

Filozofická fakulta Univerzity Palackého v Olomouci

Katedra anglistiky a amerikanistiky

Lexical Cohesion in Publicistic Texts:
a Comparative Analysis

Lexikální koheze v publicistických textech:
komparativní analýza

Diplomová práce

Autor: Ivana Tůmová

Angličtina se zaměřením na tlumočení a překlad

Vedoucí práce: Mgr. Ondřej Molnár

Olomouc 2016

*Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně
a uvedla úplný seznam citované a použité literatury.
V Olomouci dne*

.....

I owe many thanks to my supervisor, Mgr. Ondřej Molnár,
for all his patient guidance and advice,
as well as to my parents for their endless support.

LIST OF ABBREVIATIONS

CZ – Czech

EN – English

LIST OF APPENDICES

Appendix 1: Last Apartment In Paris.....	104
Appendix 2: Poslední byt v Paříži.....	106
Appendix 3: Japanese Men Are Happy.....	108
Appendix 4: Proč jsou v Japonsku muži šťastní	109
Appendix 5: Television Moscow-Style.....	111
Appendix 6: Moskevská televize	113
Appendix 7: LBJ Betrayed in Portrait.....	114
Appendix 8: Prezidentův portrét	116

TABLE OF CONTENTS

List of abbreviations.....	4
List of appendices	4
Table of contents	5
List of tables	8
1. Introduction.....	9
2. Cohesion	11
2.1 Parameters of textuality	11
2.2 Cohesion and coherence	13
2.2.1 Cohesive tie	15
2.2.2 Classifications of cohesion.....	16
2.2.2.1 Classification by Halliday and Hasan.....	16
2.2.2.2 Classification by Beaugrande and Dressler	18
2.2.2.3 Classification by Dooley and Levinsohn.....	19
2.3 Lexical cohesion	21
2.3.1 Classifications of lexical cohesion	21
2.3.1.1 Halliday and Hasan’s model.....	21
2.3.1.2 Hasan’s revisited model.....	22
2.3.1.3 Hoey’s model of lexical repetition	23
2.4 Cohesion in Czech and English	25
2.4.1 Lukeš’s classification of cohesive devices in Czech and English.....	26
2.4.1.1 Componential grammatical cohesion	26
2.4.1.2 Organic grammatical cohesion	29
2.4.1.3 Lexical cohesion	29
3. Hoey’s method	32
3.1 Lexical patterns.....	32

3.2	Principles of the method	33
3.3	Analyzing a text.....	35
3.4	Categories of lexical repetition.....	37
3.4.1	Simple repetition	38
3.4.2	Complex repetition.....	39
3.4.3	Excluding chance repetition	40
3.4.4	Simple paraphrase	41
3.4.5	Complex paraphrase	41
3.4.6	Hyponymy, hyperonymy and co-reference	42
3.4.7	Repetition with non-lexical items	43
3.4.8	A note on ellipsis.....	44
3.5	Limitations and constraints of the method.....	44
3.6	Hoey's method applied	45
3.6.1	Berber Sardinha.....	45
3.6.2	Oliveira, Ahmad and Gillam	46
3.6.3	Károly.....	47
3.6.4	Molnár	47
4.	Characteristics of the texts	49
4.1	Column in the English tradition.....	49
4.2	Publicistic style and the Czech functionalist tradition.....	50
4.3	Sample texts.....	51
4.3.1	The author	51
4.3.2	The translator.....	51
4.3.3	Texts used for the analysis	52
5.	Research design	53
5.1	Research questions.....	53

5.2	Adjusting the sample texts.....	54
5.2.1	Omissions and sentence borders	55
5.2.2	Adding the reference	56
5.3	Using Hoey's method of analysis	57
5.3.1	Four categories of lexical cohesion	57
5.3.2	Open-set lexical items only?	58
5.3.3	Proper nouns.....	59
5.3.4	Negation in Czech	60
5.3.5	Gradation	61
5.3.6	Subjectivity of analysis	61
6.	The analysis	62
6.1	Analyzing the texts	62
6.1.1	The limitations.....	62
6.1.2	Eliminating the mistakes	63
6.2	Japanese Men Are Happy (Proč jsou v Japonsku muži šťastní).....	66
6.3	The Last Apartment In Paris (Poslední byt v Paříži).....	72
6.4	Television Moscow-Style (Moskevská televize).....	81
6.5	LBJ Betrayed in Portrait (Prezidentův portrét).....	88
6.6	Summary.....	95
7.	Conclusion	100
	Appendices.....	104
	Summary	118
	List of references.....	123
	Annotation.....	126

LIST OF TABLES

Table 1: Dooley and Levinsohn's classification of cohesive devices.....	20
Table 2: Hasan's categories of lexical cohesion	22
Table 3: Repetition mechanisms according to Hoey.....	30
Table 4: A quantitative matrix	36
Table 5: Record of links in the four categories and bonds with particular sentences	64
Table 6: Example of a quantitative matrix.....	65
Table 7: Quantitative matrix – Japanese Men Are Happy.....	68
Table 8: Quantitative matrix – Proč jsou v Japonsku muži šťastní.....	70
Table 9: Summary of links and bonds – Japanese Men Are Happy (Proč jsou v Japonsku muži šťastní)	72
Table 10: Quantitative matrix – The Last Apartment In Paris.....	75
Table 11: Quantitative matrix – Poslední byt v Paříži.....	78
Table 12: Summary of links and bonds – The Last Apartment In Paris (Poslední byt v Paříži).....	80
Table 13: Quantitative matrix – Television Moscow-Style.....	84
Table 14: Quantitative matrix – Moskevská televize.....	86
Table 15: Summary of links and bonds – Television Moscow-Style (Moskevská televize)	88
Table 16: Quantitative matrix – LBJ Betrayed in Portrait.....	91
Table 17: Quantitative matrix – Prezidentův portrét.....	93
Table 18: Summary of links and bonds – LBJ Betrayed in Portrait (Prezidentův portrét)	94
Table 19: Summary of links and bonds in sample texts.....	96

1. INTRODUCTION

This Master's Diploma Thesis aims to investigate the means of lexical cohesion in Czech and English publicistic texts. A method presented by Michael Hoey serves as the methodological framework for the analysis.

The theoretical part of the thesis presents cohesion as one of the parameters of textuality, cohesive tie as an instance of cohesion and follows with various classifications of cohesive devices. The first classification presented is the most widely used one by Halliday and Hasan. Alternative classifications by Beaugrande and Dressler and Dooley and Levinsohn follow. In the next section, the focus is drawn on lexical cohesion and the classification of its means. It also starts with Halliday and Hasan's classification and Hasan's revisited model. The last classification described in this section is Hoey's classification of means of lexical repetition in text.

This part also includes a more detailed description of Hoey's method. According to Hoey, repetition is the most frequent and most important means of cohesion. The method uses the number of links and bonds created by lexical items on the basis of simple or complex lexical repetition and paraphrase as the basis for mapping the cohesive ties within a text. The method was originally created for the analysis of specialized texts; however, it has been applied to various fields, as listed in this section. The last part of this chapter explains the limitations of the applicability of Hoey's method.

The last chapter of the theoretical part serves for characterization of sample texts – columns by Art Buchwald and their Czech translations by Jan Jiráček. Column writing has a different tradition in Anglophone countries, where column has a prominent position among the news writing, from that in the Czech newspapers, where it is much less frequent. The Czech classification therefore follows the tradition of functional styles and includes a column as one of the genres of the publicistic functional style.

The practical part describes the research project and its results. The thesis aims to test the applicability of Hoey's method to expressive texts. The analysis applies this method of lexical patterns in text to four pairs of equivalent texts in Czech

and English. On the basis of the analysis, the thesis aims to classify the differences in number and type of the cohesive ties made by the means of lexical cohesion in Czech and English.

The first chapter of the practical part presents the design of the research project and the changes that were necessary for the application of Hoey's method. The texts had to be adjusted so that they were identical – there has to be the same number of sentences and their boundaries must be equivalent as well. Also, Czech does not have to express the subject in a sentence: it is indicated by the verbal suffix or inferred from the context. For the analysis, these subjects were inserted into the text, so that they can be included in the co-reference chain in Czech as well. Other problematic areas, such as negation, gradation or the classification of proper names are also discussed in this chapter. The hypotheses for the research project are stated in this chapter. The last part of this chapter deals with the issue of subjectivity in the analysis.

The next chapter focuses on the analysis itself. It starts by stating the limitation of the research project – that is, the possibility of mistakes in the analysis, and the mechanisms that help eliminate these mistakes. The next part of this chapter discusses the sample texts one by one. The specific problems that needed to be tackled when analyzing them are described and the results for each sample text are provided in the form of a quantitative matrix. The English and Czech version are compared at each sample text. The last part of this chapter sums up the results of the analysis, states the validity of the hypotheses, discusses the findings of the research project and states tentative conclusions.

The last part of the thesis, Chapter 7, provides a concluding overview of the thesis and the final remarks on the research project, its design, limitations and results. It also presents suggestions for further research in this field.

2. COHESION

This chapter deals with the phenomenon of cohesion. It overviews some of the approaches towards cohesion, provides its definition and description, as well as three classifications of cohesive devices – a widely accepted classification by Halliday and Hasan (1976), the one by Beaugrande and Dressler (1981) and Dooley and Levinsohn's (2000) classification. The chapter follows with a more detailed description of lexical cohesion and its classifications by Hasan and Hoey. The last part of the chapter discusses comparatively cohesive devices in Czech and English.

2.1 Parameters of textuality

In linguistics, a text refers to any passage, spoken or written, which constitutes a unified whole (Halliday and Hasan 1976). Halliday and Hasan (ibid.) follow with a description of a text as a unit of language in use. Its nature is semantic; a text is a unit of meaning, not a structural unit. This means that a text does not consist of sentences – it is not a unit of above-sentence level; it is realized by them.

According to Halliday and Hasan (1976), a stretch of speech or writing must meet certain parameters in order to be considered a text. They introduce the concept of texture as a crucial property of every text. Texture is therefore the property that distinguishes a text from a non-text (ibid.). Beaugrande and Dressler (1981, 3) define a text as a “communicative occurrence which meets seven standards of textuality” and they subsequently list the following seven parameters constituting textuality: situationality, informativity, intentionality, acceptability, intextuality, cohesion and coherence. Tárnyiková (2002) defines these standards as follows:

situationality – a cover term for all the factors that connect the text with a relevant situation (i.e. a whole complex of relevant factors making up the context of situation or situational context)

informativity – a notion applied to the content of the text, its semantic load and the way it can be recovered

intentionality – a language configuration must be intended to be a text

acceptability – [a language configuration must be] accepted as such

intertextuality – perceived in two ways: first, as our previous experience with other texts of a similar kind; second, as “the way in which one text echoes or refers to another text”

cohesion – a property of text primarily associated with the surface structure of text, to be distinguished from coherence, representing an underlying connectedness of the text

coherence – the underlying logical/semantic connectedness of the text units, based on such concepts and relations as cause-and-effect, sequencing of events into identical temporal frames, logical deduction, entailment, prediction etc. (Tárnyiková 2002, 29-71, [emphasis added])

All the seven standards of textuality always work in interplay. Tárnyiková (2002) emphasizes that a text is not independent of its context: there is always a connection between a linguistic form and the purpose or function a text serves.

The parameter of *situationality* connects a text with its situational context. As Tárnyiková (2002) points out, a text is both situation-bound and situation-determined. The context of situation in which a text occurs influences its semantic load and formal properties, as well as the way the information is presented and interpreted by the recipient (*ibid.*). This is true especially with spoken texts, where exophoric reference is interpreted according to the situation. The context of situation is also important for the distinction of literal and non-literal meaning: according to the situation a recipient decides whether a literal interpretation is possible and adequate or not (*ibid.*).

Informativity concerns information density of a text, as well as givenness or predictability or a text event. Informativity is established by the proportion between the items contributing to thematic development and the items indicating vagueness, such as hedges, discourse markers etc. which do not contribute

to informativity of a text (Tárnyiková 2002). Information density of a text varies according to genres and text types – e.g. the information density of an advertisement is high, while phatic communication has low degree of information density (ibid.). The recoverability of information is a pragmatic phenomenon depending on the recipient's "general knowledge" (ibid.).

Intentionality and *acceptability* are two parameters differentiating between a text and a non-text, as a text must be intended to be a text and accepted as such as well. Intentionality is thus connected to text production, while acceptability is based on the attitude of the recipient (Tárnyiková 2002). Beaugrande and Dressler (1981) state that these two parameters involve certain degree of tolerance towards deficiency in cohesion and coherence of a text, as long as it is perceived as a piece of purposeful communication. Tárnyiková (2002) shares this view. As she remarks, acceptability must not be confused with grammaticality: a text may be ungrammatical, but acceptable and, on the contrary, a grammatically correct stretch of speech or writing may be unacceptable for its recipient.

Intertextuality concerns either general expectations about a text based on previous experience with the particular text type, its typical structure, the standardized way of presentation of information etc., or a particular instance of one text echoing another, explicit or implicit (Tárnyiková 2002). This type of intertextuality "establish[es] a relation to a cultural tradition by placing a given text within the cultural framework and adding cultural value to a text" (ibid., 69). It is a phenomenon used for creating parodies: the knowledge of the original text (called "pre-text" by Tárnyiková) is a prerequisite for understanding correctly the mocking echo (called the "active text") (ibid.).

The last two parameters, *cohesion* and *coherence* are dealt with in greater detail in the following section.

