Building a Lexical Corpus: A Glossary of Film Terminology

Martin Blažek

Univerzita Tomáše Bati ve Zlíně
Fakulta humanitních studií
Ústav anglistiky a amerikanistiky
akademický rok: 2012/2013

ZADÁNÍ BAKALÁŘSKÉ PRÁCE
(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Martin BLAŽEK**
Osobní číslo: **H10090**
Studijní program: **B7310 Filologie**
Studijní obor: **Anglický jazyk pro manažerskou praxi**
Forma studia: **prezenční**

Téma práce: **Výstavba lexikálního korpusu: filmová terminologie**

Zásady pro vypracování:

Studium odborné literatury
Získávání odborných materiálů z oblasti lexikologie a filmu
Výstavba specializovaného glosáře zaměřeného na filmovou terminologii
Vyhodnocení získaných dat
Sumarizace výsledků a závěr

Rozsah bakalářské práce:

Rozsah příloh:

Forma zpracování bakalářské práce: tištěná/elektronická

Seznam odborné literatury:

Beaver, Frank Eugene. Dictionary of Film Terms: The Aesthetic Companion to Film Art. 4th ed. New York: Peter Lang Publishing, 2009.
Béjoint, Henri. Modern Lexicography: An Introduction. Oxford: Oxford University Press, USA, 2000.
Bowker, Lynne, and Jennifer Pearson. Working with Specialized Language: A Practical Guide to Using Corpora. London: Routledge, 2002.
Cruse, Alan. Meaning in Language: An Introduction to Semantics and Pragmatics. 3rd ed. Oxford: Oxford University Press, USA, 2011.
Landau, Sidney I. Dictionaries: The Art and Craft of Lexicography. 2nd ed. Cambridge: Cambridge University Press, 2001.
Lipka, Leonhard. English Lexicology: Lexical Structure, Word Semantics, and Word-Formation. Narr, 2002.
Kennedy, Graeme. An Introduction to Corpus Linguistics (Studies in Language and Linguistics). London: Addison Wesley Longman, 1998.

Vedoucí bakalářské práce: **Mgr. Hana Atcheson**
                                 Ústav anglistiky a amerikanistiky
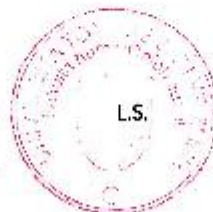
Datum zadání bakalářské práce: 30. listopadu 2012

Termín odevzdání bakalářské práce: 3. května 2013

Ve Zlíně dne 28. ledna 2013

L.S.

doc. Ing. Anežka Lengálová, Ph.D.
děkanka

PhDr. Katarína Nemčoková, Ph.D.
ředitelka ústavu

PROHLÁŠENÍ AUTORA BAKALÁŘSKÉ PRÁCE

Beru na vědomí, že

- odevzdáním bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby [1];
- beru na vědomí, že bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k nahlédnutí;
- na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3 [2];
- podle § 60 [3] odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- podle § 60 [3] odst. 2 a 3 mohu užít své dílo – bakalářskou práci - nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- pokud bylo k vypracování bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tj. k nekomerčnímu využití), nelze výsledky bakalářské práce využít ke komerčním účelům.

Prohlašuji, že

- elektronická a tištěná verze bakalářské práce jsou totožné;
- na bakalářské práci jsem pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.

Ve Zlíně 28. 4. 2013

(2) Disertační, diplomové, bakalářské a rigorózní práce odevzdané uchazečem k obhajobě musí být též nejméně pět pracovních dnů před konáním obhajoby zveřejněny k nahlížení veřejnosti v místě určeném vnitřním předpisem vysoké školy nebo není-li tak určeno, v místě pracoviště vysoké školy, kde se má konat obhajoba práce. Každý si může ze zveřejněné práce pořizovat na své náklady výpisy, opisy nebo rozmnoženiny.

(3) Platí, že odevzdáním práce autor souhlasí se zveřejněním své práce podle tohoto zákona, bez ohledu na výsledek obhajoby.

2) zákona č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, § 35 odst. 3.

(5) Do práva autorského také nezasahuje škola nebo školské či vzdělávací zařízení, užije-li nikoli za účelem přímého nebo nepřímého hospodářského nebo obchodního prospěchu k výuce nebo k vlastní potřebě dílo vytvořené žákem nebo studentem ke splnění školních nebo studijních povinností vyplývajících z jeho právního vztahu ke škole nebo školskému či vzdělávacímu zařízení (školní dílo).

3) zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, § 60 Školní dílo:

(1) Škola nebo školské či vzdělávací zařízení mají za obvyklých podmínek právo na uzavření licenční smlouvy o užití školního díla (§ 35 odst. 3). Odpírá-li autor takového díla udělit svolení bez vážného důvodu, mohou se tyto osoby domáhat nahrazení chybějícího projevu jeho vůle u soudu. Ustanovení § 35 odst. 3 zůstává nedotčeno.

(2) Není-li sjednáno jinak, může autor školního díla své dílo užít či poskytnout jinému licenci, není-li to v rozporu s oprávněnými zájmy školy nebo školského či vzdělávacího zařízení.

(3) Škola nebo školské či vzdělávací zařízení jsou oprávněny požadovat, aby jim autor školního díla z výdělku jím dosaženého v souvislosti s užitím díla či poskytnutím licence podle odstavce 2 přiměřeně přispěl na úhradu nákladů, které na vytvoření díla vynaložily, a to podle okolností až do jejich skutečné výše, přitom se přihlédne k výši výdělku dosaženého školou nebo školským či vzdělávacím zařízením z užití školního díla podle odstavce 1.

2

ABSTRAKT

Ve své bakalářské práci vysvětluji pojem korpus a popisuji proces jeho tvorby. Hlavní zájem zaměřuji na specifika tvorby korpusu a další jeho vlastnosti v jazyce dneška. Teoretickou část začínám vysvětlením pojmů korpus a korpusová lingvistika. Pokračuji jejich praktickým využitím a příklady a dále se zaměřuji na tvorbu speciálního korpusu. Pojítkem mezi teorií a praxí je v mé práci kapitola věnovaná procesu výstavby specifického korpusu, v mém případě filmového glosáře. Praktickou částí mé práce je pak glosář filmové terminologie vytvořený na základě postupu popsaného v části teoretické.

Klíčová slova:

Korpus, korpusová lingvistika, glosář, film, terminologie.

ABSTRACT

The aim of my Bachelor thesis is on lexical corpus. This topic itself would be extremely wide, since the corpora are great many and more and more are being created as the languages of the world develop. Therefore my further aim, my deeper field of study, is in film terminology. It is one of the recently developed ones (only has been around since the second half of nineteenth century), yet it is still growing because the innovation in film industry is unstoppable.

My theoretical part is strongly based in the works of world's renowned linguists and is mainly concerned in the process of creation of a special corpus. The link between theory and practice is in the description of glossary creation process and practical part itself is a glossary of film terminology with both Czech and English terms, translation, and explanations.

Keywords:

Corpora, Corpus linguistics, Glossary, Film terminology.

ACKNOWLEDGEMENTS

CONTENTS

INTRODUCTION

Corpus linguistics is one of the fastest growing fields of language study. The use of Internet and high-performance personal computers enables almost anyone to take a look at what corpus research is capable of. I see my thesis as an understandable basic guide to corpora and its research. It should help anyone who is interested in corpus linguistics but has yet no insight in this field, to gain all the information needed for basic corpora study.

The first chapter of theoretical part of my thesis sheds light on what corpora and corpus linguistics are, their history, what their aim is and how both of them can be used in practice.

Second chapter of my thesis focuses on the process of creating a specialized corpus. It depicts what language for special purposes is and how it is connected to specialized corpora. It also describes the process of designing and compiling a specialized corpus as well as the process of its construction. Basic corpus processing tools are also mentioned in order to give the reader a view of what tools are available for a corpora search.

Third chapter of my thesis is aimed at the process of glossary creation. This part is the last step between theoretical knowledge and practical use. It provides the reader with four most important steps, starting with document (text) selection and resulting in a bilingual corpus based glossary.

The practical part of my work is a glossary of film terminology derived from corpus. The texts used to build my glossary are Internet based glossaries of film terms and the corpus created from these consists of thousands of words. This allows my glossary to be used as a guide for anyone who is interested in the study of film science.

The glossary itself consists of the term extracted from my corpus, its English and Czech explanation and, where available, a Czech translation of given term. The users of my glossary might be film science students as well as amateurs interested in film terminology.

I.   THEORY

1   CORPUS LINGUISTICS

1.1   The Definition of Corpus and Corpus Linguistics

The description of Corpus linguistics is the most essential part of my work simply because it gives the reader of my thesis the basic background of what I am trying to achieve. Linguistics, as a scientific study of human language, has been around for a very long time and great many sub-fields of this discipline have emerged during its existence. One among many of these sub-fields is Corpus linguistics.

The term corpus itself comes from Latin and stands for body. Graeme Kennedy describes Corpus as: "In the language sciences a corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description." (Kennedy 1998, 1)

Another explanation of the same phenomenon can be found on Corpus Linguistics Web Site. It indicates that corpus can also be described as a collection of texts, systematical or random. The number of corpora collected can be very large, yet no matter how large, they cannot represent the entire language. Corpus linguistics is used to analyze these texts and to answer various linguistic questions. It also offers an insight into the dynamics of language which is what makes it one of the most valuable and widely used linguistic methodologies.

