumably due to the fact the TV-show originates from the United States. In this meaning, the word was most often used in TV programmes and in magazines.

For the communication meaning, British English had a lot more examples compared to the American data. This category was made up of *cheers* as a way of ending an email and by *cheers* as a way to end a phone call; ways of ending communication between (at least) two parties. This result is not particularly informative as *cheers* can be observed to be used much more often in this context by the British rather than the Americans. Still, it is good that the numbers support the intuition.

The use of *cheers* as a noun or "a shout of applause or encouragement" (as defined by the Merriam-Webster dictionary; https://www.merriam-webster.com/dictionary/cheer) did not have such a big of a difference when it comes to the ratio between the American and British examples. This use occurred more in the spoken portion of the data, more specifically in sports commentary.

An interesting thing I came across was the use of *cheers* as a name or a title. The sentences in which it was used, divide into three categories: as a name hinting to a political term; as a first name; and as the last name of someone.

Since the COCA corpus is being updated frequently and BNC corpus is a closed corpus (which is no longer updated), the results could have been different if other corpora had been used. As a beginner, the best idea at the time of the research was to choose the most standardised corpora. However, future research could expand the present study in various ways. First of all, future research may focus only on the spoken section or the more informal section of the corpora, since this is where it seems the pattern cheers occurs more often. It would be interesting to see if the different meaning categories differ in their distribution. Another interesting avenue for further research would be taking a look at other varieties of English, for example Australian English or Canadian English. One of the possible corpora that can be used to achieve this aim is the GLOWBE corpus (Corpus of Web-Based Global English) or the ICE corpus (International Corpus of English). Yet another possibility is to investigate how the pattern is used in learner English using the ICLE corpus (International Corpus of Learner English).

6. CONCLUSION

Corpus-based research is a field which has undergone many different changes throughout the years as new developments have been made and corpora and technology are ever-changing. Polysemy, however, has been researched a lot but the definitions do not differ in substantial ways. Nevertheless, with the advent of technology and the emergence of corpora, a new field of study has come to exist – semantic studies that are based or driven by corpus research.

This thesis attempted to analyze and decode 1000 different variations from sentences which used the word c*heers*. The word form can be mean different things: thanking, toasting, leaving, a way to say goodbye or end a phone call, form of noun in plural, verb form etc. The aim of the thesis was to see how different meaning categories of *cheers* are used in American and British English. In order to fulfil the aim, two corpora were used: The British National Corpus and The Corpus of Contemporary American English.
The results were somewhat predictable to some extent, due to intuition but there were also some surprises.

71.9% of all of the analysed cases fell under the category of 'other', which sadly made a big portion of the data less relevant in terms of the research question. The category 'not applicable' or 'NA' made up 14.1% of all the data, which again made more of the data less relevant.

The searched categories, thanking and toasting, made up 7.2% and 4.5% per cent of the data, accordingly. This means the most relevant parts of the data made up 11.7% of all the collected data. Some cultural differences appeared after the analysis - for example, people of Britain use cheers as an expression said when leaving a lot more than Americans. Similar to this, people of Britain use cheers as a way of thanking much more than people of America.

However, a category that was used more by the Americans than by the people of Britain, was toasting.

All of the data was either spoken or written and the percentages of that were 25.8% and 74.2%, accordingly. This means that two thirds of the data was written and only one third was spoken. Since the used corpora were the BNC (British National Corpus) and COCA (Corpus of Contemporary American English), one of which is 90% made up of written data and the other claiming to be equally divided between the two registers, it was is impossible for the registers to be equally represented.

REFERENCES

Atkins, B. T. Sue and Micheal Rundell. 2008.The Oxford Guide to Practical Lexicography. Published in the United States by Oxford University Press Inc., New York.

British National Corpus, available at http://www.natcorp.ox.ac.uk/, accessed November 1, 2017.

Corpus of Contemporary American English, available at https://corpus.byu.edu/coca/, accessed November 1, 2017.

Diessel, H. 2017. *Usage-based linguistics.* Oxford Research Encyclopedia of Linguistics. New York: Oxford University Press.

Evans, V. 2005. *The meaning of time: Polysemy, the lexicon and conceptual structure.* Journal of Linguistics 41(1): 33−75.

Glynn, D. 2014. *Polysemy and synonymy. Cognitive theory and corpus method.* Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy. John Benjamins Publishing Company.

Gries, St. Th. 2006. *Corpus-based methods and Cognitive Semantics: The many senses of to run.* In St. Th. Gries, & A. Stefanowitsch (Eds.), Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis (pp. 57–99). Berlin & New York: Mouton de Gruyter. DOI: 10.1515/9783110197709

Gries, St. Th. 2010. *Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics.* The Mental Lexicon. John Benjamins Publishing Company.

Gries, St. Th. 2015. *Polysemy.* Handbook of Cognitive Linguistics, 472-490. Berlin & Boston: De Gruyter Mouton.

Tyler, Andrea and Vyvyan Evans. 2001. *Reconsidering prepositional polysemy networks: the case of over.* Language 77(4): 724−765.

RESÜMEE

TARTU ÜLIKOOL

ANGLISTIKA OSAKOND

Signe-Renate Saar

A CORPUS-BASED STUDY ON THE POLYSEMY OF C*HEERS* IN AMERICAN AND BRITISH ENGLISH

KORPUSE PÕHINE UURIMUS SÕNA C*HEERS* POLÜSEEMIAST AMEERIKA JA BRITI INGLISE KEELES

Bakalaureusetöö

2018

Lehekülgede arv: 36

Annotatsioon: Korpusepõhised uurimused on valdkond, mis annab hea ülevaate kindla valitud keele kindla keeleüksus(t)est. Käesoleva tööga püüan anda vaadet sõna *cheers* erinevatele kasutusvõimalustele Ameerika ja Briti inglise keeles. Mainitud sõna valisin sõna polüseemia tõttu. Polüseemia tähendab, et sõnal on mitu tähendust. Minule põhiliselt huvi pakkuv valdkond oli näha, kas ja milliseid erinevusi on sõnal inglise keele kahes põhivariandis. Näiteks soovisin ma teada saada, kas *cheers*'i kasutatakse tänuavaldusena rohkem Briti või Ameerika inglise keeles.

Käesolev tees koosneb seitsmest sektsioonist. Sissejuhatus annab ülevaate polüseemiast kui terminist ja annab üldise seletuse läbiviidud korpusepõhilisele analüüsile. Teesi esimene pool pakub teoreetilist arutlust polüseemia üle ja keskendub keele korpusepõhilise ja traditsioonilise käsitluse võrdlusele ja lisaks tegeleb ka kognitiivse keeleteaduse valgustamisele. Teesi teine pool keskendub sõna *cheers* korpusepõhilisele uurimusele Ameerika ja Briti inglise keeles. Selgitatakse kategooriaid ja tõstetakse esile põhilise tulemusi.

Märksõnad: polüseemia, korpuse põhine, analüüs, inglise keel

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Signe-Renate Saar,

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

Korpuse põhine uurimus sõna *cheers* polüseemiast Ameerika ja Briti inglise keeles,

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.

allkiri:

Signe-Renate Saar

Tartus, 22.05.2018

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:
Signe-Renate Saar

Tartus, 22.05.2018

Lõputöö on lubatud kaitsmisele.

Juhendaja allkiri:
Jane Klavan

Tartus, 22.05.2018