2.2 Cohesion and coherence

The present thesis deals with cohesive relations in a text. The following section presents the phenomenon of cohesion and its relation to coherence in greater detail. The sub-sections deal with cohesive ties and classifications of cohesion respectively.

Cohesion means the interrelation between the items of a text: the interpretation of one item depends on other items (Halliday and Hasan 1976). Tárnyiková (2002) uses Daneš's (1985) definition of cohesion as manifestation of isotopic relations that substantially contribute to the inner connectivity of the text. As Beaugrande and Dressler (1981) put it, cohesion means mutual connectedness of surface components of a text with accordance to grammatical forms and conventions. In other words, "cohesion refers to the range of possibilities that exist for linking something with what has gone before" (Halliday and Hasan 1976, 10).

Hoey (1991, 3) defines cohesion as "the way certain words or grammatical features of a sentence can connect that sentence to its predecessors (and successors) in a text." From this, it is clear that Hoey puts emphasis on the fact that cohesive devices can establish connections *among sentences*: this fact is the fundamental one for Hoey's method of analysis of lexical patterns in text, as described later.

Cohesion and coherence are two very closely interrelated concepts. Tárnyiková (2002) points out Halliday and Hasan's distinction between cohesion and coherence, the former being property of the surface level of text contributing to the surface linkage of text elements, the latter referring to the underlying semantic connections of text items and their link to the real world. As Hoey (2001, 51) puts it, "coherence of a text is reflected in and signaled by the cohesion in the text." Beaugrande and Dressler (1981), who list cohesion and coherence as two of seven parameters of textuality, describe the difference between the two as follows:

[cohesion] concerns the ways in which the components of the surface text, i.e. the actual words we hear or see, are *mutually connected within a sequence*, [while coherence] concerns the ways in which the components of textual world, i.e. the configuration of concepts and relations which *underlie* the surface text, are *mutually accessible and relevant*. (1981, 3-4, [original emphasis])

Another difference between the two concepts lies in their objectivity. As Hasan (1984) remarks, coherence is not an absolute concept: the assessment of a text as coherent or incoherent always depends on its recipient. Hoey (1991) reflects this difference as well: he claims that while cohesion is a phenomenon that may be

objectively recognized, coherence is subjective and every recipient may thus perceive and assess it differently. As Tanskanen (2006, 21) puts it, “[c]ohesion can be regarded as a property of the text, while coherence depends upon the communicators’ evaluation of the text.”

It is important to note that a text may be coherent without any explicit cohesive ties on the one hand, and incoherent although cohesive on the other – in other words, “cohesive ties are not by themselves criterial of coherence” (Hoey 1991, 12), as the interpretation of the meaning of a text does not depend on text-internal relations only, but on the context of situation as well. This corresponds with the discourse perspective, which is built upon the view of text as more than only a sum of sentences (ibid.).

As mentioned by Halliday and Hasan (1976), cohesion may be realized through grammar or vocabulary. This is reflected in their classification where they list lexical cohesion as a separate category. As Halliday and Hasan (ibid.) remark, cohesive devices may work both within a sentence and among sentences. They note that structure is one of the means of creating texture. All grammatical units are therefore internally “cohesive” because they are structured (ibid., [original quotation marks]). But a text is not a structural unit: it normally extends beyond one sentence. Cohesion in this dimension therefore refers to non-structural text-forming relations (ibid.). These relations are semantic in nature: cohesion occurs whenever the interpretation of one item is governed by other items of a text – that is, by items in the co-text, or even by phenomena outside the text – in the context of situation. However, as Halliday and Hasan (ibid.) point out, in this case the relation created is exophoric and thus not cohesive, as it does not fulfill the property of linking two items within a text.

2.2.1 *Cohesive tie*

An instance of cohesion is called a cohesive tie (Halliday and Hasan 1976). Cohesive ties may be either semantic or grammatical in their nature (ibid.). Cohesion is not a structural relation and thus can spread across sentence boundaries (ibid.). This, however, does not mean that cohesive relations occur between adjacent sentences only: cohesive ties create a complex network throughout the text.

2.2.2 *Classifications of cohesion*

As mentioned above, approaches to textuality, cohesion and coherence differ. There are thus also various classifications of cohesive devices. This section presents three of them significant for the present thesis.

2.2.2.1 *Classification by Halliday and Hasan*

Halliday and Hasan (1976) set five categories of text connectedness: reference, substitution, ellipsis, conjunction and lexical cohesion. These five categories can be summarized as follows:

Reference: two linguistic elements are related in what they refer to.

Substitution: a linguistic item is not repeated but is replaced by a substitution item.

Ellipsis: one of the identical linguistic elements is omitted.

Conjunction: a semantic relation is explicitly marked.

Lexical cohesion: two elements share a lexical field.

(Sanders and Pander Maat 2006, 591, [emphasis added])

The list above reflects Halliday and Hasan's classification of cohesive devices as grammatical or lexical (with the first four categories being grammatical in nature). Halliday and Hasan (1976) define reference as a property of a linguistic element whose result is that this element is not interpreted on its own but depends on another item for its interpretation. One entity can thus enter a discourse multiple times through reference. This entity an element refers to can be text-internal (endophoric reference) or text-external (exophoric reference). As noted earlier, exophoric reference is not regarded as contributing to cohesion of a text, as it does not link text-internal items (ibid.). Endophoric reference may be illustrated by the following example¹:

Example 1: So does the young French girl. ... But Miss Schneider,
a child of the French bourgeoisie...

¹ The following examples are taken from the sample text used for the analysis.

Substitution, according to Halliday and Hasan (1976, 88), differs from reference in that “substitution is a relation in the wording rather than in the meaning” operating on the lexico-grammatical level. Ellipsis is then treated as a kind of substitution – “a substitution by zero” (ibid.). The following two examples illustrate substitution and ellipsis respectively:

Example 2: He *wants* the apartment in the worst way. *So does* the young French girl.

Example 3: No wonder American husbands *are irritable and hard to get along with at the end of the day*. You would *be*, too, ...

The cohesive force of conjunctions lies, as Halliday and Hasan (1976) state, in their specific meanings. Conjunctions presuppose there are other parts present in the discourse linked by them (ibid.).

Example 4: Contrast this to the average American home where the American wife not only refuses to bow to her husband when he comes home, *but* in some cases won't even give him a bath. *And* when she does give him a bath ...

The last category is labelled lexical cohesion by Halliday and Hasan (1976). As the label suggests, this category deals with lexical items creating cohesive ties in a text. In other words, the choice of vocabulary also contributes to cohesion of a text (ibid.). Halliday and Hasan do not deal with lexical cohesion in such depth as with the previous four categories, although they note some important facts about the nature of this phenomenon. They also note that the basis of this type of cohesion is repetition (ibid.).

Example 5: What are the major differences between the American and *Japanese woman*? For one thing, the *Japanese woman* is much more concerned about the welfare of her husband.

Halliday and Hasan's classification of the devices of lexical cohesion is presented later, in the section 2.3.1.1.

Halliday and Hasan's classification has been nevertheless criticized. Widdowson (2004), for instance, remarks that Halliday and Hasan do not clearly state the

differences in cohesive function related to the differences between the types of cohesion in their classification. There are many bordercases that may fall into either the category of substitution or reference, as Halliday and Hasan claim, but in Widdowson's view there is a lack of reasoning behind the distinction of these two categories. He also uses the examples from Halliday and Hasan to show that their distinction between reference and substitution is not clearly set. Also, Halliday and Hasan define reference as a relation on the semantic level and substitution as a relation on the lexicogramatical level – but cohesion deals with semantic items. This distinction is thus in Widdowson's view not clear and as he suggests, if based on this definition, substitution should not be classified as a type of cohesion. Another point Widdowson disagrees with is the label "reference", as this phenomenon is based on the relation of two items sharing the referent and therefore, as he suggests, should be labeled "co-reference". He also criticizes the approach Halliday and Hasan take to exclude situational parameters of text from their investigation, that is, to exclude context. This means that coherence cannot be investigated, as it relates the text to extra-textual phenomena. (Widdowson 2004, 63-72)

Hoey (1991, 5) also comments on Halliday and Hasan's classification – in his view, the category of conjunction should be "better treated as a part of a larger system of semantic relations between clauses". Also, he points out that "reference does not *mark* semantic relations; it *is* a semantic relation and occurs whenever an item indicates that the identity of what is being talked about can be retrieved from the immediate context" (ibid., [original emphasis]). Hoey discusses the difference between substitution and reference as well. In his view, substitution is a rather rare phenomenon and the importance of the distinction between the two categories for textual analysis "is not readily apparent" (ibid., 6).

2.2.2.2 Classification by Beaugrande and Dressler

Beaugrande and Dressler (1981) offer an alternative classification of cohesion taking a functional approach. They divide cohesive devices into two groups: (1) devices increasing stability – recurrence and partial recurrence, parallelism and paraphrase and (2) devices of reduction – pro-forms and ellipsis. They also list

devices signaling relations among events and situations in text – tense, aspect, junction, functional sentence perspective and intonation and stress.

This classification also reflects the importance of recurrence, i.e. repetition, as the basis of cohesive ties. Unlike Halliday and Hasan, whose classification was criticized for unclear correspondence between types of cohesion and their functions and distinctions, Beaugrande and Dressler's (1981) approach classifies the devices according to the function they serve in the text – they either repeat what was said before (recurrence), or fulfill the principle of language economy (reduction). Also, they connect a text with its context, as shown by the relation-signaling devices.

Beaugrande and Dressler's (1981) category of partial recurrence corresponds to Hoey's (1991) category of complex repetition (for Hoey's classification see section 2.3.1.3), as it refers to the repetition of two items that share the same basic word-component but are of different parts of speech. Their category of paraphrase corresponds to Hoey's simple paraphrase, as it covers synonymy.

2.2.2.3 Classification by Dooley and Levinsohn

Another alternative classification is provided by Dooley and Levinsohn (2000). They define six categories of cohesive devices. The categories defined by Halliday and Hasan are included, but Dooley and Levinsohn account also for the devices signaling relations mentioned by Beaugrande and Dressler, such as inflectional forms or intonation. Table 1 lists Dooley and Levinsohn's classification (Dooley and Levinsohn 2000, 13):

Cohesive devices:

1. descriptive expressions alluding to entities mentioned earlier

2. identity

- repetition
- lexical replacement
- pronouns
- other pro-forms
- substitution
- ellipsis

3. lexical relations

- hyponymy
- part-whole
- collocation

4. morphosyntactic patterns

- consistency of inflectional categories
- echoic utterances
- collocation

5. signals of relations between propositions

6. intonation patterns

Table 1: Dooley and Levinsohn's classification of cohesive devices

As illustrated in this chapter, cohesion refers to the means of linking parts of a text together. It is a complex phenomenon that can be approached from different perspectives: various classifications of cohesive devices are possible according to the functional perspective (such as the presented classification of Beaugrande and Dressler 1981) or the language level on which a cohesive tie is established (such as Halliday and Hasan 1976). The present thesis aims to investigate one particular type of cohesion – lexical cohesion. This particular type of cohesion is dealt with in detail in the following chapter.

2.3 Lexical cohesion

As already mentioned, the devices of cohesion may be grammatical or lexical in nature. The importance of lexical items and their relation to the overall meaning of a text is pointed out cf. by Sinclair (2004). Lexical cohesion is defined by Halliday and Hasan (1976, 274) as “cohesive effect achieved by the selection of vocabulary”. In their work, they nevertheless do not investigate this field into greater depth (cf. Christiansen 2011; Hoey 1991). The model of lexical cohesion they propose was later revisited by Hasan (1984). According to Morley (2009), lexical cohesion is created by lexical items sharing the same semantic field and helps to tie the text together. This area of investigation can, as Christiansen (2011) points out, be difficult and tricky, as theoretical definitions may fail to be valid for all cases of real language use.

2.3.1 *Classifications of lexical cohesion*

The following section presents the two classifications mentioned above, as well as Hoey’s classification, which is the basis for his model of lexical cohesion analysis.

2.3.1.1 *Halliday and Hasan’s model*

Halliday and Hasan divide the devices of lexical cohesion into two groups: (1) reiteration and (2) collocation. Reiteration covers repetition of a word, use of synonym (such as flat – apartment), superordinate (wash – take a shower/bathe) or general word (brutality – rape)² (Halliday and Hasan 1976). They do not, however, deal with lexical cohesion in greater detail. For instance, it is not clear from their description what can or cannot be regarded as synonymy.

Halliday and Hasan’s usage of the term “collocation” is discussed by Hoey (1991, 7-8): he remarks that in spite of a clear description of collocation as “a relationship a lexical item has with items that appear with greater than random probability in [textual] context”, some of their examples of this type of lexical cohesion are unclear. As Hoey (ibid.) adds, intuition as a tool for the assessment

² The examples are taken from the sample texts.

whether or not a particular case should be counted as collocation is not particularly reliable.

2.3.1.2 *Hasan's revisited model*

Hasan worked further on the classification of lexical cohesion and revisited largely the model described above. In particular, she devoted her attention to the classification of what was called “collocation” in the above mentioned model and incorporated some sub-categories that in the original model were treated simply as the cases of “collocation” (Hoey 1991). Her model of lexical cohesion consists of two groups of categories – general and instantial.