Definition of corpus as a linguistic tool can be found in the work of Nadja Nesselhauf. "A corpus can be defined as a systematic collection of naturally occurring texts (of both written and spoken language)." (Nesselhauf, 2005) Further in her text she explains that the systematic nature of corpus is based on principles which are not rooted in linguistics but in other scientific disciplines. The term corpus as we know it can be used to describe any collection of texts chosen according to some pattern or common features, however, its meaning has a "narrower sense" in modern linguistics and corpus is mainly used "to refer to systematic text collections that have been computerized." (Nesselhauf, 2005)

Her definition of Corpus linguistics is that "Corpus linguistics is a method of carrying out linguistic analyses." (Nesselhauf, 2005) This linguistic discipline has the ability to investigate and answer many different questions and also carries a high potential in further development of the study of language. With the application of computer software it has become one of the most important and frequently used methods in linguistic research. (Nesselhauf, 2005)

The most simple and understandable definition of corpus can be done through creating a computerized record of personal texts. If I decided to digitalize my complete email correspondence with my close friend, I would be able to create my personal email corpus. This would be an example of a "Personal Corpus" which is valuable to a single person, however not of great value to linguistics. In linguistic studies, corpora are gigantic texts and in these, linguists can find many different varieties of chosen language.

Concerning English there is a difference between "e.g. the Corpus of English native speakers and the Corpus of English as second language speakers or the Corpus of English learners." (Mukherjee 2009, 4) According to Mukherjee, another aspect that has to be taken into consideration while discussing corpora is the time period in which it is based. The Corpora of Present Day English (English of everyday use in the 21st century) are many. So are those of Modern English period. However there are not many corpora concerning Old, Middle and Early Modern English periods. Mukherjee also believes that the size of corpora is one of its crucial characteristics.

Another important fact Mukherjee points out is that the texts in corpus must be chosen according to a pattern and based on common features. A collection of random texts cannot be called corpus. "A corpus must consist of texts chosen according to certain principles, and its usefulness will be judged on how well these principles are represented." (Mukherjee 2009, 4)

1.1.1 Types of Corpora

Alongside the biggest corpora, which are designed for linguistic purposes and carrying the most linguistic data, it is also necessary to mention smaller units of corpora, which also have certain importance.

Special corpora are not concerned with Standard English as a whole. They mainly focus on specific types of language and different phenomena connected to it. On one hand there are for example specific business correspondence corpora, closely connected to one specific fragment of language, and on the other hand there are corpora such as Corpus of London Teenage language (COLT), concerned with specific age group and language used by this age group in a very specific area. (Mukherjee 2009, 6)

Parallel corpora are another prime example of specific corpora type. Such corpora consist of texts which are translated from another language and become part of a corpus of the language we study. "For example, a French newspaper article is translated into English

and included in the English corpus, while an English article is translated into French and incorporated into the French corpus." (Mukherjee 2009, 7) These corpora are a great aid to comparative linguists and translators, since a direct comparison can be done in a matter of seconds.

In order to trace grammatical and lexical changes linguists use Diachronic corpora. Texts rooted in different time periods can be found in such corpora and the connection between time and language is easy to find. (Mukherjee 2009, 7)

Very specific and yet again very useful type of corpora Mukherjee mentions is Learners corpora, which consist of texts written by students trying to acquire a second language. In such texts, both linguists and students (authors) can find the common mistakes that are made in language acquisition. This comes in handy for language teachers and students.

1.2   History of Corpus Linguistics

The first attempt to design a corpus-based linguistic discipline dates back to the end of 19th century. "German linguist J. Kading used a large corpus consisting of about 11 million words to analyze distribution of the letters and their sequences in German language."(Corpus Linguistics Web Site) The size of this corpus was revolutionary and beyond compare at the time.

As indicated on the Corpus Linguistics Web Site, even though the birth of modern corpus linguistics dates to 1960's with The Brown Corpus designed by H. Kucera and W. N. Francis, the linguistic discipline as we know it today, with millions or sometimes even billions of words, dates to late 1980's and is connected to the development of high-performance computers. With further emendation of personal computers, it is now possible for almost anyone to have access to these computer stored corpora and to examine them.

Graeme Kennedy claims that: "Corpus linguistics did not begin with the development of computers but there is no doubt that computers have given corpus a huge boost by reducing much of the drudgery of text-based linguistics and vastly increasing the size of the databases used for analysis. It should be made clear, however, that corpus linguistics is not a mindless process of automatic language description." (Kennedy 1998, 2)

Nevertheless, Noam Chomsky, one of the best known linguists of 20th century, claims that Corpus is a list of imprecise data which can be distorted by the fact that in some cases words would occur in sentences, which are false or can be omitted completely. When

Chomsky made this statement in 1960's Corpus linguistics was an extremely slow and impractical linguistic field. However I believe that with the invention of capable and affordable computers, it became more useful and therefore more valid. (Šudová 2009, 11)

1.3   The Targets of Corpus linguistics

The basic target of Corpus linguistics can be expressed through asking a set of questions, answers to which can be found in the use of Corpus linguistics:

"How can we best exploit the opportunities which arise from having texts stored in a machine-retrievable form? What linguistic theories will best help structure corpus-based research? What linguistic phenomena should we look for? What applications can make use of the insights and improved descriptions of languages which come out of this research?" (Kennedy 1998, 3)

As the Corpus Linguistics Web Site says, to answer these questions it is necessary to evaluate the possibilities that Corpus linguistics offer. For example, the fact that Corpora are created from natural and real-life based texts makes the study of dialects, sociolinguistics and stylistics one of the main domains of Corpus linguistics. Today's studies of Corpora are all computer based and therefore extremely fast and precise. Also the time needed to find certain word or a phrase (key word, key phrase) is reduced to a minimum.

It can also be found on the Corpus Linguistics Web Site that the study of patterns and collocations has been made sufficient through the use of computer Corpora. Instead of searching for one particular word or phrase it is now possible to search for whole sets of phrases and words and their occurrence in language. At the same time, various parameters of search and tools used to conduct such search can be used, which allows the result to be precise.

In his book, Graeme Kennedy points out that corpus, as the body of language, "is the main source of evidence for linguistic description and argumentation. " The research activity conducted throughout Corpora is mainly based on quantity of usage of certain phenomenon and its occurrence. Therefore, he does not see Corpus linguistics as a new linguistic branch but more as a source of information which is valuable to all linguistic disciplines. "Linguists have always needed sources of evidence for theories about the nature, elements, structure and functions of language, and as a basis for stating what is possible in a language. At various times, such evidence has come from intuition or

introspection, from experimentation or elicitation, and from description based on observation of occurrence in spoken or written texts. In the case of corpus-based research, the evidence is derived directly from texts."(Kennedy 1998, 7)

1.4   The Practical Usage of Corpus Linguistics

One of the main advantages of Corpus linguistics, as Prof. Dr. Karin Aijmer (Aijmer et al. 2009, 3) claims, can be found in students using corpora of a language they are trying to acquire. The fact that situations captured in computer databases are realistic makes them perfect study aid for non-native speakers. Based on this fact, the language itself becomes more acceptable and easier to grasp by their reader.

Daniel Krieger specifies the use of Corpus linguistics in language teaching and as he claims, Corpus linguistics can be used in three specific fields in teaching: syllabus design, which can be modified according to specific needs of students based on corpus research; development of materials for students, which can be compiled according to needs of specific group of students, again based on research of corpus consisting of these students work; and last but not least, the classroom activities. "The teacher can guide a predetermined investigation which will lead to predictable results or can have the students do it on their own, leading to less predictable findings. This exemplifies data driven learning, which encourages learner autonomy by training students to draw their own conclusions about language use." (Krieger 2003, 1)

Importance of use of corpus based data is also visible in the study of grammar. Corpora are bases for both theoretical and practical approach towards grammar studies, as they present a great number of grammatical phenomena and also carry empirical data for grammatical theory. The quantitative research carried out through corpora is fundamental for grammar and other linguistic fields. (McEnery and Wilson 2001, 27)

Another linguistic branch which gains a lot of significant data from corpora is sociolinguistics. As seen on the Corpus Linguistics Web Site, sociolinguistics search corpora for differences and similarities which occur based on social status and specific placement of the text creator. The collection of both spoken and written utterance from all over the world enables linguists to compare texts and reach reliable conclusions.

From the point of view of my thesis, the most important aspect of practical corpora use is in its use in lexicography. "Corpus linguistics plays an important role in compiling, writing and revising dictionaries as within a few seconds, the linguist can get examples of

words or phrases from millions of spoken and written texts. And since corpora continue to grow and are constantly being expanded with new texts, lexicographers have an instant access to up-to-date information." (Corpus Linguistics Web Site)

On the other hand, the Corpus Linguistics Web Site states, that there are some questions, which Corpus linguistics is not able to answer. The co-occurrence of words is easy to find in corpora but, for example, the increase in usage of certain word over a period of time can not be explained by Corpus linguistics.

2   BUILDING A SPECIALIZED CORPUS

Building a corpus is not a simple process and there are many mistakes which can be made. The second part of my thesis gives a detailed description of the process of corpus creation with particular examples and hints. It introduces the LSP (Language for Special Purposes) and points out the differences between this particular type of language and the Language for General Purposes (LGP). It also shows how to design a special purpose corpus and how to compile this type of corpus. The most important part of this chapter is the introduction to the creation of glossaries using corpora and a description of the process of using specialized software tools to extract terms from corpora. (Bowker and Pearson 2002, 26)

2.1   Language for Special Purposes

As mentioned above, one of the key points of this part is Language for Special Purposes. The description of this phenomenon, according to Lynne Bowker would be: "(LSP) is the language that is used to discuss specialized fields of knowledge." The use of plural Languages for Special Purposes would actually be more accurate, since there are many LSPs which are connected to many different fields of knowledge. Each field has its own unique set of language patterns which can occur in other fields but it might also be used in one field only. LSP and LGP very often overlap, as some terms from LSP become words of everyday use and therefore part of LGP. This process, as Bowker states, can be called de-terminologization. Words such as Anorexia, which were part of LSP used only by doctors and psychiatrists, become part of LGP through their use by non-specialists (journalists, media).