General: repetition (realize – realizes)

 synonymy (movie – film)

 antonymy (hot – cold)

 hyponymy (wash – bathe)

 meronymy (flat – room)

Instantial equivalence (aging American – Marlon Brando)

 naming (film – “Last Tango In Paris”)

 semblance (as romantic as a TV dinner)

Table 2: Hasan's categories of lexical cohesion (Hasan 1984, 202)

The conclusion that Hoey (1991) derives from Hasan's work is that the number of cohesive ties in a text does not have to correspond to the “level” of coherence as assessed by the text's recipient. In other words, it is not only the number of links that determines the degree of cohesiveness established by the items of a text: Hoey (1991) classifies means of lexical cohesion not only according to the type of link established, but also according to its strength, as explained in detail in Chapter 3.

2.3.1.3 Hoey's model of lexical repetition

As Hoey (1991) states, the study of cohesion means the study of lexis and its patterns. In his view, lexical repetition contributes significantly to the organization of a text. But as Molnár in his study based on Hoey's methodology remarks, text organization "does not imply that each text is imprisoned in a straitjacket of fixed structures. It rather shows a recurrent incidence of traceable patterns as networks in each text that are conducive to its comprehension, i.e. to coherence" (Molnár 2012, 31). These patterns can be investigated on the basis of Hoey's classification. Lexical repetition is divided by Hoey (1991, 53-65) into four categories: simple lexical repetition, complex lexical repetition, simple paraphrase and complex paraphrase.

Simple repetition occurs when an item is repeated within a text. This category allows for alteration within the scope of a closed grammatical paradigm. Singular and plural forms of the same lexical item, for instance, are treated as simple repetition. The category of simple repetition is restricted only to open-set lexical items. Hoey (1991) thus excludes for example articles, auxiliaries or prepositions – that is, grammatical words.

Example 6: American *women* could learn a lot from *Japanese* wives. ... For one thing, the *Japanese woman* is much more concerned about the welfare of her *husband*. ... Another area in which *Japanese women* excel is giving their *husbands* baths.³

Words that share a lexical morpheme but either differ in form or are formally identical but have different grammatical function fall into the category of *complex repetition* (Hoey 1991). For example, a verb and a noun derived from it will be treated as a case of complex repetition – even in case of zero derivation, when a verb and a noun are formally the same.

Example 7: ... giving their husbands *baths*. ... anyone who has been *bathed* by a Japanese woman ...

Simple paraphrase "occurs whenever a lexical item may substitute for another in context without loss or gain in specificity and with no discernible change

³ The examples in this section are taken from the sample texts.

in meaning” (Hoey 1991, 62). In other words, this category covers synonymy. Hoey also emphasizes that the assessment of a relation and its classification depend highly on context (ibid.).

Example 8: He wants the *apartment* in the worst way. ... They meet by accident in the empty *flat* ...

Complex paraphrase is the most complex of the four categories – and the most difficult one for assessment as well. Hoey himself claims it to be “a can of lexical worms” (1991, 64). This category covers the cases “when two lexical items are definable such that one of the items includes the other, although they share no lexical morpheme” (ibid.). This covers cases of antonymy where the two antonyms do not have the same lexical morpheme and cases of “link triangle”.

Example 9: ... they will be considered *inferior*. ... A wife who knows how to bathe her husband in the Japanese style is a *superior* person ...

The link triangle occurs either when two items of a text have a relation to the same third item present in the text or when this third item is not present in the text. The first case occurs for instance if a text contains expressions “wash”, “bathe” and “take a shower” – “bathe” and “take a shower” may be regarded as co-hyponyms and “wash” may be regarded as a hyperonymic expression to the former two. Such a connection is established even if one of the three words is not explicitly present in the text. In other words, if two relations with an item are present (between items A, B and A,C), the third relation is created as well (between items B and C) (ibid.). These relations can refer to hyponymic or superordinate repetition, where the hyponym and hyperonym have the same referent. This leads to co-reference and substitution.

Co-referential chains, established by multiple links created by co-reference, such as “the movie” – “Last Tango” – “it”, also belong to the category of complex paraphrase.⁴ In general, Hoey excludes pronominal forms of the first and the second person from entering the co-referential chain. In this research, however, this rule is not followed. The reasoning for this choice, as well as more detailed

⁴ Hoey (1991) avoids the label „co-referential chain“, although it is a widely accepted term. He nevertheless includes such instances.

description of Hoey's methodology and its application for this research task, follow later.

2.4 Cohesion in Czech and English

Cohesion is a phenomenon all languages have in common. There are, however, different devices used to realize cohesive ties (Lukeš 2004).

According to Lukeš (2004), there are three main factors influencing the manifestations of cohesion in various languages: (1) difference in grammatical type, (2) different conventions on the application of similar rules and (3) different conceptualization of the discourse world.

The first case, the difference in grammatical type, refers to systemic differences between languages; for instance, the grammatical category of definiteness differs in Czech and English. While this category is present in English, Czech lacks it completely and uses other means instead, such as demonstratives (Lukeš 2004).

In the second case, the languages concerned share a grammatical category, but its use in these languages differs. An example could be the usage of passive in Czech and English: in Czech, the usage of passive is rather limited to the academic writing (Lukeš 2004). As Knittlová (2000) remarks, the possibility of its usage in Czech is determined by the functional style and the genre of a text.

The third category covers the usage of discourse connectives. As Lukeš (2004) puts it, the difference between Czech and English in this respect lies in the way these languages express logical relationships, as well as attitude of the transmitter. While Czech has means to express both explicitly, "English tends to limit itself to stating the logical relationship only and it is more implicit in comparison to Czech"⁵ (Lukeš 2004, 4).

It is pointed out that while in general, cohesion is understood to facilitate the comprehension of a text; different languages appear to have different demands on what is the "correct" or sufficient level of cohesion. This suggests that cohesion should be treated not simply as the linkage of parts of a text, but "as a function

⁵ Lukeš supports this claim by examples from the field of second language learning, where Czech students of English tend to overuse connectors such as "therefore" and "so", while the writing of English students of Czech tends to lack sufficient connectives expressed explicitly.

of mental representations” (ibid.) and thus is closely related to the concept of coherence.

Taking the previous remarks into account, the present thesis aims to investigate comparatively how Czech and English differ in establishing cohesive ties. The following section lists particular differences in this field as described by Lukeš (ibid.).

2.4.1 Lukeš’s classification of cohesive devices in Czech and English

Lukeš (2004)⁶ presents an overview of the means Czech and English use to establish cohesion. In his classification, cohesive devices are divided into three groups: (1) componential grammatical cohesion, (2) organic grammatical cohesion and (3) lexical cohesion.⁷

2.4.1.1 Componential grammatical cohesion

The first category, componential grammatical cohesion, consists of the following cohesive devices: pronominal anaphora and co-reference, definiteness, negation and substitution and ellipsis.

As Lukeš points out, there is a difference in the use of third person pronouns. In Czech, there is prevalence of zero pronoun forms. The category of gender is often expressed by the verbal suffix: in the past tense forms, there is no ambiguity in gender. In English, the co-reference of “it” is ambiguous (as the scope of reference may vary: for instance in “Mary broke a vase. It was terrible.”⁸ the reference is ambiguous between the vase and the situation as a whole). The use of zero pronoun forms in Czech “predicts that the antecedent must be highly accessible”. In English, zero pronouns can be used only “in extremely high accessibility contexts”. Full pronouns in combination with word order can be used for emphasis in Czech, while in English the means of emphasizing are stress and intonation. The difference between the usage of reflexive pronouns versus possessive pronouns in the third person serves for disambiguation of reference

⁶ The data collected by Lukeš were “identified in the literature, excerpted from real life data and identified in a small parallel corpus (4000 words of source language; referred to as pilot corpus to distinguish it from my main corpus) compiled preliminary to my research.” (Lukeš 2004, 12)

⁷ The main source for this chapter and its sub-sections is Lukeš 2004. All citations, paraphrases and examples, unless indicated otherwise, come from Lukeš 2004, 12-33.

⁸ my own example

in Czech. This is not possible in English. The relative pronoun “který” can be used for both human and non-human antecedent in Czech, while English differentiates between “what” and “who”.

This difference is very important for the text analysis: due to the tendency described above, the Czech and English version of a text may differ in the way the subject is expressed. If the subject is not expressed explicitly in Czech and is inferred from the co-text, context or according to the form of the particular verb, while in English the subject is explicitly present, the two versions are different in the number of the instances of a particular item, which may lead to difference in the number of links. The two versions thus may need adjustments in order to present the information in the same way. This issue, as well as Hoey’s reasoning for the adjustments, is discussed in detail later with regard to the research design in Chapter 5.

The category of definiteness is represented by the usage of articles and determiners in English and is connected to the notion of context-new and context-old information. In Czech, there are different means of expressing this notion, as the category of definiteness is missing – word order, based on the principle of functional sentence perspective, and lexical means, such as the usage of demonstratives. Unlike Czech, English uses possessive pronouns for body parts, family members and the like (Lukeš illustrates this by the equivalents “my father” and “otec”). Also, Czech uses dative in some cases where possessive is used in English (such as in “I broke his arm.” versus “Zlomil jsem mu ruku.” mentioned by Lukeš). This category is not of particular interest for the thesis, as Hoey’s method is applied (prevalently) to open-set lexical items.

Czech and English also differ in the way they express negation. The most obvious difference is the fact that Czech allows for two negative items in a clause (as in Czech “nic tam není” versus English “there is nothing in there” or “there is not anything in there”). The use of negation influences the communicative dynamism of a sentence. “The distribution of degrees of communicative dynamism over the sentence elements [i.e. in the written form of communication] ... is an outcome of the interplay of three factors: context, semantic structure [and] linear arrangement.” (Firbas 1971, 138) Apart from the means of word order, the requirements of functional sentence perspective (FSP) can be fulfilled by stress

and by lexical means. The functional sentence perspective in Czech, due to relatively free word order, is the leading principle for the distribution of information in a sentence according to the principle of linearity. In English, due to “relatively fixed word order”, the leading principle is grammatical. As for the thesis, the most important fact concerning negation is the systemic difference between the two languages – synthetic versus analytic means of expressing negation will be further discussed in Chapter 5, as they affect the application of Hoey’s method.

Another field Lukeš (2004, 19) remarks upon is the fact that “Czech has a vast array of indefinite pronouns at its disposal but they seem to be optional and mostly used for emphasis.” In some instances, Czech underspecifies the reference (one of his examples is “tu pruhovanou” versus “the striped one”). Also, Czech, unlike English, must deal with the fact that different prepositions need their head to occur in different case and therefore cannot use ellipsis in these cases (e.g. “before and after the lesson” versus “před hodinou a po ní”). English uses substitution of verbal forms by “do” and “so”. Czech lacks these forms and must therefore employ other means. In addition, “[b]oth languages also have a number of elliptical constructions which are lexical in nature and may or may not have an equivalent. ... Czech has a whole range of systemic elliptical constructions following modal verbs ... which are not available in English.” (This may be illustrated by his example of Czech “Otec musí na schůzi.” where English cannot omit the verb.)

This specific use of reference in Czech, as well as the differences concerning substitution and ellipsis, must be assessed in the sample texts before the analysis as well, as they may also result in different numbers of links in Czech and English versions of the analyzed texts. The above mentioned prepositional phrases that require different cases of their head are, however, ignored, as the repetition of the head occurs within a sentence and therefore does not establish multiple links (for the reasoning and a detailed explanation of the method see Chapter 3, Hoey’s method).

2.4.1.2 *Organic grammatical cohesion*

This category “covers all devices whose main function is to provide cohesive links”. It deals with both syndetic and asyndetic connections. The reason for this classification of conjunctions is that Lukeš adopts the view of Halliday and Hasan, who exclude them from the first category. In their view,

“[c]onjunctive elements are cohesive not in themselves but indirectly by virtue of their specific meanings; they are not primarily devices for reaching out into the preceding (or following) text, but they express certain meanings which presuppose the presence of other components in the discourse.” (Halliday and Hasan 1976, 226)

The difference between Czech and English in this respect lies in the usage of asyndetic connection. As already stated, Czech tends to explicitly state the relation of clauses and thus prefers to use connectives in cases where asyndetic connection is used in English. Czech also tends to express logical relations as well as attitude of the transmitter, while English expresses the former only.

As it is clear from the previous brief description of Hoey’s model (in section 2.3.1.3), this category is not accounted for in Hoey’s system of analysis, the reason being again the concern with open-class lexical items.

2.4.1.3 *Lexical cohesion*

In this area, Lukeš points out the difference in conventions between Czech and English with respect to repetition. English appears to be less sensitive to repetition of lexical items, while Czech prefers to reiterate an item by means of synonymy, generalization or hyperonymy instead of repetition. The reason for this difference may be the fact that English as an analytic language is forced to repeat many words for grammatical reasons.

Lukeš also refers to Hoey’s approach to repetition as the basic cohesive mechanism and lists his chart of repetition mechanisms:

simple lexical repetition (writer – writers)

complex lexical repetition (writer – writings, happy – unhappy)

simple mutual paraphrase (writer – author)

simple partial paraphrase (writings – volume)

antonymous complex paraphrase (author – plagiarist)

other complex paraphrase (author – writings)

substitution

co-reference

ellipsis

Table 3: Repetition mechanisms according to Hoey (Hoey 1991, 65, 83 in Lukeš 2004, 30)

In the list above, the types of repetition are listed according to their respective “strength”: simple lexical repetition is regarded by Hoey (1991) as the strongest type of means of lexical cohesion, ellipsis being the weakest form of it.

The above mentioned tendency suggests that the English texts have greater amount of items classified as Hoey’s strongest two categories (simple and complex lexical repetition) than the Czech ones. From this fact it may thus be implied that English is more cohesive than Czech in this respect. This is an interesting fact for the present thesis, as it deals with the patterns of repetition in texts and aims to investigate the amount and prevalent type of devices of lexical cohesion present in Czech and English version of the sample texts.