Another usage of LSP can be seen in the process of language learning. Word bank of each language student is based on LGP, since it is the most useful communication tool, mainly because it is understandable to all native speakers. Adding words from LSP to LGP contributes to the ability to participate in dialogue on more specific topics. (Bowker and Pearson 2002, 26)

LSP, even thought it is a specialized language field, is not only used by specialists as I mentioned earlier. There are three main groups of users: experts, who have studied the field thoroughly and are trained and experienced in both the field itself and its language; semi-experts, in most cases students of the field, who gained an insight into the field, but have not used it in practical life and non-experts, who for some reason use LSP without deeper insight. This includes translators or technical writers, who have to be familiar with usage of

LSP, but their knowledge is not deep enough to e.g. debate on an academic level. (Bowker and Pearson 2002, 27-28.)

Bowker also indicates that the use of corpora can be a fast and reliable way to gain an overview of certain LSP. Through the creation of specialized corpus and with the use of corpus analysis software it is possible to research certain LSP corpus for e.g. which words are most commonly used, which words are only used in this particular LSP etc.

2.2   The Design of Special Purpose Corpus

The most intricate part of corpus creation is collecting the data. There are many different criteria that have to be taken into consideration and the most important ones are the size of corpus that is to be created, the use of extracts or full texts or the number of texts used to create a corpus.

As Bowker points out, the size of corpus is defined by many factors and the final decision about the size of a corpus is left to be made by its creator. Depending mainly on the needs and aims of the project, the amount of data (texts) available and amount of time the author is willing to spend on the creation. In this case it can be said, that the size of corpus is not as important as the quality of texts used to create it. Concerning the purposes of my work, the ideal background text for the creation of glossary is of course a number of glossaries on the same topic. (Bowker and Pearson 2002, 45)

Another size vs. quality comparison can be used in the case of texts used to create a glossary. Partial texts or text extracts are not always useful because the particular words linguists have interest in can be spread throughout the whole text and its part is not sufficient. The use of full and complete texts is always better since there are no words omitted and the corpus built form such texts is therefore more useful. (Kennedy 1998, 71)

Last factor is the number of texts, which are used in corpus creation. Number of various texts written by different authors with different background and sources is a valid source for corpora creation and for learning LSP. On the other hand creating corpora based on single author and his texts can be misleading simply because it does not show which words are really used in author's field of interest, but merely the words preferred by this particular author. (Kennedy 1998, 71)

2.3   Special Purpose Corpus Compilation

This part of my work sheds light on the process of corpus compilation. Once the design and layout of corpus is complete the collecting of texts begins. Choosing the most suitable texts to be included in corpora construction and to form the final corpus itself is an important issue. The outcome can differ from expectations at the beginning of corpus building process but it is still valuable. The conclusion and interpretation of data at the end of the process must contain all the possible inaccuracies made during the process itself, e.g. old texts were used, authors reputation is unknown etc. (Bowker and Pearson 2002, 58)

The finding of data can be difficult even with all the possibilities of Internet and sometimes not many texts can be found. This scenario is likely to happen if the field of study is extremely specialized and therefore only known and fully understood by a small group of users. (Bowker and Pearson 2002, 59)

Most common source of texts is, as mentioned before, the Internet. "There are two main tools that you can use to locate candidate texts for inclusion in your corpus: search engines and subject directories." (Bowker and Pearson 2002, 61) Search engines, such as Google, Yahoo or local Seznam are commonly known and widely used by all Internet users. They are also the quickest tool for search of texts. The search is very easy, it begins with typing in what you are searching for and finding first results. The search can be conducted more thoroughly by typing in more specific request or by using terms closely connected to the topic chosen. Search through subject directories is much narrower but at the same time it offers more sub-fields and other nuances which can make the search itself more successful and precise. (Bowker and Pearson 2002, 62)

Important point that Bowker also mentions is that the choice of texts is one of the major issues in corpus creation. The criteria set in the stage of corpus design now work as a sieve, which lets all the useful and valuable texts inside the corpus and keeps the less valuable ones out. The decision making process is essential in this case and it can be helped by sticking to simple rules. Web pages are not only written texts but also pictures, videos etc. and these multimedia often determine the meaning behind the page. For this reason, it is better to choose text-based web pages and stick to these. The quality and veracity of these pages can often not be verified, there might be no editorial process involved in creation of these pages (in contrast to printed texts e.g. books). "Therefore, you must take care when selecting texts from the Web for inclusion in your corpus by considering factors such as whether or not the site has been approved by an official

organization or whether the author is a subject field expert." (Bowker and Pearson 2002, 63)

2.4   The Process of Corpus Building

The corpus building process is done through various computer applications, starting with simple, on-line ones and ending with professional ones which are very expensive. Free software is more likely to have bugs and not work properly or to be too complicated for use by amateurs. The easiest way to obtain a quality software tool for corpus building is by downloading a free trial version. Most corpus building software is available for free download and through trying out different tools each person can find the one that fits them best.

John Sinclair gives a helpful list of steps which makes the creation of useful corpus easier. He says that the first step should be making a backup copy of each text used in exactly the format in which it was first found. If the text is not computerized, hard copy must be kept for further reference. To avoid future problems with text/file conversion, it is advised to keep computerized texts in simplest format possible. Formats such as HTML or DOCX might not be recognizable for some programs. On the other hand simple Notepad file in TXT is easy to read and is supported by wide range of software linguistic tools. (Sinclair 2004, 99)

Identification of a text is also a feature which makes it easier to create a corpus. The easiest way to identify every single text used in the corpus building process is to code or number them and make a reference list in which all the basic information about text, such as author, year of publication etc., can be found. The text has to be pre-processed in terms of format conversion etc. Modern computer software does the pre-processing automatically. (Sinclair 2004, 99)

When the bases for corpus is built, it is vital to make a backup copy and then keep updating the copies if any change to the corpus is made. Problems with results might develop in time and it is best to keep track of each change made to the corpus. (Sinclair 2004, 99)

2.4.1   Internet Based Data Collection

As mentioned before web is the biggest source of information, but there are certain rules that have to be obeyed in order to keep this source valuable. For lexicographic purposes it is necessary to see the Internet itself as a resource for further study, not as a source.

Using the most advanced technology available a great deal of work has already been done on all sorts of texts and therefore, it is not necessary to carry out deep research of Internet as a whole. In simple terms, many corpora are already created and rather than creating new ones, it is possible to take the ones that already exist and create a corpus which is simplified and only contains information that has already been verified. (Atkins and Rundell 2008, 78)

2.5   Corpus Processing Tools

As the past proves, the search and analysis of corpora could be done without use of computers. However it is the computers that give the search speed and accuracy. Mastering the use of such computer programs is not easy but it can be done. It is essential to start with simple tasks and work one's way up.

2.5.1   Word List

Enthusiasts who want to find their way to corpus research are likely to begin their work with producing a word list. It is the simplest tool of every corpus processing program, which creates a list of all the words that are used in certain corpus. The software I used for my research is called WordSmith and it offers a free download from the Internet and also a free trial (for limited period of time). Creation of word list is done by selecting the texts which are to be used and simply clicking the "WordList" function feature. The result of such text examination offers three different outputs: "statistical analysis, frequency-ranked word list and an alphabetically ordered word list." (Bowker and Pearson 2002, 109)

Statistical data comes in a chart and shows the characters used, token, their frequency etc. The most valuable outcome can be seen in the charted count of words according to their length (number of letters). LSP based texts are more likely to contain number of very long words, since specialized terms are often taken from another languages, on the other hand simple LGP texts are more likely to contain shorter words of everyday use.

The content of frequency list will probably be quite misleading, yet understandable. The list of most frequently used words contains functional words, which help build the text and which make it comprehensible. There are usually no lexical words, which would carry meaning themselves, present in the top positions of word list. The outcome of such analysis is rather in how the text is created than what it is made of. To point out a specific example, In Scientific American corpus the word "I" does not place in the chart of most

frequently used words. The explanation is that scientist do not use first person in their work. (Kennedy 1998, 245)

Key words of a text can also be identified using one of the tools available in the software. This feature enables the user to point out what the text is about in one simple step. Key words list, unlike the frequency list, shows words which do not occur very often in other texts. It is based on comparison of texts rather than breaking down one text. (Kennedy 1998, 245)

The feature of alphabetical listing simply alphabetizes all the words used in a text and gives a number of their occurrence. The outcome can have many interpretations and is most importantly used to gain access to all words beginning with the same letter. This is helpful in comparing how often certain words tend to occur in a text based on their resemblance or how often such words tend to be placed close to each other. (Kennedy 1998, 247)

Word clusters can also be examined through the use of this tool. Practical application of this feature can be found in studying LSP, phrases and phrasal verbs. It can also chart all the multiword units in a text and noun phrases as well, which is again useful for further studies of specific features of each text and even the whole language. (Bowker and Pearson 2002, 118)

2.5.2 Concordance

There are two basic types of concordances connected to corpus creation and examination: monolingual and bilingual. Monolingual concordance can be found in texts written in one language and its main use is to find the presence of corresponding words in their natural context. On the other hand the bilingual concordance allows the search of word in both the language of origin and its translation. (Kennedy 1998, 251)

Bowker points out that monolingual concordance is based on the search of a chosen pattern in a corpus. The outcome comes in a table and can vary depending on used software. The word of choice is displayed in context which is usually eighty characters long, but it can be set to show the entire sentence or even paragraph in which the word occurs to expose the whole setting surrounding the word. The outcome of search can be alphabetically ordered based on the word preceding or following the word of concordance and this order can be set to up to three words in front of or behind the word we desire to study.

The right choice of search pattern allows the search to be precise and based exactly on the demands of a linguist conducting the search. Most of the research done in the field of concordance is closely connected to the earlier mentioned lists of words, but concordance allows a more thorough search based on frequency of use of words. The most frequently occurring words are most likely to be the basis for a further study such as the search for context words. Narrowing down the search to a more specific level is easily done through the use of concordancer by simply adding a context word. The retrieval of information is then based on the concordant word itself and another word of choice which narrows down the search. (Kennedy 1998, 256)

Collocates can also be found using concordancing software. "Some concordancers offer an additional facility which frequency ranks the words that appear in the vicinity of your search pattern. This facility computes and displays the most frequently occurring collocates. Collocates are words which typically occur in the vicinity of your search pattern." The monolingual concordance is, based on above mentioned possibilities of use, a tool to find patterns of words occurrence in text, its meaning and the context in which it is usually set. (Bowker and Pearson 2002, 125-126)

What Bowker also says is that bilingual concordances utilization is in the examination of parallel corpora (original text and its translation). The way this tool is used is the same as in monolingual concordance. "The difference is that bilingual concordancers also retrieve the corresponding translation segments." (Bowker and Pearson 2002, 127) The outcome of search of both texts, the original and its translation, is displayed at the same time and allows pointing out the interconnection not within single language but in both languages used. The main application of this tool is in checking for correctness of translations, and in language learning. Language enthusiasts use it in order to see patterns in language and their translations, as if they were using a bilingual dictionary. (Bowker and Pearson 2002, 129)

3   CREATING A GLOSSARY

Final chapter of the theoretical part of my work defines the term glossary and shows how glossaries are created. Its main focus is on the process of glossary creation with the use of corpus linguistics and corpus processing software. It is the final link between the theory of my thesis and its practical part.

There are many definitions of what glossary is and they all have one thing in common. It is that glossary is a list of terms and can consist of words in either one or several languages.

To quote Lynne Bowker: "A glossary is essentially a list of terms in one or more languages." The Oxford Dictionary's definition is that a glossary is "an alphabetical list of words relating to a specific subject, text, or dialect, with explanations; a brief dictionary." Whatever the definition, one thing is for sure: glossary contains a list of words or phrases in one or more languages, their translation, and explanatory part, concerning these words and their meaning.

There are many varieties of glossaries available. Simplest glossaries are most likely to only contain a word and its translation with no further information. These glossaries are designed with the assumption that reader is familiar with the topic and simple translation is satisfactory. More user friendly glossaries have translation part, explanatory part (preferably in both languages) and equivalents as well as synonyms (if available). (Bowker and Pearson 2002, 138)

3.1   Process of Glossary Creation

There are several methods which lead to the creation of glossary. The method used by professional terminologists consists of several steps, starting with text collection, through data examination and ending with the process of writing a glossary. However there are easier ways to create a glossary and since I do not consider myself a professional terminologist, I decided to use the simpler way used in Lynne Bowker's book.

3.1.1   Document Selection

The collecting of documents is a relatively simple process, since Internet offers countless number of texts. Nevertheless it is the relevance of text that is important and sometimes hard to determine. Choosing the right texts to include in the corpus is crucial. Base corpus

composed of 30 000 words is a good start and can provide the researcher with a lot of beneficial information. (Bowker and Pearson 2002, 145)

3.1.2   Choice of Terms

The definition of term in Oxford Dictionary is : "a word or phrase used to describe a thing or to express a concept, especially in a particular kind of language or branch of study;" (Oxford Dictionary Web) Term extraction can be done in many different ways, the easiest one being by browsing the texts and manually picking terms. This term extraction style is extremely time-demanding. That is why it is more favorable to use corpus processing software.

One of the tools, as mentioned before, is frequency list. It points out the most frequent terms (consisting of one, two or three word-phrases) and can show the surroundings of each term with the use of concordance tool. At the end of this process, a list of most frequent word clusters from all texts is created and is made use of in the glossary building process. (Bowker and Pearson 2002, 145-146)

Keyword list is another option of term extraction and is even more precise than above mentioned frequency list. "The advantage of using this type of facility is that it will generally give you an even better indication than frequency lists of what a text is about." (Bowker and Pearson 2002, 147)

3.1.3   Terminological Profile

As indicated on the Docsymmetry.com Web site, once the choice of terms is done and the very base of glossary is created, it is necessary to provide the user with some additional information about each term. This additional information should essentially be the definition of each term, since reader might not be familiar with all the terms and their use, and also a translation if bilingual glossary is created.

While creating such description it is best to take explanatory parts from other glossaries and dictionaries and compare them. The use of already existent explanations to create new ones is highly effective and profitable.

Other aspect that has to be taken into account is the use of words in phrases and sentences. Glossary which contains such information and example is much more valuable, especially for students of language. The real life use of a term or phrase is the best example available. (Bowker and Pearson 2002, 153-156)

3.1.4 Bilingual Glossary

Adding a second language into glossary can be done in couple of hours. The creation of another corpus is necessary, in this case in language to which the original terms are to be translated. The choice of text is based on the topic of glossary and the theme of texts in both languages must be identical.

The process itself is exactly the same: the terms are already chosen, so equivalents in second language have to be found and listed. The occurrence of each word in second language word list (frequency list, key word list) will be similar to the first one and the terms that will pop-up are likely to match the ones from list number one. As soon as an exact equivalent is found, the search for definitions begins. Again, best definitions are created in combination of use of already existent definitions in second language and personal creativity. (Bowker and Pearson 2002, 161)

3.2 Summarizing the Process of Glossary Creation

Bowker believes that while the corpus-based approach to glossary creation might not be able to answer all the questions that arise; it is still a beneficial tool for linguists and language students. Through the creation of corpora, the use of corpus processing software and most importantly through the creation of glossary itself, both linguists and students can learn a lot of valuable information. "A corpus-based approach allows you to identify subject-specific term candidates quickly and easily, and, most importantly for you as a language student, to identify typical collocation patterns." (Bowker and Pearson 2002, 162)

II. ANALYSIS

4 THE GLOSSARY OF FILM TERMS

The practical part of my work describes the process of creation of my own glossary based on glossaries available on the Internet. I chose to use five different glossaries which are listed further in my work. These glossaries together created a corpus consisting of hundreds of thousands of words, from which I had to extract the necessary terms and their definitions. Result of the whole process can be seen further in my thesis, namely in my own glossary.

4.1 Creation of My Glossary

The following part of my work describes how my own glossary was created, from first search of the Internet using Google.com to the finishing touches on the content of my glossary. As I mentioned before, for the purposes of my corpus building and glossary creation I decided to use five glossaries I found on the Internet. Using Google.com, I found suitable glossaries by simply typing "Glossary of Film Terms" in the search bar. This search resulted in finding thousands of texts, from small ones with only few terms explained, to more complex ones with hundreds of terms. To use glossaries of all different sizes and to represent the whole spectrum of glossaries available in my corpus, I chose two smaller ones, one medium-sized glossary and two of the biggest Internet based glossaries available.

My corpus consisted of several thousands of words and it was not difficult to search. Most of the terms I used in my glossary co-occurred in at least two texts, some of them even in all five. The terms I describe were chosen by me, personally, and the descriptions I provide are compilation of two or more descriptions available in my input glossaries. Czech explanations are based on translation of English ones and are entirely my work. I was only able to find a couple of Czech written glossaries and the terms in them did not correspond with the ones in my English written glossaries.

As to the translation part, I partially used Internet translators, such as Translate.google.cz or slovnik.seznam.cz, yet I never put the result given by these translators directly into my glossary. Consulting the results with Jiří Vedral's dictionary was a great aid and most of the translations are based on his book. However, there are some terms I was not able to find good translations for, with these I left the translation column empty.

All the terms are in alphabetical order and there are many interconnections between them. All the interconnected terms are described in my glossary; hence there is no need to find extra information elsewhere.

5 MY GLOSSARY OF FILM TERMINOLOGY

TERM	CZECH EQUIVALENT	ENGLISH EXPLANATION	CZECH EXPLANATION
Actor	Herec	Person representing a character in a film.	Osoba ztvárňující roli ve filmu.
Adaptation	Adaptace, zpracování	Transformation of work of art into film (Book into film through screenplay).	Přetvoření uměleckého díla do jiné formy umění.
Ambiance	Zvuková kulisa	The mood and feeling of particular scene set by sound.	Zvuková atmosféra příslušné scény.
Anamorphic	Anamorfotický	Technique of optical imagery which produces different magnification in both horizontal and vertical dimension. The picture itself seems "squeezed".	Technika využívaná k provádění horizontálních a vertikálních změn v obraze.Obraz samotný potom vypadá "zmáčknutě".
Angle	Úhel	Perspective in which camera records its object.	Úhel, pod kterým kamera natáčí snímaný objekt.
Art Director	Filmový výtvarník, architekt	Person responsible for the design of a set, its constriction and perfect look in the film.	Osoba zodpovědná za návrh a výstavbu scény a její dokonalý vzhled ve filmu.
Aspect Ratio	Poměr stran obrazu	The ratio of width to length in which the film is shot or in which it later appears on screen.	Poměr stran výšky a šířky, ve kterém je film natočen nebo následně promítán v kině či televizi.

Audition	Hlasová herecká zkouška	The process in which an actor, who tries out for a role, reads or plays out part of the script to the director.	Proces, při kterém předvádí herec své umění režisérovi tím, že předčítá ze scénáře a ztvárňuje tak roli.
Auteur Theory	-	In film criticism: Director revealing his world-view through hints in his work. Designed by Francois Truffaut.	Ve filmové kritice: autor díla, nejčastěji režisér poukazuje na svůj osobitý pohled na svět pomocí scén ve filmu, které jej přímo reflektují.
B-Movie	Béčkový Film	A film shot with little production costs, second grade actors and in short time, usually of small artistic value.	Film natočený v krátkém časovém úseku a s nízkými výdaji. Role ztvárňují druhořadí herci. Béčka mají většinou nízkou uměleckou hodnotu.
Backlot	-	The real-life backgrounds and exteriors built in the area of film studios (streets etc.)	Obrovská zázemí pro natáčení exteriérních scén vybudovaná přímo na pozemcích studií (ulice, atd.)
Barney	-	A blanket-like cover used to lower the noise of moving mechanism inside cameras.	Polstrovaný kryt kamery, užívaný ke snížení hluku vydávaného mechanickými částmi.