(The last category Lukeš describes is the availability of lexical items in Czech and English. Obviously, there are items that do not have and equivalent in the other language: Lukeš illustrates this difference e.g. by “respectively” which lacks a direct equivalent in Czech. Here, Lukeš also points out the differences in poetic language – in Czech, the sanctions for rhymes are much stricter than in English and, on the contrary, Czechs are much less sensitive to alliterations than the English. Another area in which the two languages differ is that of verbal puns. In Czech, these are “virtually absent from prose and especially from expository texts such as newspapers and academic writing”. These are, however, not of interest to

this thesis, as it does not aim to discuss the differences in lexical cohesion with respect to the quality of translation.)

Chapter 2 aimed to describe cohesion as one of seven parameters of textuality. As demonstrated, there are various approaches towards cohesion and various classifications. Lexical cohesion, the sub-type of cohesion based on the relations of lexical items that was dealt with in greater detail in this chapter, is the basis for Hoey's methodology described in the following chapter.

3. HOEY'S METHOD

As the previous chapter demonstrates, there are various approaches towards cohesion. Various classifications of cohesive devices have been provided, as well as a more detailed account of one type – lexical cohesion. One such approach towards lexical cohesion is presented by Michael Hoey.

In his publication *Patterns of Lexis in Text* (1991), Hoey presents his method for investigating the devices of lexical cohesion and patterns created within a text. The interplay of these devices creating cohesive patterns can be observed on the basis of repetition of lexical items in a text. Hoey's analysis thus serves as a tool for examining the manifestations of lexical cohesion in a particular text. The aim of the method is to create readable and coherent abridgements of non-narrative texts resulting from the analysis of a text from the point of view of lexical cohesion.

The following chapter presents basic principles of the method: it defines terms "link" and "bond" and the application of the method for creating abridgements. The four categories presented earlier are described in greater detail. The last part of the chapter mentions Hoey's method applied in various research projects.

3.1 Lexical patterns

As mentioned in detail in chapter 2, cohesion refers to the surface linkage of items of a text. According to Hoey (1991), lexical cohesion is the most frequently occurring and productive subtype of cohesion. Hoey (ibid.) presents a system of text analysis on the basis of lexical cohesion. The analysis serves as a tool for creating abridgements. The fundamental basis of Hoey's methodology is the view that lexical items creating networks within texts on the basis of repetition establish a kind of structure within the text and help facilitate comprehension to the recipient (Molnár 2012).

Apart from his own research, Hoey (1991) also supports his claim that cohesion is based on lexical repetition by the results of Halliday and Hasan's (1976) analyses. These show that over forty per cent of cohesive ties (that is, the cohesive linkage of two items of a text) classified by them are created by the means of lexical cohesion and if conjunction is discounted, the devices of lexical cohesion form

almost fifty per cent. Hoey (1991) also points out that the number should in fact be even higher, as if one item is repeated multiple times, every instance has a relationship with all the other ones. In other words, “lexical cohesion is the only type of cohesion that regularly forms multiple relationships” (ibid., 10).

3.2 Principles of the method

Cohesive devices manifest surface-level relationships between items of a text. The term “cohesive tie”, as defined Halliday and Hasan (1976), is used for instances of cohesive connection. Hoey accepts this view, but in the context of his method or lexical cohesion analysis, he labels these relationships as “links”. Hoey (1991) explains reasons for not using “ties” in the case of a connection between two lexical items – firstly, the term is used by Halliday and Hasan to label instances of cohesive relations *in general*. Secondly, Hoey (1991) perceives “tie” as being directional, while “link” in his view lacks this directionality and denotes that the two items “share” a connection.

One item in a sentence can create links with items in multiple sentences at once. The only restriction on these is that the same item in one sentence cannot create links with two items of another sentence. An “item” does not have to be one word only – in some cases, a phrase may repeat what was said by one word previously etc. (Hoey 1991).

According to Hoey (1991), three instances of a link create a “bond”. His reasoning for this three-link criterion establishing a significant connection between two sentences is that the analysis is based on lexical patterning and an isolated link is therefore not important. As Hoey (ibid., 91) puts it, “[s]ince we are interested in repetition that appears to serve some text-organizing function, we will only concentrate on those cases of linkage that, within a text, show an above-average degree of connection.” The number three, however, is not an absolute value. Hoey (ibid.) recommends adjusting the number to a particular text, although he does not give any definite value according to text typology or genre. Also, the above mentioned criterion of three links allows for excluding the marginal and doubtful cases of links (as explained in detail later) at this step of analysis, when identified links are clustered into bonds. If any arguable case of bond is still identified, it must be marked in the results (ibid.).

On the basis of the analysis, the sentences building a text can be assessed according to the number of bonds they enter as “*marginal*” or “*central*”. Central sentences are those that show a high amount of repetition, i.e. they have a high number of connections with other sentences. Central sentences are important for the thematic development of a text. Marginal sentences have fewer connections with other sentences. They have lower information value, as the information they present is not essential for the text and its use (Hoey 1991). From this, it is clear that central sentences carry the most important information presented in the text and they can thus be used in order to create its abridgement. It is also possible to trace thematic development in a text. The number of bonds a sentence has with previous and following sentences indicates its relation to the topic: a *topic-opening* sentence is bonded to a high number of following sentences and if it is not the first sentence of a text, while a *topic-closing* sentence is bonded to a high number of previous sentences. In other words, a topic-opening sentence introduces a topic discussed later in the text and a topic-closing sentence closes the topic. A sentence may have a status of both a topic-opening and a topic-closing sentence: in such case, it closes one topic and starts another.

Hoey states four points summarizing the essential principles underlying his method:

1. If two sentences of a text are appropriately connected by the means of repetition, they make sense together or display some relationship regardless of the distance between them.
2. The sentences sharing lexical items can be used to observe the thematic development within the text.
3. The method allows for tracing the sentences opening and closing a particular topic.
4. The number of links a sentence enters indicates its centrality or marginality within the text.

(Hoey 1991, 75)

The method is aimed to provide abridgements for non-narrative texts. Ahmad and Benbrahim (1995) confirm the applicability of Hoey’s analysis of lexical patterns

as the basis for text processing systems that generate “executive summaries”. The terms “abridgement” or “summary” refer to “brief statements of the main ideas in a text or passage” (Longman Dictionary 2010, 582). The reason for creating summaries is clear: sparing the recipient from reading the whole text without the possibility to check its usefulness beforehand. The summary should be significantly shorter than the original text. Molnár (2012) for instance adopts 38 % as the maximum length of an abridgement.

Hoey (1991) proposes three possible methods for creating an abridgement. The first method is based on deletion of marginal sentences. The remaining sentences build a readable summary of the original text. An alternative method creates a shortened version by taking only the central sentences. The third method takes a topical sentence and all sentences linked to it to create a summary. Here, Hoey (ibid.) points out that it may be necessary to supply conjunctions. Also, the verbal voice and modality of the sentences may need to be adjusted in order to make a coherent, intelligible abridgement. (ibid.) The choice of the appropriate method depends on a particular text. Molnár (2012), for instance, demonstrates that the first method is inappropriate for a text with low number of marginal sentences, as the abridgement is too long.

3.3 Analyzing a text

As the previous section explains, the analysis of lexical patterns in a text is not the aim of Hoey’s method – it is a tool for creating abridgements. This section briefly summarizes the procedure of analysis of a text according to Hoey’s (1991) methodology. Subsequently, it discusses the types of output the analysis provides and the principles for their interpretation.

The following four steps summarize the procedure of the analysis of a text according to Hoey’s method:

1. The analyzed text may need to be modified in order to supply reference before the analysis. (ibid.)
2. The sentences are numbered. A sentence is marked orthographically: it starts with a capital letter and ends with a punctuation mark, usually a full stop. (ibid.)
3. The links are identified and later clustered into bonds.

4. The output of the analysis are matrices, topological diagrams and two-figure coordinates documenting networks of bonds within the text. (ibid.)

The horizontal column of a matrix shows the relations of a sentence with the preceding ones, while the vertical column relates it to the following ones. An itemized matrix shows particular cases identified as links, while a quantitative matrix shows the number of occurrences and thus allows for identification of the density of links with particular sentences.

The following quantitative matrix shows the number of bonds between the first ten sentences of the English version of the sample text *Last Tango In Paris*.

	1									
2		2								
3		1	3							
4		1		4						
5					5					
6						6				
7						1	7			
8					1			8		
9							1		9	
10					1					10

Table 4: A quantitative matrix

In topological diagrams, the position at the top of the diagram means the initial sentence. A diagram thus graphically represents the network of cohesive connections within a text created by individual bonds between the sentences. (These are represented by their number.) An example of a topological diagram was taken from Molnár (2012):

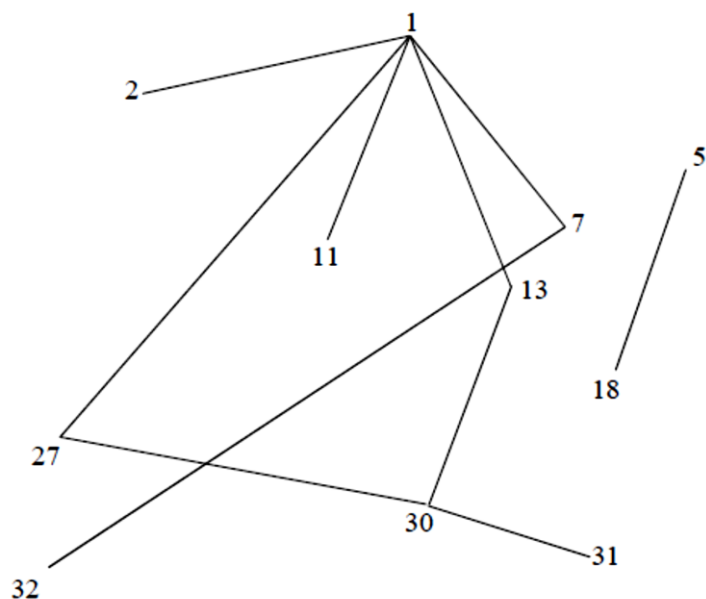


Figure 1: A topological diagram

The two-figure coordinates give two numbers for each sentence, the first indicating the number of previous sentences bonded to the particular sentence, the second showing the number of bonds with following sentences.

For example, a coordinate **22** (4,4) shows that the sentence 22 shares four bonds with previous sentences and four bonds with the following sentences.

When observing the two-number coordinate indicating the bonds a sentence is part of, topic-opening and topic-closing sentences may be identified as well (ibid.), the former having a high value of the first number within the coordinate, the latter being those with a high second number. All these serve as sources for the interpretation or lexical patterning.

3.4 Categories of lexical repetition

As already presented, Hoey's method is based on examination of one type of cohesive devices within a text – the devices of lexical cohesion. This subtype of cohesion, as Halliday and Hasan (1976) state, is manifested through the choice of vocabulary: it is the repetition of lexical items that creates lexical cohesion.

Hoey (1991) defines four basic types of repetition of lexical items: simple and complex repetition and simple and complex paraphrase. He also discusses other possible ways lexical cohesion can be manifested in a text. The following section

discusses them in greater detail. It also provides reasoning for including or excluding certain cases, as explained by Hoey. The adjustments of the method for the present study are discussed later in Chapter 5.

The order of the categories is not random. The assessment phase of the analysis is based on the view that certain types of lexical repetition are “stronger” than others. As already mentioned, Hoey (*ibid.*, 83) lists them in this order:

1. simple lexical repetition
2. complex lexical repetition
3. simple mutual paraphrase
4. simple partial paraphrase
5. antonymous complex paraphrase
6. other complex paraphrase
7. substitution
8. co-reference
9. ellipsis

The method “give[s] priority to lexical links over grammatical links” (Hoey 1991, 84). For this reason, if there are both repetition and substitution present between two sentences, the former is counted and the latter is ignored, as well as the cases of sentence-internal links (*ibid.*).

3.4.1 Simple repetition

The least complicated of the four categories is simple (lexical) repetition. Within this category, a link is established between two tokens of a type (Hoey 1991). It occurs if an item is repeated in a form that differs only within a closed grammatical paradigm. The link can be established between two open-set lexical items only. Multiple occurrences of prepositions, conjunctions and the like are thus ignored.

In the following example for instance, the repetition of “woman” and “Japanese” are classified as simple repetition, while the two occurrences of “the” are ignored.

Example 10: What are the major differences between the American and *Japanese woman*? For one thing, the *Japanese woman* ...

This category is, however, not unproblematic. The problem lies in the fact that the same word may have different meaning in each occurrence. Hoey (1991) does not require a complete identity of meaning between two occurrences: the requirement of shared collocational environment in these cases should be sufficient to exclude two occurrences of an item having significantly different meaning. This requirement excludes polysemous words used in two different meanings as well.

An example of such case could be a transitive verb “sanction”, which has two opposing meanings – (1) permit or approve for something and (2) impose a penalty on someone.⁹

Another problem with this category is that there are “doubtful cases” that are neither purely lexical nor grammatical. An instance of this doubtful case may be an adverbial such as “sometimes”. There is no reason for not treating this item as lexical, if repeated. The reason Hoey (ibid.) gives is that a single link is not of much importance in isolation; a cluster of links is significant. These doubtful links therefore usually cease to matter once the bonds are identified. The way these cases were accounted for in the study is discussed later, in Chapter 5.