Best-boy	-	Technical assistant responsible for light cables on set. Term comes from whaling.	Technický asistent zodpovědný za kabeláž spojenou s osvětlením na place. Původně používán ve velrybářském průmyslu.
Bit-part	Malá úloha	Small role only lasting for one scene, usually with few lines of acting.	Malá role trvající většinou po dobu jedné scény a s minimem textu.
Blockbuster	-	Movie that has huge financial success (over 200 million dollars in USA only).Term comes from war terminology-bombing.	Film, který zaznamená velký finanční úspěch (výdělěk 200 milionů dolarů jen v USA). Termín původně užíván ve válečné terminologii jako označení pumy.
Blooper	Přebrept	Mistakes in film such as mis-spoken dialogues or factual mistakes made by actors.	Přeřeky a faktické chyby učiněné herci během natáčení a ponechané ve filmu.
Bomb	-	A film that is financially unsuccessful, a disaster for investors.	Extrémně komerčně neúspěšný film, zkáza pro investory.
Box-Office	Pokladna	The amount of money paid by movie goers, divided into domestic gross and worldwide gross.	Množství, které zaplatili návštěvníci kin za film, dělené obvykle na domácí a celosvětový hrubý výdělek.

Call Sheet	Denní dispozice	A schedule given to every member of the film crew. It points out what time they should be on set and what they should do.	Harmonogram, který dostávají během natáčení všichni členové filmového štábu. Je v něm zaznačeno kdy a kde má který člověk být a co má dělat.
Cast	Obsazení	All the actors appearing in a certain movie.	Všichni herci vystupující v daném filmu.
CGI	-	Computer-Generated Imagery, used to create 3D graphics and images in film.	Počítačem vytvořené obrazy, užívané ve filmech k tvorbě 3D grafiky a obrazu.
Character	Postava	Individual in a story; can be real or fictitious. Portrayed by actor.	Postava v příběhu, může být reálná i fiktivní. Ztvárněna hercem.
Clapboard	Klapka	Board with information about each shot in a movie. Has a moving part on top, which makes the specific clapping sound in the beginning of every shot.	Tabulka s pohyblivou vrchní částí, která vydává specifické klapání, zahajující každý záběr filmu. Jsou na ní napsány informace o každém záběru kvůli identifikaci.
Close-Up	Detail	A close-distance shot in which the main object is magnified. It is used to show importance or to build tension.	Záběr z blízka, ve kterém objekt snímání vypadá zvětšeně. Používá se pro zdůraznění důležitosti nebo pro umocnění napětí.
Composition	Kompozice	The arrangement of elements in a scene. Movement, lighting, setting etc.	Uspořádání jednotlivých prvků ve scéně. Pohyb, světlo, umístění scény atd.

Crane Shot	Záběr z jeřábu	Camera shot taken from height using a mechanical arm or an electronic device.	Záběr kamery braný z výšky, kdy je kamera umístěna na mechanické paži nebo elektronickém zařízení podobném jeřábu.
Cyberpunk	Kyberpunk	Sub-genre of science fiction named after and often unwinding from cybernetics and punk. Influenced by Raymond Chandler's novels.	Jeden z podžánrů science fiction pojmenovaný po a často se odvíjející z kybernetiky a punku. Ovlivněn romány Raymonda Chandlera.
Day-for-Night Shot	Americká noc	Shots made during the day, that is converted to look like night shot using filters, lenses etc. The illusion of moonlight.	Záběry natočené během dne, které jsou pomocí užití různých filtrů a čoček přetvořeny tak, že vypadají jako dělané v noci. Vytváří iluzi měsíčního svitu.
Depth of Field	Hloubka ostrosti záběru	Range of field, area or distance in which objects captured by camera are in acceptable focus. Usually the area 1/3 in front and 2/3 behind the object.	Rozsah záběru kamery, ve kterém je snímaný objekt v přijatelném zaostření kamerou. Obvykle je to oblast 1/3 před objektem a 2/3 za snímaným objektem.
Direct Sound	-	The audio of a film is recorded at the same time as the image.	Zvuk je nahráván ve stejném čase jako obraz filmu.

Director's Cut	-	The film version, which director feels is the best, without the interference of film studio. Also can refer to a more polished and often longer version of a film released years after initial release.	Verze filmu schválená režisérem k vydání před zásahem filmového studia. Může to být také prodloužená a vylepšená verze filmu vydaná dodatečně, někdy i léta po premiéře.
Dolly Shot	Záběr z kamerového vozíku	Shot in film made with moving camera installed on special rails. The perspective of such shot is changed due to movement.	Záběr filmu vytvořený pohybující se kamerou, která je tlačena po speciální koleji. Perspektiva záběru je pozměněná díky pohybu kamery.
Double	Dubl, dvojník	An actor who is hired to replace the star actor in a scene which is dangerous, physically demanding or sexual.	Herec najatý studiem, aby v nebezpečné, fyzicky náročné nebo sexuální scéně nahradil jednu z hvězd filmu.
Dubbing	Dubbing, ozvučení	Refers to adding a new soundtrack to a movie that has already been produced. Usually in order to fit the lip movement better, use better music or to translate to another language.	Proces přidávání zvukové stopy do již hotového filmu. Většinou z důvodu vylepšení synchronizace zvuku a pohybu rtů herců, použití adekvátnější hudby nebo pro překlad do jiných jazyků.

Editing	Stříhání, střih	The process of choosing best shots made from many takes and putting them together in order to gain the best creation possible.	Proces výběru ze všech pořízených záběrů a skládání těchto záběrů tak, aby vytvořili co nejdokonalejší celek.
Establishing Shot	Úvodní záběr	Long and wide-angle shot at the beginning of a scene, which visualizes the whole setting in order to give viewer a complex picture of a setting.	Dlouhý širokoúhlý záběr užitý na začátku scény, zobrazující celou lokaci příběhu. Tento záběr umožňuje divákovi získat široký přehled o tom, kam je příběh zasazen.
Executive Producer	Šéf produkce	Person responsible for film's production elements such as stars and screenwriter. Also responsible for financing of a film.	Osoba zodpovědná za financování filmu, herecké obsazení filmu a výběr scénáristy.
Exposition	Expozice	The process of revealing background story information in order to advance the storyline or to give detail necessary for unraveling the plot.	Proces odhalování pozadí příběhu, které je důležité pro urychlení vývoje příběhu samotného nebo pro rozuzlení zápletky.
Extras	Kompars, dav	People in minor roles not important for the story, only used to set the proper atmosphere e.g. soldiers in a battlefield	Lidé obsazení do pro příběh zcela nepodstatných rolí. Jsou využiti pro dotvoření atmosféry např. vojáci na bitevním poli.

Fade (Out/In)	Zatmít/Rozetmít	Gradual changing of light in a film from normal light to complete darkness or the other way around.	Pozvolná změna světla ve filmu od normálního osvětlení až po úplnou temnotu, nebo obráceně.
Feature Film	Hraný film, tématický film	Film with minimal length of 45 minutes, but usually from 60 to 120 minutes long.	Film o minimální délce trvání 45 minut, obvykle však trvající mezi 60 a 120 minutami.
Femme-Fatale	-	Woman character in a movie usually played by a world known star. Representing beauty and charisma.	Ženská postava ve filmu, obvykle hraná světoznámou hvězdou. Představuje snoubení krásy a kouzla osobnosti.
Film Grain	-	Film material used for film shooting. It differs in the amount and quality of chemicals used on celluloid. Fine grained films are sharper and need less light, over grained celluloid are used in low-light conditions.	Filmový materiál sloužící k natáčení filmů. Liší se množstvím užitých chemikálií. Menší množství chemikálie se používá pro točení dobře osvětlených scén, zatímco více chemicky bohatý materiál se používá na tmavší záběry.
Film Noir	-	Film genre usually represented by darker image, detectives, corruption and life in big cities.	Filmový žánr, který se vyznačuje tmavým zpracováním obrazu, hlavními postavami bývají detektivové zabývající se prohnilostí a životem ve velkoměstech.

Final Cut	Jemný střih	The final edition of a film, the version after final cut is ready to be released.	Poslední úprava filmu, verze která vzniká po poslední úpravě je připravena k vydání.
Follow-Up	-	Film, following the initially released film, no matter if its storyline precedes it or follows it. See: prequel or sequel.	Film, který následuje natočení původního filmu bez ohledu na to, jestli příběhová linie původního filmu na tento film navazuje, nebo jí předchází.
Gaffer	Směnář, předák	Person responsible for whole electric department and the head of lighting team. His assistant is the Best-boy.	Osoba zodpovědná za elektroniku na místě natáčení a zároveň za osvětlení při natáčení. Jeho pomocníkem je Best-boy.
Gender Twist	-	A role usually casted by one sex is casted by actor of the opposite sex. (Judi Dench as M in James Bond films)	Role obsazovaná herci jednoho pohlaví je obsazena hercem pohlaví opačného (Judi Dench jako M ve filmech s Jamesem Bondem)
Gothic	Gotický	Film genre with typical dark influences such as supernatural forces or death, placed in dark settings such as old castles. Usually horror films.	Filmový žánr specifický přítomností temných sil a smrti obvykle zasazený do tajuplného prostředí, např. starý hrad. Obvykle se jedná o horrory.

Grip	Obsluha kamerového jeřábu, ramenní páka	Person responsible for setting up dolly track, cranes or heavy scenery props. Physically demanding job. Head of grip department is Key Grip.	Osoba zodpovědná za výstavbu kolejí pro kameru, kamerového jeřábu a těžkých rekvizit. Fyzicky velmi náročná práce. Vedoucím oddělení je Key Grip.
Guilty Pleasure Film	-	Film which did not receive a good critique, but is very popular with viewers and gained large revenue.	Kritiky špatně hodnocený film, který je velmi oblíbený diváky a zaznamenal velké příjmy.
Head-On Shot	-	The action in a shot moves directly towards camera and enhances the viewers feeling of anticipation. Mainly used in 3D movies.	Akce se v záběru přibližuje ke kameře a umocňuje divákův pocit zatažení do děje. Používá se zejména u filmů ve 3D.
High Concept	-	Film that is made in order to attract attention of public using star actors or extreme action. The story is explained in a simple way and is usually based on social phenomenon.	Film vytvořený za účelem přilákat diváka, k čemuž pomáhá přítomnost slavných herců. Příběh je vyprávěný jednoduchou formou a obvykle založený na sociálním fenoménu.
Highlighting	Soustředění světelného zájmu	A bright ray of light is used in order to point out certain area of object in a shot e.g. eye or gun.	Paprsek světla namířený na specifickou oblast v záběru. Používá se pro zdůraznění důležitosti menší části v celku.

Homage	-	Imitation of a scene from one movie in another movie. Usually to show filmmaker's respect to another filmmaker's work.	Napodobení scény z určitého filmu v jiném filmu. Používá se jako výraz respektu mezi filmaři.
Hype	Mediální humbuk	Advertising and marketing of a movie which is broad to extreme level. Films promoted by stars in advertisement, television etc.	Obrovská reklamní kampaň věnovaná filmu. Jsou v ní zapojeni i samotní herci např. výstupy v televizi nebo reklamě.
In-Camera Editing	-	Amateur way of film production, in which all the scenes are shot in the same order as they appear in the final product.	Amatérský způsob natáčení, při kterém jsou scény natáčeny v přesně stejném pořadí, v jakém se objevují ve výsledném filmu.
Ingenue	Naivka	A young actress in a lead role and also a type of character played by young women. The character is usually naive, innocent and attractive.	Mladá herečka v hlavní roli a zároveň typ filmové postavy hraný mladými ženami. Tato postava většinou symbolizuje naivitu, nezkaženost a mladistvou krásu.
Insert Shot	Vřazený záběr	A short shot that breaks up a longer scene by bringing a close-up shot. Can be used either to release or increase tension.	Krátký záběr vložený do dlouhé scény. Většinou se jedná o záběr zblízka, který je užit jako odlehčení nebo naopak k zintenzivnění situace.

Intercut Shots	Záběr s mezistřihem, prostřihem	Shots of two simultaneous events are spliced together to create a single scene. It is used to create tension.	Záběry ze dvou, ve stejnou chvíli probíhajících, scén jsou spojeny do jedné scény pro vytvoření napětí.
Intermission	Přestávka	Short break during the projection of a very long movie. Is usually used as a "stretch break".	Přestávka při promítání velmi dlouhých filmů. Diváci ji využívají k protažení těla.
In the Can	-	The term is used when director is satisfied with a certain take. It can also be used to describe a set of scenes or whole film which is ready for viewers.	Termín se užívá, pokud je režisér spokojen s určitým záběrem. V širším měřítku je pak používám, když je celý film nebo některá z jeho důležitých částí hotová dle představ režiséra.
Jump Cut	Skok v obraze	Cut between two similar scenes or into different time setting used to create discontinuity.	Střih z jedné scény do podobné scény nebo do jiného časového období. Jeden z uměleckých projevů diskontinuity.
Juvenile	Mladík	Male character of young age in a role of young man. Counterpart to Ingenue.	Mladý muž ztvárňující roli mladého muže. Jeho protejškem je Ingenue.
Juxtaposition	Srovnávání	Two scenes in sequence that are being compared and contrasted. Also can be used to show the relationship between scenes.	Dvě po sobě jdoucí scény, které jsou porovnávány a je poukazováno na rozdíly mezi nimi. Další použití je pro poukázání na vztah mezi těmito scénami.

Key Light	Hlavní světlo	Primary lighting of an object. It is set on different parts of object during one scene to show depth and shadowing.	Primární osvětlení scény, které je zaměřováno na různá místa v oné scéně. Slouží k představení hloubky scény a jejího stínování.
Kick-Off	Začátek	The moment when production of a film is started.	Počátek produkce filmu.
Landmark Film	Průlomový film	Unique film, which pioneers a new artistic or performance approach.	Průkopnický film, který jako první používá nový umělecký nebo technologický postup.
Lead Role	Titulní úloha	The main character, usually there are both man and women lead role in one film.	Hlavní hrdina filmu. Ve většině filmů je zastoupen jak hlavní hrdina, tak i hlavní hrdinka.
Lens	Objektiv, čočka	Glass object inside a camera, which is penetrated by light. The light is changed before hitting the film.	Skleněná čočka uvnitř objektivu kamery, která mění světlo před tím, než dopadne na film.
Lip Sync	Synchronizace s pohybem rtů	The synchronization of actor's lip movement with actual sound of actor's voice in particular scene.	Synchronizace pohybu hercových rtů s jeho hlasem a zvukovým projevem ve scéně.
Library Shot	Archivní snímek	Shot made as a reserve, often used in different film that originally intended.	Záběr natočený "do zásoby". Není použit ve filmu pro který byl vytvořen, ale v jiném filmu natáčeném později.

Lighting	Osvětlení, zasvícení	Lighting of a scene is the amount of light and shadows used in a particular setting and manipulation with both of these.	Osvětlení scény je odvozeno od množství světla a stínu v záběru a označuje také manipulaci s oběma elementy.
Location	Místo natáčení	Place used for filming with its set of variables. Usually out of the studio (sometimes in back lot).	Místo, na kterém se film natáčí. Obvykle se jedná o místo mimo oblast filmového studia (může jít také o back lot).
Long-Shot	Celek, totál	The object is viewed from a distance and the whole shot is dedicated to the objects moving towards camera. Used in countryside scenes.	Objekt natáčení je vzdálen kameře a celý záběr je věnován tomu, jak se objekt ke kameře přibližuje. Používá se u záběrů krajiny a lidí pohybujících se v ní.
Looping	Ozvučení smyček	Post-Production process in which dialogues are recorded again in the studio and used to replace original on-set-made dialogues. Connected with lip sync.	Jeden z procesů při post produkci filmu. Dialogy jsou znovu nahrány ve studiu a originální dialogy jsou jimi nahrazeny. Spojení se synchronizací s pohybem rtů.
Make-up	Líčidlo, herecká maska	Tools and materials used in artistic preparation of actor's looks in order to fit his role. Mainly connected to masks and facial cosmetic changes.	Nástroje a přípravky použité k uměleckému dokreslení hercova vzhledu s přihlédnutím k jeho roli. Jedná se většinou o masky a kosmetické úpravy tváře.

Matte Shot	Záběr s maskou před objektivem	Technique in which a shot of dark object is placed into another shot with normal color. The dark one creates a silhouette-like picture (fall scene in Hitchcock's Vertigo) in the colorful one.	Technika, při které se natočí záběr černého objektu a tento záběr se překryje se záběrem barevným. Temný záběr tak vytvoří tmavou siluetu ve světlejším,barevném, obrazu. Např. Hitchcockovo Vertigo.
Medium Shot	Polocelek	Shot made form medium distance. The body language of actor is visible, but not his facial expression.	Záběr natočený ze střední vzdálenosti, ve kterém je herec zabírán od pasu nahoru. Vynikne tak řeč těla herce, ale výraz obličeje je zcela nepostřehnutelný.
Method Acting	Stanislavského metoda	Acting method based on the actor using his own life experience to interconnect with the character. Sometimes actors create situations in their lives that are similar to the characters.	Herecká metoda, při které herec hledá ve svém životě situace, které jsou situacím v roli. Někdy tyto situace uměle vytváří.
Montage	Montáž	The creation of artistic look, sound and feel through the use of visual editing.	Umělecký vzhled, zvuk a pocit z filmu je dotvářen pomocí vizuálních úprav a střihu.

Motion Blur	Rozmazání pohybem	The effect of fast moving objects shot with unadjusted camera. Can be used to deliver the feeling of extreme speed.	Efekt vyvolaný rychle se pohybujícími objekty, natáčenými kamerou neuzpůsobenou pro takové natáčení. Cíleně může být použito pro vyvolání pocitu extrémní rychlosti pohybu.
Motif	Motiv	Thematic element that is presented repeatedly in a film. It has significance in a story and can be represented by a symbol, line or even object.	Tématický prvek, jenž se opakovaně objevuje v různých úsecích filmu. Tento prvek má význam pro příběh a může mít podobu symbolu, věty nebo věci.
Narration	Průvodní text	The supporting text in a film, which is often given to the audience by an extra voice not participating in the film. Narrator can also be one of the characters in a film.	Průvodní text předčítaný hercem, který se ve filmu aktivně neobjevuje. Vypravěčem může být I jedna z postav ve filmu.
Negative, Negative Print	Negative	The film is made in reversed colors- white is black, black is white, and has to be reprinted into positive before projection. It is done this way to save costs.	Film je natočen s obrácenými barvami, bílá je černou a naopak. Před promítáním musí být přetvořen na pozitiv. Negativy jsou tvořeny kvůli šetření nákladů.