3.4.2 *Complex repetition*

The second category is complex (lexical) repetition. This type of repetition occurs if two items “can be paraphrased in the context of the text in which they appear in such a way as to ensure that the paraphrase of one includes the other” (Hoey 1991, 55).

Again, it does not have to be an exact paraphrase. Hoey requires only “close approximation in a context” (ibid., 56), emphasizing that two words that in general have different collocational profiles may still be counted as an instance of complex repetition in the context of a text, if they fulfil this requirement of close approximation within that particular text.

An example of complex repetition from a sample text follows:

⁹ The two definitions were taken from the Oxford Dictionaries website. <http://www.oxforddictionaries.com/definition/english/sanction>

Example 11: ... anyone who has been *bathed* by a Japanese woman will never want to take an American shower again. Unlike the American woman, a Japanese wife looks forward to giving her husband a *bath*.

Here, “bathed” and “bath” are instances of complex repetition. One of them can be paraphrased by the other, as in “anyone who has been given a bath by a Japanese woman” and “looks forward to bathe her husband”.

3.4.3 *Excluding chance repetition*

There is always a possibility that a word is repeated in a text, but the instances share no cohesive relation – that is, “the only common ground is the choice of the same lexical item” (Hoey 1991). Such cases are called “chance repetition”. Hoey (ibid.) repeatedly emphasizes that although it is useful to distinguish between text-forming and chance repetition, it is not vital, as a single link has no significance in the analysis, as it is “naturally eliminated by [its] failure to be part of clusters of links” (ibid.). He states that chance repetition forms only a small and insignificant amount of all the instances of lexical repetition and these instances are eliminated in the analysis (ibid.).

These two types of repetition can be distinguished on the basis of two criteria: common reference and contextual criterion. If two items refer to the same entity, they cannot form a chance repetition. On the other hand, if two items do not refer to the same entity, they do not automatically form an instance of chance repetition. According to the contextual criterion, two items do not share a chance repetition if they have common or related contexts, share common relationship with neighboring lexical items or when there is whole or partial parallelism between their contexts (ibid.). Here, Hoey remarks that contextual factors are always subjective and their judgement may vary. Also, the contextual criterion does not always lead to a clear answer – this is another case of the doubtful, marginal link (ibid.).

Hoey (1991) also points out that the link triangle is in operation here: if an item is repeated for three or more times, a link with one of the previous occurrences implies links with all the other occurrences as well.

3.4.4 *Simple paraphrase*

The category of simple paraphrase includes synonyms. But as Hoey (1991) emphasizes, synonyms in this sense are understood to be dependent on the context of a particular text (unlike Hasan [1984] who treats synonymy as a general relation). The condition for counting two items as synonyms is that there must be no loss or gain in meaning – and again, Hoey (ibid.) notes that subjective judgement plays a role here. In the case of simple paraphrase Hoey recommends excluding doubtful instances rather than including them in the analysis. There are two sub-types of simple paraphrase – partial and mutual (ibid.). Partial “works in one direction only” (ibid., 62).

An example from a sample text follows:

Example 12: They meet by accident in an empty *flat* and you see Brando’s mind working. He figures if he rapes the girl, she’ll go away and he’ll get the *apartment*.

The two items do not have to be in exactly the same form in order to be counted as simple paraphrase, as long as this difference in form lies within a grammatical paradigm – for instance, if one of the synonymous adjectives occurs with the –ed suffix, while the other has the –ing ending. This case, however, requires shared collocational characteristics of the two items.

3.4.5 *Complex paraphrase*

The last category is also the most complex and complicated one. The instance of complex paraphrase occurs if “one of the items includes the other” (Hoey 1991, 64). There are three sub-types of complex paraphrase – antonymy, link triangle where all three items are present in the text and link triangle where two words share a link via an a common item not present in the text.

Most antonyms, as Hoey (1991) points out, can be counted as instances of complex repetition. These are antonyms sharing a morpheme, such as “happy – unhappy”. Antonyms that do not share a morpheme are counted as instances of complex paraphrase, such as “hot water – cold floor” taken from the sample text.

The second type of complex paraphrase occurs, if an item has a link with two different items. These two items are then counted as sharing a link as well. In other words, two links establish the third one (Hoey 1991). For instance, if “woman” and “wife” are taken as synonymous in the particular context, and if “wife” and “husband” are taken as forming a link, then “woman” and “husband” share a link as well.

The third situation involves cases when the item connecting the two others is not present in the text. The condition in these cases is that “there must be an item that is capable of paraphrasing exactly *in that context* one of the items and of repeating the other.” (Hoey 1991, 66, [original emphasis]) An example of this relationship may be the link “declare – statement” taken from the sample text. The word “state” would be an instance complex repetition of “statement” and would paraphrase “declare” in the particular context. Thus, these two items establish a link.

3.4.6 *Hyponymy, hyperonymy and co-reference*

There are also other cases that Hoey counts in his system of analysis. One possibility is that an item is followed by a more general one with the same referent. A more complicated situation arises if it is followed by a more specific one, as explained below:

If the more general word follows, it cannot be said to supply any information that was not already contained in the earlier item. If it also meets the contextual criterion which we applied to lexical repetition and paraphrase, there are no principled grounds for refusing to include the relationship between these items within the general category of repetition. (Hoey 1991, 69)

Thus, for instance, “Last Tango In Paris” and “the movie” establish a link of this kind.

If, however, the items are in reverse order, the meaning of the latter is more specific than the preceding one. The link can then be counted only in two cases: either if the link triangle occurs or if the two items have an identical referent. Hoey (1991) explains that if this limitation was not accepted, then in a text where

“man” and “woman” appear at the beginning, all subsequent references to male and female beings human beings would have to be taken as establishing a link, which would undermine the results of the analysis.

3.4.7 *Repetition with non-lexical items*

Although Hoey’s model focuses on the means of lexical repetition, there are also items not lexical in nature, which fulfill the same function as lexical items and are therefore counted as well. These non-lexical items, despite being included, are rather peripheral to the analysis (Hoey 1991).

The first group contains personal pronouns. In spite of the fact that pronouns are not lexical items, they can serve the same function and as such they are treated as lexical items and therefore not excluded in the analysis, although these links are also taken as “doubtful cases” (Hoey 1991, 71). Hoey (ibid.) remarks that only third person singular and plural pronouns can enter these relations. First and second person pronouns refer “out of the text”, so they are counted only in quotations. This, as he notes, applies to non-narrative texts.

The same applies to demonstrative pronouns “this”, “that”, “these”, “those” – they are treated as lexical items, which enables them to enter links. The corresponding demonstrative modifiers are, however, excluded, unless they modify a noun head which is not part of lexical repetition or paraphrase (Hoey 1991). Otherwise, the same link would be counted twice. For the same reason, the definite article is ignored (ibid.).

Another group consists of various other cases of substitution. The substitute “one”, if occurring in the form of a premodified nominal head, is included. Another substitute that is not excluded is “do” and its variations, such as “do so”, “do it” or “do the same”. The third one is clausal “so”. Here, Hoey (1991) points out that its negative counterpart, clausal “not”, is ignored, as he considers this to be the case of ellipsis.

The last group Hoey (1991) mentions includes expressions “another”, “the other”, “(the) same”, “different” and “similar”. The first three, if accompanying a nominal head, are treated in the same way as demonstrative modifiers: they are included only if their head establishes no link in order to avoid counting the same

link twice (ibid.). If there is no head accompanying these three expressions, then “the same” is treated like “this” and the other two are understood as signaling ellipsis. The remaining two items, “different” and “similar”, “are treated in the same way as ‘other’ and ‘same’, except that they are also analyzed as lexical items capable of entering into repetition relations” (ibid., 74).

3.4.8 *A note on ellipsis*

In several places, Hoey mentions ellipsis. This phenomenon is defined as “the absence of some required stretch of language that has to be supplied by the listener or reader to make sense of the sentence encountered” (Hoey 1991, 74). Ellipsis is of importance when creating the abridgements of texts, as it indicates a reference to another part of the text and thus must be dealt with in order to create an intelligible text summary: it may be necessary to add the information referred to by means of ellipsis or to adjust the form of the particular sentence (ibid.).

3.5 Limitations and constraints of the method

There are three major issues concerning Hoey’s system of analysis. Firstly, the method was intended as a basis for automatic computational recognition of lexical cohesion. Hoey himself (1991) grants that while some cases are relatively unproblematic, the identification of links according to contextual parameters is beyond the scope of automatic computational method. This is as well the case with polysemy, in which case a word has two or more mutually related meanings (Longman Dictionary 2010). Also, some of the categories, as mentioned before, are based on individual’s subjective assessment, which is also impossible to be implemented by a computer software. These issues pose an obstacle for unproblematic usage of Hoey’s method as the basis for automatic text segmentation softwares (Hoey 1991). Examples of computer segmentation software based on Hoey’s analysis, as well as the adjustments necessary for its application, are presented in section 3.6.

Secondly, Hoey’s method is intended for analysis of non-narrative texts serving as the basis for creating abridgements. Narrative texts differ considerably from non-narrative texts and the method may therefore be inapplicable to them – the most obvious reason being that there is no use of creating abridgements of narrative

texts (Hoey 1991). The sample texts used for the present research project, columns by Art Buchwald, are aimed to serve for testing the possibility of the use of Hoey's method on publicistic texts. Here, the primary reason for using Hoey's method is not to create summaries, but to use it as a method for a comparative analysis of the devices of lexical cohesion in two language versions of the same texts.

Also, the systemic differences of languages may play an important role in the applicability of the method. Hoey's classification is based on the system of English language, which differs considerably from the other language used for the research project, Czech. There are supposed to be differences in results of the present research project on the basis of this systemic difference. (The differences in cohesive devices in Czech and English are described in detail in section 2.4).

3.6 Hoey's method applied

This section presents four research projects applying Hoey's methodology. The first paper by Berber Sardinha (1999) presents the application of Hoey's method on text segmentation. The focus of the research is segmentation of a text performed by LSM system for statistical analysis of a text in comparison to random segmentation. The second research project, conducted by Oliveira, Ahmad and Gillam (2002), is focused on the usability of summaries produced by SummariserPort program in comparison to random summaries. Károly's study (2010) deals with repetition as a means of cohesion and the role it has in a discourse. Hoey's method is contrastively applied to a corpus of Hungarian – English news texts. It investigates shifts in cohesion and coherence caused by translation of a text. Molnár (2012) applies Hoey's method on selected texts published by the EU. He aims to contrastively assess the amount of lexical cohesion in Czech and English versions of the texts and its means. Also, he investigates the reliability of the method for producing abridgments in the context of this text type.

3.6.1 *Berber Sardinha*

The paper presented by Berber Sardinha (1999) reports on a research project dealing with the development and application of a computer model for discourse

analysis on the basis of text segmentation. The computer model, called LSM (Link Set Median), was tested on a corpus of 300 texts of three different genres – encyclopedia article, business report and research article – and the results performed by the computer were assessed against random segmentation (ibid.). Hoey's method for analysis of lexical patterns in a text was used as the basis for the computational method for segmentation of texts according to "link sets" (ibid.). The aim was to test the ability of this software to divide a text into sections that would correspond to those by the text's authors (ibid.).

Statistical data prove that that LSM system outperforms random segmentation significantly. The project thus confirms the applicability of Hoey's methodology in the field of automatic text segmentation. Also, as Berber Sardinha remarks, the research project emphasizes the importance of lexical cohesion as the basis for text segmentation (Berber Sardinha 1999).

3.6.2 Oliveira, Ahmad and Gillam

In their research project, Oliveira, Ahmad and Gillam (2002) focus on pragmatic behavior of key terms in a text as the means of automatic generation of text summaries (ibid.). Their aim was to test summaries produced by the program SummariserPort against random summaries.

SummariserPort uses lexical cohesion as the basis for the selection of sentences suitable for a summary. The summaries are created on the basis of two of Hoey's four categories of lexical repetition – simple and complex lexical repetition. These two categories are according to Hoey (1991) relatively uncomplicated and therefore suitable for application by computer software. Apart from testing the performance of the above mentioned program, the research thus also tests the applicability of Hoey's method to narrative texts, in particular to 623 financial news texts.

The results of the research confirm the reliability of SummariserPort as a device for creating high-quality summaries. Also, the paper empirically tests and confirms the applicability of Hoey's method to narrative texts of this type (Oliveira, Ahmad and Gillam 2002).

3.6.3 *Károly*

Károly's (2010) research deals with the discursual role of repetition, in particular the influence of shifts in repetition patterns on coherence shifts. The research takes a text-linguistic, discourse-based approach, identifying instances of different combinations of lexical repetition within a source and a target text and examining higher level coherence shifts caused by them (ibid.). It intends to study the impact these differences have on the meaning of translated texts.

The research project is focused on these two issues: (1) it investigates whether it is possible to identify shifts in lexical repetition in non-literary translation based on frequency, quality and patterns of repetition links and bonds and (2) examines whether overt shifts of repetition lead to covert shifts of coherence in translation. The hypotheses were tested upon a corpus consisting of 40 news articles on politics and economy in two sub-corpora (Hungarian source texts and English target texts). Frequency, quality and combinations of repetitions were observed in the texts in order to study shifts in cohesion and coherence in translation.

The results of the analysis show that the number of bonds between the same sentences can differ in the source and the target text, as well as between different sentences and that a sentence that is central in one of the two text versions may be marginal in the other and vice versa (Károly 2010). As Károly (ibid.) concludes, the results of the analysis support the hypothesis that the shifts of lexical repetition in translation may affect the macropropositional structure of a target text. She also remarks that "Hoey's (1991) repetition model is capable of describing particular shifts of cohesion and coherence in translation" (ibid., 62).