Nitrate Film Base	Nitrátový film	Flammable film base made of cellulose nitrate and later replaced by acetate film base.	Hořlavý filmový materiál později nahrazený acetátovým filmem.
Non-linear Editing	-	Modern computer based editing process, in which piece of film can be found and altered at any time without the need to look through the whole film.	Moderní počítačem podporovaný systém pro úpravy, využívající rychlosti počítače k hledání a pozměňování částí filmu bez nutnosti procházet celý film.
Off-line Edit	Off-line úprava	The edition of a film is done at a lower resolution, than the resolution of a final product. It is done to save costs and preserve the negative.	Úpravy filmu probíhají v nižším rozlišení, než v jakém bude film ve výsledku vydán. Tyto úpravy šetří náklady a umožňují uchování originálního negativu.
On-line Edit	On-line úprava	Follows the Off-line edit. The sequence is put together using high-resolution media and only the parts that will be in the final version of a film are edited. This again saves costs.	Následuje Off-line úpravu. Filmová sekvence je poskládána dohromady pomocí médií s vysokým rozlišením a jen ty části filmu, které budou ve finální verzi jsou upravovány. Opět jde o úsporu nákladů.
Optical, Opticals	Optický trik	Visual device, such as split-screen, used in film. Can be created by camera or by post-production.	Možnost vizuálního nastavení, která umožňuje například rozdělení obrazu. Tvořena kamerou či postprodukcí.

Out-Takes	Vystřižené scény	Each scene is usually taken several times and only one is picked for the final version. Ones that are not used can either be stored or shown during final credits, e.g. funny bloopers etc.	Každá scéna se natáčí několikrát a jen jeden z těchto pokusů je použit ve výsledném filmu. Některé tyto scény jsou zařazeny pro pozdější použití, jiné můžou být ukázány při závěrečných titulcích. Např. vtipné přeřeky atd.
Over-the-shoulder Shot	Záběr přes rameno	Conversations are shot this way. The camera rolls over one characters shoulder and focuses on second character in dialogue. It shows the point of view of dialogue of first character.	Tento záběr se používá při natáčení dialogů, kdy kamera přejede přes rameno prvního účastníka dialogu a zaměří se na druhého. Ukazuje tak pohled na dialog očima prvního v dialogu.
Pan	Panoramovat, švenk	Camera shot is made while the camera moves clockwise or anti-clockwise around its axis. It delivers the panoramic view of scenery.	Záběr tvořený při pohybu kamery doprava nebo doleva kolem své osy. Výsledkem je panoramatický záběr celé scény.
Pan and Scan	-	Method of conversion of widescreen shot into narrower aspect ration. Letterboxing is avoided and only the essential part of shot is shown, the rest is not displayed at all.	Metoda konverze obrazu ze širokoúhlého na užší. Vyhýbá se letterboxingu a ukazuje tedy jen podstatnou část záběru na celé obrazovce, přičemž zbytek není zobrazen vůbec.

Positive	Pozitiv	Originally colored film print created from negative. Used for projection.	Kopie filmu v originálních barvách vytvořená z negativu a určená k projekci.
Post-Production	Postprodukce	The finishing process of a film during which the sound, picture and visual effects are edited. The film is also transferred into a format suitable for release.	Dokončovací proces tvorby filmu, během něhož je upravován zvuk, obraz a vizuální efekty. Film je převeden do kompatibilního formátu, aby mohl být přehráván v kinech.
Pre-Production	-	The steps made before the shooting begins, such as casting of actors, screenplay writing, budgeting and location scouting.	Činnost spojená s tvorbou filmu odehrávající se před začátkem natáčení samotného. Jedná se o obsazování herců, psaní scénáře, tvorbu rozpočtu nebo vyhledávání natáčecích proctor.
Prequel	-	Second movie in row, which precedes the first one in story time and shows the characters and plot before the events in the first movie.	Druhý v řadě natočených filmů, který ukazuje události časově se odehrávající před událostmi prvního filmu. Postavy a zápletka jsou představeny v dřívější době.
Print	Kopie filmu	Positive copy of a film which is intended for projection. Usually comes in reels.	Pozitiv filmu určený k promítání, obvykle v několika filmových cívkách.

Producer	Producent, realizátor	The logistical chief of film staff. Takes care of budgets, finances, schedules etc. Hires director and script writer and takes care of distribution of completed film.	Producent zaštiťuje všechny logistické úkony spojené s filmem od financování přes tvorbu rozpočtu až po časové plánování. Najímá režiséra a scénáristu, má na starosti distribuci filmu po jeho dokončení.
Props	Rekvizity	All the things used on a film set, that surround an actor in order to complete the scene or those that actor needs to perform his role.	Všechny věci použité v záběru společně s hercem, ať už jej pouze obklopují a dovytváří tak scénu, nebo je herec přímo využívá k hraní role.
Q Rating	-	A research done on how recognizable celebrities are to general public and how likeable a celebrity is.	Výzkum široké veřejnosti zaměřený na oblíbenost celebrit a jejich rozpoznávání obyčejnými lidmi.
R-rating	Film nepřístupný do 17 let	Based on the rating of Motion Picture Association of American, R rated films are only to be seen by persons of 17 years of age or older. Younger people can only see these films, if accompanied by an adult.	R-rating je hodnocení na základě MPAA, který neumožňuje vidět filmy osobám mladším 17-ti let. Mladší osoby mají přístup na tyto filmy pouze v doprovodu dospělého člověka.

Reel	Filmová cívka	Film strip which is wound on plastic or metal spool and is ready for projection. Typically there are 25 minutes of film on one reel.	Filmový pás namotaný na plastovou nebo ocelovou cívku připravený na promítání. Jedna cívka obvykle nese dvacet pět minut filmu.
Remake	Nová verze	New production of an old project done with new actors, screenplay etc. Narrative remains the same.	Nově vytvořený film kopírující starší, již vydané dílo. Je obsazen novými herci a založen na novém scénáři, příběh zůstává stejný.
Reverse Shot	Protizáběr, kontrazáběr	Shot taken in a 180 degree angle to the previous shot. Can be used in dialogues, combining multiple over-the-shoulder shots in order to change perspective from one actor to another.	Záběr obrácený o 180 stupňů oproti záběru předchozímu. Používá se jako kombinace několika po sobě jdoucích záběrů přes rameno. Ukazuje tak změny perspektivy od jednoho herce, účastníka dialogu, k druhému.
Room Tone	-	The background sound recording, that enables modification and audio removal in post-production. It is the sound in a location without any added film sounds.	Zvuk v pozadí každé lokace nahraný před natáčením umožňující upravovat nebo vyjmout zvuky při post produkci. Je to zvuk v lokaci před přidáním všech zvuků filmu.

Rough Cut	Hrubý střih	First edition of a film. The film is put together in the order in which it will be projected, but is not yet edited completely.	První střih filmu. Film je sestaven postupně tak, jak bude vypadat při promítání, ale žádné další úpravy ještě nebyly proveden.
Running time	Délka promítání, minutáž	The length of duration of a film. Usually around two hours for full-length movie.	Délka trvání filmu. U celovečerních filmů obvykle okolo dvou hodin.
Scene	Scéna, obraz scénáře	Shot or series of shots put together in order to create a dramatic unit or action. A piece of film's narration.	Záběr nebo soubor záběrů, které dohromady tvoří zápletku nebo akci. Část příběhu filmu.
Score	Nahrávka, partitura	Musical component of film soundtrack. In many cases made by composer for the purposes of a particular film.	Hudební kompozice filmové zvukové stopy. Často bývá vytvořena skladatelem pro účely jediného filmu.
Screening	Projekce, předvádění	Display of a film done in cinema. The film is projected on a screen.	Promítání filmu v kině za použití plátna a projektoru.
Screenplay	Scénář	A text written for film production in which all the dialogues, setting and movements of actors are depicted. Often adaptations of novels.	Text napsaný speciálně pro film popisující všechny dialogy, posazení scén a pohyby herců. Často bývají adaptací jiných děl, zejména románů.

Script	Scénář, rukopis	Script consists of screenplay, shooting script and breakdown script (very precise depiction of sequences of scenes	Text k filmu složený ze scénáře, technického scénáře a členěného scénáře (podrobný popis všech natáčecích dní a sekvencí po sobě jdoucích scén).
Second Unit	-	Secondary crew of a film responsible for shooting scenes of minor importance, such as crowd scenes, sceneries etc.	Druhotné osazenstvo filmového štábu, které se nepodílí na natáčení klíčových scén, ale pouze na davových scénách nebo natáčení scenérií.
Sequel	Pokračování	Film that continues the story of previous film. Its characters and storyline develop.	Příběhové pokračování předchozího filmu. Dochází k vývoji postav I příběhové linie.
Set	Místo natáčení, plac	Exterior or interior setting of a film shoot. It is artificially constructed, but appears natural on screen.	Exteriér nebo interiér sloužící k natáčení filmu. Je vytvořen uměle, ale ve filmu vypadá reálně.
Shot	Záběr, snímek	Smallest unit of film narrative. Reffers to a single frame of film which containes a piece of the story.	Nejmenší jednotka výpravy filmu. Jedná se o jediný snímek, který je částí příběhu filmu.

Slapstick Comedy	Filmová groteska	Laughter in audience is brought up by harmless violence among protagonists. Simple falls of actors or cake throwing is the most common feature of such comedy.	Komedie založená na neškodném násilí mezi protagonisty, které se většinou projevuje házením dortů nebo pády.
Soundtrack	Zvuková stopa, hudba k filmu	Double meaning of this term can either refer to the complete set of sounds used in a film or to a collection of songs used in film.	Tento termín nese dvojí význam. Může to buď být výběr písní, použitý ve filmu vydaný na hudebním nosiči, nebo soubor všech zvuků použitých ve filmu.
Stand-in	Náhradník	A double used during setting up a scene, which physically resembles the actor himself.	Náhradník za herce, využívaný při nastavování scény před začátkem točení.
Storyboard	Bodový scénář, boďák	Sequence of sketches and illustrations used in order to display the looks of each scene. It is used for visual mapping of scene sequences.	Skeče mající za úkol zmapovat a přiblížit vzhled každé scény a návaznost po sobě jdoucích scén.
Stunt	Kaskadérský kousek	Dangerous physical action in a film often performed by stunt double.	Nebezpečná fyzická akce ve filmu spojená často s použitím kaskadéra.