Károly's study points out that there may be differences between source and target texts arising from translation. This claim is of interest to this study, as it supports its preliminary assumptions about the differences in lexical cohesion between two language versions of a text.

3.6.4 *Molnár*

In his paper, Molnár (2012) tests some of Hoey's claims and investigates lexical patterns in the texts of the European Union, in particular the Czech and the English versions of three official EU texts. The research tests Hoey's claims

on the semantic connection between bonded sentences, two of the methods for creating abridgements (namely omitting the marginal sentences and creating a summary from central sentences only) and the position of topic-opening and topic-closing sentences within a text. Also, the research serves as a test of the applicability of Hoey's method to the Czech language (ibid.).

The results suggest that the Czech versions of the analyzed texts established fewer bonds than their English counterparts. In other words, the English texts were more cohesive than the Czech ones (Molnár 2012). As Molnár (ibid.) explains, the reasons for this fact are systemic, as well as stylistic: Czech is a synthetic language with preference for avoiding repetition, while English is an analytic language which prefers simple and complex lexical repetition.

Also, the results proved that for this particular text type the first method mentioned earlier is not valid, as the texts have high information density and only few marginal sentences occur. A summary produced by omitting marginal sentences would therefore be too long (Molnár 2012).

Molnár's findings support the assumption that the means of lexical cohesion in the sample texts used for this study will prove to differ. The systemic and stylistic differences between Czech and English seem to be a plausible reason for these differences, notwithstanding the possible influence of the style of the author and the translator.

4. CHARACTERISTICS OF THE TEXTS

The following chapter serves as the theoretical background for the texts used in the research part of this thesis. The first part provides a theoretical description of a column in the tradition of Anglo-Saxon culture on the one side and within the publicistic style (“publicistický funkční styl” in Czech) in accordance with the Czech tradition of functional styles on the other. The reason for such a difference in classification is that column writing has a much deeper-rooted tradition in English than in Czech. The last section presents a description of sample texts, their author and translator into Czech.

4.1 Column in the English tradition

The style of newspaper reporting is one of the so called pragmatic styles, which represent the reality according to political attitudes of the news writers and political or cultural focus of the newspaper (Oakland and Urbanová 2002). The central function of newspaper is to inform the reader, to report. Crystal and Davy (1969, 174) list the following general features of the language of newspaper reporting: “the need for compression of the information into a limited space, the need for clarity, the avoidance of ambiguity”.

As McNair (2008) points out, the work of a journalist is, however, not only about reporting, but also about analyzing, interpreting and commenting on news and events. Column writing involves presenting author’s opinions as well. As McNair remarks, newspaper writers (and especially those who write columns) have significant power over their readers once they establish themselves as a discursive authority, i.e. when the people take their writings as a reliable and trustworthy source of information. This sense of trust in the author arises either from their status or prominence, or from their rhetorical skill; the author can persuade the reader by his/her “elegant, erudite or witty prose” (ibid., 107).

Thematic focus of a column may vary considerably. As McNair (2008) states, topics of a column may range from polemical ones, “address[ing] the reader in tones raging from the counter-intuitive and the skeptical ... to the indignant and even the outraged” to an analytical-advisory column, presenting author’s insight into an issue or his/her advice, or the satirical column (ibid., 110).

4.2 Publicistic style and the Czech functionalist tradition

The publicistic style is the “youngest” one within the functional classification. It was first defined and described as late as 1957 by Jelínek. It shares many features with other functional styles – it is similar to the scientific style in its logical structure and the usage of connectors, as well as to the poetic style in the emotional appeal and the usage of stylistic devices (Knittlová 2000). Knittlová (ibid.) divides the style into three sub-categories: (1) informative (reporting), (2) publicistic (analytic) and (3) literary (fiction)¹⁰.

The characteristic feature of this style is the use of figurative language. As Chloupek (1993) remarks, publicistic style makes use of common similes, metaphor and metonymy, as well as biblical and classical idiomatic expressions and allusions. A more comprehensive description may be found in Čechová et al. (2008), who list the following features: figurative speech, terms, idioms and collocations, clichés, proverbs, general expressions and allusions. As they point out, there are two opposing tendencies in their use – automatization and foregrounding.

Chloupek (1993) defines the functions of publicistic style as follows: (1) building awareness, attracting the reader, promotion of an idea/event/product, (2) specific reporting in the fields of politics, economics, production, sports and culture and (3) information service to the reader. In other words, using Bühler’s typology, the publicistic style has the informative, as well as operative function (Čechová et al. 2008). The aim of publicistic texts is to inform the readers of varying social status, age and education quickly, precisely and effectively, to convey ideas in a comprehensibly, unambiguously, immediately and convincingly and to attract the readers and persuade them about the attitudes represented in the text (ibid.).

Column as a genre is in the Czech tradition described as follows: a short and brief unit of written publicistic style, formally creating one column on the page. It serves for the author’s reflection of reality. As a particular insight into life it does not aim to solve the most important issues, nor does it have to present the most significant contemporary information. It is, nevertheless, up-to-date, as it serves

¹⁰ This classification is used by other authors as well, although the labels of the three categories differ: Čechová et al. (2008, p. 245) for instance label them “reporting, analytic and publicistic-fictional”.

as an immediate reaction of the author. It is usually specific, sometimes humorous, funny or critical (Čechová et al. 2008).

4.3 Sample texts

The sample texts used for the research project are columns written by Art Buchwald. They come from the collection *Have I Ever Lied To You?* published 1968 and its Czech version *Copak jsem vám někdy lhal?* published 1992. This section provides basic information about the author and the translator and follows with the overview of the texts used for the analysis as well as the reasoning for their choice.

4.3.1 *The author*

Arthur “Art” Buchwald¹¹ (1925-2007) was an American humorist and satirist, best known for his columns written for The Washington Post. The topics range from contemporary domestic political issues and international political scene to social issues and everyday life.

Due to a turbulent and unstable family background, he was brought up by foster parents. In the WWII he volunteered for the US Marine Corps. After the war he enrolled at the University of Southern California, but eventually left without graduating his degree. He received an honorary doctorate from the school in 1993. Soon after leaving the school, he started working for the European edition of New York Herald Tribune. Soon he gained popularity on the both sides of the Atlantic and his career was on the rise. He was awarded the Pulitzer Prize in 1982. In the later period, his popularity began to decrease: some even claimed his columns are tiresome and not funny anymore. Buchwald died in 2007 of kidney failure.

4.3.2 *The translator*

Jan Jiráček¹² (*1958) is a Czech translator. His translations from English are mostly film screenplays, fiction and scientific publications on media studies. His translations of fiction include John Grisham’s *The Chamber (Cela smrti)*,

¹¹ source of the biography: https://en.wikipedia.org/wiki/Art_Buchwald and <http://www.theguardian.com/media/2007/jan/19/guardianobituaries.pressandpublishing>

¹² source of the biography: http://www.obecprekladatelu.cz/_ftp/DUP/J/JirakJan.htm

John Updike's *The Witches of Eastwick* (*Čarodějky z Eastwicku*) or Kurt Vonnegut's *Bluebeard* (*Modrovous*).

4.3.3 *Texts used for the analysis*

The aim of the thesis being the investigation of lexical cohesion in Czech and English, as well as testing the applicability of Hoey's method to publicistic texts, the genre of column was an outright choice.

The particular author was chosen on the basis of the fact that a whole collection of his columns was translated into Czech. This was seen as an advantage for two reasons – firstly, the author was famous and highly appreciated, which suggests that the sample texts are not of doubtful quality, and secondly, all the texts were translated by the same person, so the chance of “interference” of the style of various translators was thus limited.

When choosing the particular texts for the analysis, the main criterion was the “fidelity” of translation. In some cases, probably due to the gap between the time of publication and time of translation, as well as the cultural gap between the two cultures, some passages were added, omitted or adjusted by the translator in the Czech texts. To allow for a comparative analysis, the texts that differ the least from each other in the two languages were chosen for the analysis. In some cases, adjustments in both English and Czech versions were necessary. These are discussed in detail in section 5.2.

The texts chosen for the research project are the English and the Czech versions of four Buchwald's columns: *The Last Apartment In Paris/Poslední byt v Paříži*, *Japanese Men Are Happy/Proč jsou v Japonsku muži šťastní*, *Television Moscow-Style/Moskevská televize* and *LBJ Betrayed in Portrait/Prezidentův portrét*.

5. RESEARCH DESIGN

The second part of the thesis is devoted to the practical research based on Hoey's methodology, using the necessary theoretical background presented in the previous chapters. This chapter describes the research design. It formulates the research questions and discusses the procedure of the research regarding the adjustments made to the texts and Hoey's method for this particular use. The limitations of the research are explained.

The following chapters present the research itself, the former presenting the analysis of texts and the latter summing up and interpreting the results.

5.1 Research questions

As the theoretical background presented in the previous chapters exemplifies, lexical cohesion, a type of cohesion established through the choice of vocabulary, is a pervasive and important phenomenon and one of the seven parameters of textuality (as discussed earlier in Chapter 2). The means of establishing lexical cohesion vary from language to language. Hoey's method is one of the approaches used for the investigation of lexical cohesion in texts. The presented research project assesses comparatively the means of lexical cohesion in English columns by Art Buchwald and their Czech translations by means of Hoey's methodology.

The research is aimed at answering the following three research questions regarding the devices of lexical cohesion in Czech and English columns.

1. The applicability of the methodology for analysis presented by Hoey is less problematic in English than in Czech.

This research question is based on the assumption that the systemic differences between Czech and English may cause problems when applying the method to Czech texts, as the two languages differ both in the grammatical categories present in the language system and in their manifestation. One of the areas assumed to be problematic is affixation in Czech, for instance when negating verbs.

2. The English versions of the sample texts contain more instances of bonds than their Czech counterparts (strong claim.)

The Czech versions of the texts contain more links, but fewer bonds than the English versions (weak claim).

This assumption is based on the previous research, in particular on the findings of Molnár (2012), who remarks that in his research of EU texts, the Czech versions contained more links but fewer bonds. The reason was that in Czech some of the sentences were divided into two. This fact influenced the choice of vocabulary in these sentences. The number of instances of lexical cohesion was thus affected, as vocabulary is the basis for lexical cohesion.

The same problem applies here, as in the Czech translations of the sample texts some of the sentences were divided as well.

3. The English versions of the sample texts show more instances of simple repetition than their Czech counterparts. The Czech versions, on the contrary, will display more links in the other categories – that is, complex repetition and simple and complex paraphrase.

The last research question again reflects the systemic difference of Czech and English. English is more tolerant to the repetition of the same lexical items in a text, while in Czech repetition is not a preferred stylistic device; Czech prefers repeating the items “in other words”. This difference is supposed to result in higher numbers of links in the category of simple repetition.

5.2 Adjusting the sample texts

The most obvious changes to the texts were omissions of certain parts and changes in the sentence borders resulting from the translator’s decisions when translating into Czech. The translator omitted those parts of the texts that were not relevant for the recipient of the target text due to cultural specificity or the gap between the time of production of the source text and the target text. Also, in the Czech translations longer sentences were in some cases divided into two. This led to the non-conformity of the Czech versions with the English originals. That is why the above mentioned changes were necessary, as the comparative interpretation of the results of the analysis requires the texts to be (approximately)

the same. Also, in the Czech texts it was necessary to add referents where they were not expressed explicitly.

5.2.1 Omissions and sentence borders

It is of course not desirable to adjust the translations of particular expressions. Such intrusions and “improvements” of particular items would result in an unnatural text, while the research aims to analyze the original Czech and English versions of the sample texts. Both versions of the text nevertheless need to have the same number of sentences and must match in the information they convey, so as to provide comparable results. As for the first case, the sentence borders had to be adjusted in some cases – several sentences in the Czech translations were joined together or split so as to match the English originals. In order for both language versions of the texts to convey the same meaning, the omissions of certain parts were necessary as well, due to the fact that in translation some parts of the text were omitted or added. Here, two approaches were possible – either omitting the part of the text not present in the other language version, or again taking the English original as a yardstick and adjusting the Czech version to it. The latter approach would, however, require translating the missing parts and adding them to the “official” translation by Jan Jirák. I decided to use the first approach, as I did not intend to intrude into the wording of the texts. Following is the example of the original text *Television Moscow-Style* and the adjustments made to the other language version¹³.

The English version:

(25) When we got back to the room eight hours later I immediately turned on the set.

(26) The David Susskind-type panel was still going on, and the guy was still doing the talking.

The adjusted Czech version of a text looks like this:

(25) Když jsme se o osm hodin později vrátili, okamžitě jsem se vrhl k televizoru. ~~Dva z programů vysílaly monoskop, ale~~

¹³ The numbers in the brackets indicate the order of sentences in a text.

(26) [D]iskuse řízená chlápkem podobným Davidu Susskindovi ještě běžela a chlápek pořád ještě mluvil.

From the example it is clear that a part of the sentence 25 was omitted in Czech, as it is not present in the original English version. In Czech, the sentence borders were adjusted as well, as the sentence 26 originally had a different opening.

5.2.2 Adding the reference

Apart from omitting parts of the texts and changing the sentence borders, it was also necessary to supply the referent where it was missing in the Czech versions, as in Czech the subject of a sentence may not be expressed explicitly and is inferred from the verbal suffix or from the co-text. The following examples were taken from the text *Prezidentův portrét*, the Czech version of *LBJ Betrayed in Portrait*:

(6) [JÁ] Jsem přesvědčen, že prezident má plné právo odmítnout vlastní portrét, neodpovídá-li jeho představám o sobě samém.

(7) Prezidentově podobě, jak [JÁ] ji znám z fotografií, zůstal portrét rozhodně mnohé dlužen.

Here, the reference to first person is obvious from the verbal suffix. The first person pronoun here refers to the author of the text. In the second example from the same text, the situation is different:

(18) Mohl by [HURD] začít tím, že prezidenta zachytí s úsměvem ve tváři.

(19) Místo knihy by přitom [PREZIDENT] mohl třímat výsledky průzkumu veřejného mínění Harrisova nebo Gallupova ústavu z roku 1964.

Here, the referent of sentences 18 and 19 is not clear from the form of the verb. From the co-text it is clear that while the former sentence refers to Mr. Hurd, the painter, the latter refers to President Johnson.

The cases where the reference had to be added were treated as entering co-reference links only; they were not counted into the category of simple repetition,

as the referent was not explicitly present in the text. The first category of simple repetition, as already explained, establishes the strongest type of links. For this reason the “added” referents, not being explicitly present in the text before the adjustment, were not treated as such, as in my opinion it would undesirably alter the results of the analysis concerning the category of the strongest links.

5.3 Using Hoey’s method of analysis

This section explains how the methodology was used. It provides reasoning for the changes to Hoey’s method and discusses its limitations and the problematic issues of its application. Again, systemic differences between Czech and English play an important role here.

5.3.1 Four categories of lexical cohesion

Hoey (1991) classifies four types of repetition patterns – simple repetition, complex repetition, simple paraphrase and complex paraphrase. Apart from these, Hoey also lists cases of co-reference, hyponymic/hyperonymic relation, substitution and ellipsis and a link triangle (with the linking item either present or missing in the text) as instances of lexical cohesion. All these cases listed by Hoey separately from the four categories were counted under the category of complex paraphrase in the analysis.

As already mentioned, Hoey (1991) considers simple repetition to be the strongest type of lexical cohesion, followed by complex repetition and simple paraphrase, the weakest form being complex paraphrase (including the above listed cases). In some cases, an item could potentially enter links in more categories, which differ in strength, so that there could occur a link with the item that is either weak or strong, depending on the category in which the link is counted (as there can be only one link between an item of one sentence and items of another sentence). The analysis proceeded from the strongest type, simple repetition, to the weaker types. This order prevented counting weaker links where stronger ones occur as well. The types of cohesive devices and their strength according to Hoey (1991, 83) are ordered from the strongest to the weakest:

1. simple lexical repetition
2. complex lexical repetition
3. simple mutual paraphrase
4. simple partial paraphrase
5. antonymous complex paraphrase
6. other complex paraphrase
7. substitution
8. co-reference
9. ellipsis

In the analysis, the category 1 refers to “simple repetition”, the category 2 to “complex repetition”, categories 3 and 4 are included in “simple paraphrase” and categories 5-9 belong to “complex paraphrase”.

5.3.2 *Open-set lexical items only?*

Hoey’s method is based on the connections between lexical items in a text. As already described in Chapter 3, Hoey (1991) states that the basis for the analysis is counting the links between open-set lexical items. He nevertheless admits that even items that do not belong to this category may enter such links – the example may be pronominal co-reference. It is therefore possible to include such items as pronouns. Hoey (1991) suggests that only the third person pronominal reference should be counted, unless the first or second person occurs in direct speech. In that case, there is a shift of reference between the first/second person of the direct speech and the third person – in other words, two different personal pronominal references share the same referent. The example from *Japanese Men Are Happy* illustrates the situation:

(21) Most American wives will run the water, hit their husbands a couple of times with a washcloth and then hand him a towel and say,

(22) “Dry yourself.”

It is evident that all the underlined expressions refer to the same entity – the American husband. The second person pronoun in direct speech thus has the same reference as the third person pronoun outside it (here, of course, the third person

pronoun is marked only to emphasize the reference; it could not be counted into the co-reference chain, as the sentence already contains the nominal referent).

In the sample texts, the pronominal reference was counted as part of the co-reference chain. Apart from third person pronouns, the texts contain the first and second person pronouns as well. These instances of pronominal reference mark the author and the recipients of the text. They are the means of establishing the contact between the author and the recipients and are therefore repeatedly used throughout the texts. These instances were counted as links, even though they do not occur in direct speech. On the contrary, the first and second person pronominal reference in direct speech was, in accordance with Hoey, counted to the respective third person reference.

Another category difficult to tackle were numeral expressions. The numerals occur in the texts either on their own, as in “two hours”, or as a part of expressions like “the first time”. Also, as the examples illustrate, there are more types of numerals in the texts (ordinal and cardinal). I decided to treat these as open-set lexical items and count them as creating links in the texts. It was also necessary to decide how to treat these two types of numerals: for instance, whether “one” and “first” should be treated as simple or complex repetition. In accordance with Hoey’s suggestion, repetition of two items of the same part of speech creates simple repetition, but Hoey also defines simple repetition as a case when two words share the same word basis and part of speech and the difference in their form lies within the scope of grammatical paradigm, such as “four – fourth”. In the case of “one – first”, the items do not share the same stem, but as the difference between them is not that of different part of speech, these items were counted as the first category of lexical cohesion.

5.3.3 *Proper nouns*

The sample texts¹⁴ contain references to people, places and cultural elements. These are problematic to count, as Hoey (1991) does not suggest any particular principle of categorizing such items.

¹⁴ The whole sample texts are provided in the Appendix. Further comments on each text are provided in the chapter 6.

The least problematic were place names, such as “The White House” or “Moscow” and references to cultural elements, such as the film “Last Tango In Paris” in the sample text *Last Apartment In Paris* and the respective Czech equivalents. These expressions allowed for counting them as the instances of simple repetition, as there was no variation in the particular expressions referring to the respective phenomena. If there was a pronominal reference to these items, it was counted as an instance of co-reference, thus belonging into the fourth category, complex paraphrase.

A problematic issue was the categorization of proper nouns referring to people. In some cases, there was a considerable variation in the expressions referring to a particular person, such as “Marlon Brando, Marlon, Brando” or “Johnson, Mr. Johnson, President Johnson, President”. As the latter example taken from the text *LBJ Betrayed in Portrait* shows, the proper noun occurs also in combination with a general classifier.

The problem was further complicated by the fact that these expressions were combined with pronominal reference as well. In the cases where a proper noun was combined with a classifier, I decided to prefer the classifier over the proper noun and classify such instance as simple repetition. In the other cases, I decided to count all the instances like “Brando, Marlon, Marlon Brando” in the category of complex paraphrase together with their pronominal references, as counting separately “Marlon”, “Brando”, “Marlon Brando” etc. as instances of simple repetition and then combining them with each other and with the pronominal references within the category of complex paraphrase manually would be very complicated and prone to mistakes in count.

The possessive case, such as “president’s”, was counted as simple repetition with “president”. The instances of “his” versus “he” were treated in analogy to this.¹⁵

5.3.4 Negation in Czech

Another area where Hoey’s method was problematic to apply was when a negative verb occurred in Czech. In English, the negative “not” occurs either separated from the other words, or in the contracted form, it is connected to the

¹⁵ The pronominal reference, as already stated, was counted as complex paraphrase.

auxiliary verb, so that the lexical verb retains the same form. In Czech, verbal negation is expressed by the verb taking a negative prefix “ne-“, for example “vysílali – nevysílali”. The basis of the word remains the same, so in analogy to English, these instances were counted as simple repetition.

5.3.5 Gradation

In both Czech and English, there is a regular way of gradation of adjectives and adverbs. But both languages also contain irregular expressions, where the regular grammatical paradigm cannot be applied, for instance “dobrý – lepší” in Czech and the equivalent “good – better” in English. The regular cases would be counted as simple repetition (e.g. “warm – warmer”), as they fulfill Hoey’s (1991) definition of simple repetition – they are of the same part of speech, share the same stem and the difference between their forms is within the grammatical paradigm. Again, as with the numerals described above, the irregular cases of gradation obviously do not fulfill all these requirements set by Hoey, as they are not derived from the same lexical basis. But as the stems they are derived from are complementary to each other, these adjectives and adverbs are treated in the same way as the regular cases – that is, as simple repetition.

5.3.6 Subjectivity of analysis

The last general remark regarding the application of Hoey’s method I would like to make is on the subjectivity of evaluation. As Hoey (1991) remarks, judging whether two particular words in a particular text can or cannot be treated as synonyms is always a subjective matter.

I would like to emphasize that not only the assessment of synonymy is subjective. The same might apply in some cases to antonymy, the link triangle and hyponymy/hyperonymy as well. The replication of the analysis by a different person may thus lead to slightly different results in the above mentioned cases.

6. THE ANALYSIS

The following chapter describes the procedure of the comparative analysis and describes the sample texts one by one, emphasizing their specific features and the way these influenced the application of Hoey's method.

6.1 Analyzing the texts

This section provides the description of the procedure of the analysis as it was done in the research project. The following sections discuss the limitations of performing the analysis manually and list the mechanisms used to eliminate the possibility of wrong count.

The procedure of the analysis could be described in the following steps:

1. Adjusting the sample texts.
2. Analyzing the texts with regard to the four categories of lexical cohesion, as defined by Hoey – proceeding from simple repetition via complex repetition and simple paraphrase to complex paraphrase. Recording the particular links in each category. Every link is recorded at both sentences establishing it.
3. Determining the bonding between the sentences: counting the number of links between each two sentences and establishing the bonds (three links constitute a bond). Recording the bonding with every sentence.
4. Transferring the recorded bonds into the form of a quantitative matrix. Checking whether the bonds were recorded correctly (and possible re-counting and corrections). Counting and categorizing the bonds. (Up to this point, the bonds are counted “twice”, as they are recorded at both sentences creating them. Now the total amount of links and bonds is divided by two in order to get the correct number of bonds.)
5. Comparison of the data for the English and the Czech texts in the four categories of lexical cohesion and the number of bonds.

6.1.1 The limitations

The analysis was done manually. Here, it is important to note that the manual analysis of a text is a very complicated and time-consuming process which does

not allow for bigger amounts of text to be analyzed by one person. Firstly, each item of every sentence must be assessed against all the other sentences, so that with a higher number of sentences the number of items to be assessed would grow considerably. Analyzing manually a longer text, for instance a text in the extent of a scientific article or a short story, would be very demanding if not impossible. Secondly, analyzing a longer text manually would be prone to “overlooking” links or double-accounting simply because it is not possible to concentrate on the analysis, all the possible combinations and the assessment of each potential case of linkage for the whole time needed to perform the analysis.

Using a computer to speed up the process would also bring certain disadvantages. As already proven, even the category of simple or complex repetition is far from being unambiguously assessable, as it is necessary to assess the lexical items even if they are not open-set and decide whether to include them or not in that particular case. Also, there are irregularities in the grammatical paradigms that further complicate the analysis. The third and fourth category depends on context to a large extent and needs to be assessed by a human.

6.1.2 Eliminating the mistakes

The manual analysis of texts brings two risks. Firstly, the assessment is inevitably subjective, as noted earlier, particularly in the categories of synonymy, antonymy or link triangle, where the classification of the items depends on the recipients’ evaluation of an item in the particular context. In other words, the recipient decides whether s/he perceives a particular item in a particular text as synonymous to another one or not.

Secondly, there is always the risk of making mistakes in count. The loss of concentration during the process of analysis may lead to overlooking a link or counting it twice or to assessing an item incorrectly and thus counting a link where none should be counted or vice versa. Such a mistake may occur in individual cases of linkage. But, as Hoey (1991) himself remarks, it is the patterning of bonds that is important for the analysis, not an individual link, as it is necessary for two sentences to establish three links with each other in order to be bonded. In view of this, an error in count on one particular occasion therefore does not necessarily invalidate the whole analysis (ibid.).

In order to ensure the correct results, all the instances of linkage and bonds are recorded for every text, so that it is possible to return to a particular case, if needed. Apart from the possibility of re-counting the results of each analysis or re-assessing a particular instance of linkage, another mechanism eliminating the possibility of mistakes is the record of links with each particular sentence and the subsequent transfer into the quantitative matrix. This is possible due to the fact that there are always two items sharing a link or a bond.

At first, the links are recorded for every sentence separately – for instance if there was a link between sentences 12 and 16 of a text, “16” would be recorded to sentence 12 and “12” would be recorded to sentence 16 in the particular category of lexical cohesion in which the link was established. Later, the bonds with other sentences are counted for every sentence – for instance, if “16” occurred three times at sentence 12, a bond would be recorded with sentence 16. The sentence 16 would then be supposed to have three links with sentence 12. The process of transfer of the bonds into the quantitative matrix then serves not only as a means of presenting the results comprehensibly and clearly all in one place, but also to check that the bonds recorded with one sentence match its counterparts. The following case is given as an illustration: A bond with sentence 16 was recorded at sentence 12. It is therefore assumed that sentence 16 also has a recorded bond with sentence 12. When transferring the sentence 12, a bond is recorded with sentence 16. When transferring sentence 16, the presence of a bond recorded with sentence 12 serves as the proof that the bond has been established and counted correctly at both the items. The examples from the text *Television Moscow-Style* are given below.