Sword and sorcery	-	A fantasy film genre set in ancient time. Swordfights and witchcraft are common features.	Jeden z fantasy filmových žánrů, který se vyznačuje množstvím soubojů s meči a užitím čarodějnictví.
Take	Záběr	Continuously recorded performance of a scene. Each scene is is shot in several takes and director picks up the best one.	Nepřerušované natáčení scény. Každá scéna je natáčena na několik pokusů a režisér rozhoduje o tom, který vypadá nejlépe.
Teaser Trailer	Napínavý foršpan	Trailer released long time before the actuall release of a film giving a short view of what film looks like.	Filmová ukázka zpřístupněná několik měsíců před samotným filmem, která potenciálním divákům ukazuje krátký přehled toho, jak bude film vypadat.
Theme	Námět	The central idea of a film, its plot and motif.	Hlavní myšlenka filmu, obsahující zápletku a motiv.
Title Role	Titulní úloha	Actor or actress playing main role of a story in a film.	Herec nebo herečka v hlavní roli příběhu filmu.
Trailer	Trailer, Foršpan	Pieces of film composed together and designed as an advertisement for potential viewers. Before premiere of a film, several trailers are released.	Útržky filmu propojené dohromady a tvořící reklamu pro potenciální diváky filmu. Před premiérou filmu bývá vydáno několik trailerů.

Turnaround	-	Film that has become inactive. Usually because the project has been abandoned by a studio.	Filmový projekt, jenž se studio rozhodlo dále nepodporovat.
Two-shot	Dvojdetail	Shot in which two characters, often in dialogue, are in detail from chest up. Used to highlight differences between these two actors.	Záběr, ve kterém jsou dvě postavy v detailu od prsou nahoru. Využívá se ke znázornění rozdílu mezi těmito postavami.
Undercranking	-	During this process a frame rate of a camera is altered (slow down) so that the final film appears to be in fast motion.	Proces při které se zpomalí natáčecí rychlost kamery tak, že výsledný natočený materiál vypadá zrychleně.
Visual Effects	Vizuální efekty	Sub-category of special effects done in post-production. Refers to all the alterations made to the original film.	Podkategorie speciálních efektů doplňovaných při post produkci. Označují se jím všechny změny učiněné na originálním filmu.
Wardrobe Master	Vrchní kostymér	Person responsible for all the costumes and clothing used on set.	Osoba zodpovědná za všechny kostýmy a oblečení použité na place.
White Balance	Vyvážení bílé	Electronic setting of white color balance in a camera.	Elektronické nastavení vyvážení bíle barvy na kameře.

Widescreen (projection)	Široké plátno, širokoúhlý	Projection of a film on a wide screen. Its aspect ratio has to be changed to a wider format.	Promítání filmu na široké plátno, vyžadující změnu poměru stran obrazu původního záznamu filmu.
Word-of-mouth	-	The publicity a film achieves through being mentioned by people who had seen it. Can be positive or negative.	Publicita, kterou filmu získává nezávislé zmiňování jeho kvality nebo nekvality mezi diváky a potenciálními diváky.
XXX	-	The rating used for pornographic films.	Označení používané u pornografických filmů.
Zoom	Transfokátor	A shot which is made by using movement of camera lens in order to either zoom in (get closer) or zoom out (get further) from the object.	Záběr vytvořený za pomoci pohybu přibližovacích čoček kamery. Může dojít k přiblížení nebo oddálení snímaného objektu.

6   ANALYSIS SUMMARY

The practical part of my work consists of corpus analysis and creation of glossary based on such analysis. The corpus I created includes tens of thousands of words and was built by combining five different glossaries available on the Internet. With search conducted throughout this corpus, done by WordSmith software, I was able to pick over one hundred and forty most frequently occurring terms.

By using Internet translators and dictionaries of film terminology in print I was able to translate most of the terms into Czech Language. English explanations given are compilations of several explanations available in the glossaries I used as the bases for my own corpus based glossary. Czech explanations were created by translating the English ones. The whole glossary is a table with four columns: first column contains English term found in two or more of my input glossaries, second column provides Czech translation (where found), third column is meaning explanation in English and fourth column is meaning explanation in Czech.

CONCLUSION

My thesis provides its reader with significant information connected to Corpus linguistics, corpus building process and the use of corpora in glossary construction. This work can be seen as a simlple guide to Corpus linguistics and the possibilities of its utilization, which can be used by amateurs to gain an overview of a modern linguistic method.

Theoretical part of my work is aimed at the explanation of what Corpus linguistics is, whre it came from and how it can be beneficial to its user. The theory I provide is based on works of renowned corpus linguists and should be a summary of the best introductory works available. The processes and research explained in my work are simplified and made more understandable to non-specialist readers. This means that the theoretical part should give any Corpus linguistics amateur an easy way to obtain more information on the topic and its possible practical use.

The link between theory and practice in my thesis can be seen in the explanatory part to the process of glossary creation. Based on this theoretical knowledge, reader should be able to create simple corpora from which glossaries can be built.

The analytical part of my work is an example of how glossary can be derived from corpora. Through the use of computer software a list of most frequently occurent terms is created and using glossary building knowledge gained from specialized literature a glossary with Czech translation and both English and Czech explanation is created. This glossary can further be used by students or translators in order to gain an insight into the field of film terminology.

BIBLIOGRAPHY

Aijmer, Prof. Dr. Karin, ed. 2009. *Corpora and language teaching*. Amsterdam: John Benjamins Publishing Company.

Amott, Lyndsey. How to Write Glossaries. Docsymmetry.com. http://www.docsymmetry.com/glossaries.html(accessed April 19, 2013).

Atkins, B.T. Sue, and Michael Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press, USA.

Beaver, Frank Eugene. 2009. *Dictionary of Film Terms: the Aesthetic Companion to Film Art*. 4th ed. New York: Peter Lang Publishing.

Béjoint, Henri. 2000. *Modern Lexicography: An Introduction*. New York: Oxford University Press, USA.

Bowker, Lynne, and Jennifer Pearson. 2002. *Working with Specialized Language: a Practical Guide to Using Corpora*. London: Routledge.

Bronzite, Dan. A Glossary Of Screenwriting Terms and Filmmaking Definitions. Movie Outline.http://www.movieoutline.com/articles/a-glossary-of-screenwriting-terms-and-filmmaking-definitions.html(accessed April 10, 2013).

Corpus Linguistics: What is Corpus Linguistics? Corpus Linguistics. http://www.cl2011.org.uk/(accessed April 4, 2013).

Cruse, Alan. 2011. *Meaning in language: an introduction to semantics and pragmatics*. 3rd ed. Oxford: Oxford University Press, USA.

Dirks, Tim. Filmsite: Film Terms Glossary Illustrated. Filmsite. http://www.filmsite.org/filmterms.html(accessed April 4, 2013).

Film Land: The Dictionary of Film, Audio and Video Terminology. Film Land Web.http://www.filmland.com/glossary/Dictionary.html (accessed April 5, 2013).

Fort, Karën 2011. Corpus Linguistics: History. Lecture, INIST, France. November 18.http://www.loria.fr/~fortkare/CorpusLinguistics_01_history.pdf (accessed April 3, 2013).

Internet Movie Database: Glossary. IMDb. http://www.imdb.com/glossary/ (accessed April 5, 2013).

Kelly, Shauna. How to Create a Glossary in Microsoft Word. Shauna Kelly.http://www.shaunakelly.com/word/layout/glossary.html (accessed April 19, 2013).

Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics (Studies in Language and Linguistics)*. London: Addison Wesley Longman.

Krieger, Daniel. 2003. Corpus Linguistics: What It Is and How It Can Be Applied to Teaching. *The Internet TESL Journal* 9, no. 3 (March): page nr. http://iteslj.org/Articles/Krieger-Corpus.html (accessed April 10, 2013).

Landau, Sidney I. 2001. *Dictionaries: the art and craft of lexicography*. 2nd ed. Cambridge: Cambridge University Press.

Lipka, Leonhard. 2002. *English Lexicology. Lexical Structure, Word Semantics, and Word-Formation.* Tübingen: Narr.

McEnery, Tony, and Andrew Wilson. 2001. *Corpus Linguistics*. 2nd ed. Edinburgh: Edinburgh University Press.

Mukherjee, Prof. Dr. Joybrato. 2009. A Guide to Corpus Linguistics. diss., University of Giessen. In University of Giessen,http://fss.plone.uni-giessen.de/fss/faculties/f05/engl/ling/help/materials/restricted/clguide.pdf/file/A%20Guide%20to%20Corpus%20Linguistics.pdf(accessed April 8, 2013).

Nesselhauf, Nadja. 2005. Corpus Linguistics: A Practical Introduction. Lecture, University of Heidelberg, Germany. October 2005. http://www.as.uni-heidelberg.de/personen/Nesselhauf/files/Corpus%20Linguistics%20Practical%20Introduction.pdf(accessed April 6, 2013).

Oxford Dictionaries:The World's Most Trusted Dictionaries. Oxford Dictionaries.http://oxforddictionaries.com/definition/english/glossary (accessed April 16, 2013).

Sinclair, John. 2004. Appendix: How to Build a Corpus. In *Developing Linguistic Corpora: a Guide to Good Practice*, ed. Wynne M, 96-101. London: Arts and Humanities Data Service.http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm (accessed April 10, 2013).

Springhurst: A Basic Glossary of Film Terms. The Springhurst Project.http://www.springhurst.org/cinemagic/glossary_terms.htm#Index (accessed April 7, 2013).

Šudová, Barbora. 2009. Role of Language Corpora in Today's Linguistic Study of the English Language: Corpus Linguistics. Bachelor's thesis, University of Tomas Bata in Zlin.

Tognini-Bonelli, Elena. 2001. *Corpus linguistics at work*. Amsterdam: John Benjamins
    Publishing Company.

Vedral, Jiří 2002. *Anglicko-český slovník filmový a televizní* Prague: JTP.