Sentence number	Simple repetition	Complex repetition	Simple paraphrase	Complex paraphrase	Bond with
8	1, 2, 3, 4,6, 9, 22, 22 , 23	5, 11	17	11, 13, 16, 18, 19, 22	22
22	8, 8, 11, 14, 14, 14 , 15, 20, 21, 21	11	-	1, 2, 3, 4, 5, 6, 8, 9, 11 , 13, 16, 18, 19, 23, 23	8, 11

Table 5: Record of links in the four categories and bonds with particular sentences

As explained above, the “double record” of each link is a key concept when analyzing a text manually, as it allows for three various methods of checking the results, as well as re-evaluating and re-counting any particular case, if necessary.

The following sections present specific remarks on every sample text used for the analysis. The focus is drawn on the features relevant for the analysis, such as the problematic cases of count etc., but the Czech translation is commented upon as well where deviations from the original etc. were observed. The quantitative matrices are provided separately for each text. Their interpretation and discussion of results follows in the next chapter.

6.2 Japanese Men Are Happy (Proč jsou v Japonsku muži šťastní)

This text ironically describes the happiness of Japanese households, where the woman is in inferior position to her husband, contrary to the American marriages.

The singular and plural forms of the same words are counted as simple repetition, e.g. “wife” in sentence 2 and “wives” in sentence 1 are regarded as two forms of the same word basis and therefore counted as the first category, the same as the Czech “manželky” in sentence 1 and “manželka” in sentence 10.

The negative verbal forms in Czech were treated simply as a different form of the verb, thus constituting a simple repetition link, such as “ukloní” in sentence 12 and “neukloní” in sentence 19. In English this is not applicable, as the negative is expressed either separately by “not”, or in contracted form with an auxiliary verb.

An example of Czech language not corresponding to the rules set by Hoey on the basis of English are the words “manžel” and “manželka”. The only formal difference is the suffix, so if not taking into account the meaning, these could be treated as simple repetition. Obviously enough, this is not possible, as they refer to two different entities within the text and therefore are counted separately – all instances of “manžel” are taken as simple repetition, the same as all instances of “manželka” (irrespective of the grammatical case). As explained previously, the instances of possessive case were counted as establishing a link of simple repetition with the nominative case of the same item. This applies to both Czech and English. Thus, e.g. “husband” in sentence 2 and “husband’s” in sentence 6

enter into the first category and analogically, “manžela” in sentence 2 and “manželovo” in sentence 5 are also recorded as an instance of simple repetition.

The gerund form “stepping” in “one whom any husband would be proud of stepping on” in sentence 27 and “step” in “won’t have to step” in sentence 6 are counted as an instance of complex repetition.

The instances of “woman” and “wife” (both in singular and plural) in English and “manželka” and “žena” in Czech were treated as synonyms in this particular text, as there was recognized no specific pattern of their use (for instance one carrying positive and the other negative connotations). As already stated, the singular and plural forms were counted as simple repetition and therefore even “wife” and “women” were treated as synonymous – that is, within the third category, simple paraphrase.

“Anyone” in sentence 9 and “nikdo” in the Czech version of the same sentence is treated as being a co-referent of “husband” in the context of this particular text, which is:

EN: “...anyone who has been bathed by a Japanese woman will never want to take an American shower again.”

CZ: “...nikdo, koho jednou myla japonská žena, se už nechce vrátit ke sprše.”

The combination of “slave” and “superior” in sentences 2 and 27, as well as “master” and “inferior” in sentence 2 and 25 were treated as instances of antonymy, as in the fourth category it is possible to count even different parts of speech as entering the complex paraphrase link. This was not possible in Czech due to translation: the pair “otrok” in 2 and “nadřazený” in 25 could be taken as antonyms in this particular context, but the other pair working in English is not present in Czech, as the equivalent to “inferior” is “ponížit se” in Czech, which was perceived as a slight shift in meaning and the link was therefore excluded.

The Czech translation, as it is evident from the transcript provided in the Appendix, contains a typo in sentence 2 – “osvození” instead of “osvobození”.

and therefore there are not enough links with any other sentence to create a bond. These sentences can thus be regarded as marginal.

The sentences with the highest number of bonds are sentences 8, 10, 19, 25 and 27 (all of them establish more than 15 bonds with other sentences). If we applied Hoey's method for creating abridgements and the threshold was 15 and more bonds, we would get this gist of the text:

Another area in which Japanese women excel is giving their husbands baths.

Unlike the American woman, a Japanese wife looks forward to giving her husband a bath.

Contrast this to the average American home where the American wife not only refuses to bow to her husband when he comes home, but in some cases won't even give him a bath.

American women are afraid that if they offer to bathe their husbands, they will be considered inferior.

A wife who knows how to bathe her husband in the Japanese style is a superior person, and one whom any husband would be proud of stepping on when he gets out of bed in the morning.

In order to get a more complete message, the threshold could be lowered to include more sentences, for instance those having 5 or more bonds etc. This topic will not be covered in greater detail, as it is not the aim of the thesis. But even from such a brief example it is clear that these few sentences give a summary of the main topic of the sample text.

If the same criterion is applied to the count of central sentences as with the English version – that is, at least 15 bonds with other sentences or more, only four sentences can be treated as such. These are sentences 8, 10, 19 and 27. (Sentence 25 which fulfills this criterion in English is excluded in the Czech version, as it is below the threshold). The summary of the text would thus look like this:

Jiná oblast, v které japonské ženy vynikají, je koupání manžela.

Na rozdíl od americké ženy se japonská manželka na koupání manžela těší.

V americké domácnosti se manželka nejen manželovi neukloní, když přijde domů, ale někdy ho dokonce i odmítne vykoupat.

Žena, která ví, jak vykoupat manžela v japonském stylu, je bytostí nadřazenou a každý muž bude pyšný na to, že si na ni může stoupnout, když ráno leze z postele.

Again, the Czech version gives a gist of the main topic of the text, even though one sentence is missing compared to the English version.

When compared with each other, the Czech version of the sample text differs from the English one concerning the cohesive linkage in each category. The following table gives an overview of the linkage and bondage of both language versions¹⁶:

¹⁶ Here, the numbers of links and bonds refer to the final count – that is, the output numbers of the analysis are recorded after being divided by two (as each link is recorded twice at first, as explained earlier).

	simple repetition	complex repetition	simple paraphrase	complex paraphrase	number of links	number of bonds
EN	253	19	44	341	657	101
CZ	183	28	49	330	590	100

Table 9: Summary of links and bonds – Japanese Men Are Happy (Proč jsou v Japonsku muži šťastní)

It is obvious from the table that the English version has a significantly higher number of links in the category of simple repetition, while the number of links in the category of complex repetition is higher in the Czech version. This corresponds to the tendency of Czech to avoid repeating the same words multiple times.

What is interesting is the fact that while the number of bonds is almost identical, the number of links is higher in English, contrary to the expectations formulated previously.

6.3 The Last Apartment In Paris (Poslední byt v Paříži)

The Last Apartment In Paris is a humorous commentary upon a very famous movie, The Last Tango In Paris. Buchwald “re-interprets” the plot from a funny perspective of housing shortage In Paris.

Apart from mentioning the city of Paris as a part of the text’s title and of the name of a famous film, the city is mentioned once again separately in both Czech and English. Here, the name of the film had to be regarded as a separate multi-word unit and “Paris” (“Paříž” in Czech) in this unit was not counted as entering into a simple repetition link with “Paris” (“Paříž”) in sentences 6 and 7 (occurring in both language versions). These two instances referring to the city itself, on the contrary, were counted as simple repetition.

Once again, the Czech version contains a case of a verb occurring in both positive and negative form – “získali” in sentence 7 and “nezíská” in sentence 21. Analogically to the previously described case, these verbal forms were counted as entering into a simple repetition link.

Similarly, the complementary forms of the first person pronoun “we” and “our” were treated as two forms of the same word. If these were open-set lexical items, they would be counted as simple repetition, but as pronouns they were categorized as items establishing links of complex paraphrase. Hoey (1991) does not include first person pronouns unless they occur in direct speech, but as already explained, in this analysis they were included into the count.

“Others” in sentence 3 of the English version was treated as an instance of co-reference with “critics” in sentences 2 and 4, equal to “jiní” at the same place in the Czech version and its co-reference link with sentences 2, 4 and 5.

Due to the reasons stated in greater detail previously, the variants of the names of people were counted together with the pronouns referring to them as entering into co-reference. This is one particular case in which the method could be changed in case of future application, but as already explained, counting separately every “Marlon”, “Brando” and “Marlon Brando” as simple repetition and then linking them with each other and with the pronominal reference manually would be very prone to mistakes in count.

The Last Apartment In Paris is one of the sample texts that contain direct speech. Here, Hoey (1991) advises to account for the shift of personal reference and include first and second person pronouns, counting them to their respective referents. The direct speech occurs twice. The second case, sentence 37 in both texts is unproblematic. The reference is clear and nothing prevents the sentence from being treated as a separate unit establishing links with the other sentences, although it is a marginal sentence, as it is not bonded to any other one in either Czech or English.

The first case, on the contrary, was a bit problematic.

EN: This infuriates Brando and he throws her down on the bed and keeps muttering, “It’s mine. It’s mine.”

CZ: Bran[da] to ale rozčílí natolik, že ji povalí na postel a začne mumlat: „*Moje, moje...*“

The reference to “it” in English was not quite clear and the Czech version is even more confusing, given the context – it is not clear whether “moje” refers to “the

bed”, “the flat” or “Miss Schneider”. Of course this is one of the humorous parts of the text, where the original situation in the film is interpreted in a different context, so this ambiguity serves as the means of conveying humor. For the analysis, however, the unclear reference is problematic. For this reason I decided not to treat this sentence as a separate unit, as the transcript shows. All the other cases of direct speech in all the sample texts were treated separately, as this problem with unclear reference does not occur in any other place.

One more remark on this sample text is devoted to the Czech translation and the mistakes occurring in the text. It is evident that in some cases, the name “Brando” was misspelled. Thus, for instance, we can find “Brandonovu” instead of “Brandovu”. It is not quite clear what the cause of this mistake was – whether it is simply a typo occurring on several occasions, or the influence of wrong inflection of the proper noun. What is even less clear is the change in the subject of sentence 38. The English subject “I” was substituted in Czech by “ona” (not present explicitly, but inferable from the verbal suffix), which shifts the subject from first person singular referring to the author to the third person singular referring to “Miss Schneider”. Again, there are two possible reasons for this change – either a mistake on the part of the translator, or a deliberate change of the subject, but here again the reason for such a change is not quite obvious.

The English version of the sample text contains eight marginal sentences, as the matrix above shows. These are sentences 1, 15, 19, 24, 25, 26, 37 and 38. These sentences do not share any bond with other sentences.

The number of bonds in this text is quite low in general. If taking 10 bonds as a threshold for the selection of the central sentences, only two would match this criterion – sentences 7 and 39. A summary would thus look like this:

Only those who have ever searched for an apartment In Paris can appreciate what Brando and Miss Schneider go through for this lovely flat near Seine.

I don't know if "Last Tango In Paris" is a great movie or not, but I believe that director Bertolucci has made an important social statement about one of the real outrages of our time, which happens to be the housing shortage in France.

Although these two sentences seem to sum up the idea of the author from the point of view of a person familiar with the whole text, for a reader not familiar with the whole story the information would probably be rather misleading.

If the threshold is lowered to 5 bonds, there are nine sentences with a sufficient number of bonds – sentences 5, 7, 11, 12, 22, 23, 29, 30 and 39. The following examples give the summary of the text for the latter case:

"Last Tango In Paris" is not, as has been described, the story of an aging American (Marlon Brando) and a young girl (Maria Schneider) in a desperate sexual battle for survival.

Only those who have ever searched for an apartment In Paris can appreciate what Brando and Miss Schneider go through for this lovely flat near Seine.

They meet by chance in an empty flat and you see Brando's mind working.

He figures if he rapes the girl, she'll go away and he'll get the apartment.

We see a tiny hotel he lives in and realize why Brando is so intent on getting the apartment.

Miss Schneider goes off with her fiancé and we discern why she wants a new place to live.

One takes place against the wall and Miss Schneider realizes if she ever gets the flat she's going to have to buy a lot of wallpaper.

Rather than being frightened by Brando's brutality, Miss Schneider becomes more determined than ever to wrest the key away from him.

I don't know if "Last Tango In Paris" is a great movie or not, but I believe that director Bertolucci has made an important social statement about one of the real outrages of our time, which happens to be the housing shortage in France.

In this case, the summary does not work very well, as it does not reveal the point of the text, but it covers most of the important information in the text.

The number of marginal sentences is higher in Czech compared to English – the Czech version contains ten such sentences. These are sentences 3, 8, 19, 24, 25, 26, 27, 34, 37 and 38.

If applying the same two thresholds as in the English text, the first case (10 bonds as the minimum) would result in a resume consisting of only one sentence – sentence 7 (as the other sentence included in the English version, sentence 39, has only 9 bonds in Czech). This sentence could serve as a gist of the point the author makes in the text, but it would not be sufficient as a summary.

Jenom člověk, který se někdy pokoušel najít si v Paříži bydlení, může pochopit, čím vším musel Brando a slečna Schneiderová projít, aby [ONI] získali vyhlídnutý byt nad Seinou.

The second case with threshold lowered to 5 bonds, the summary would consist of eight sentences – 5, 7, 11, 12, 14, 29, 35 and 39.

Jenom člověk, který se někdy pokoušel najít si v Paříži bydlení, může pochopit, čím vším musel Brando a slečna Schneiderová projít, aby [ONI] získali vyhlídnutý byt nad Seinou.

Poslední tango není, jak nám kritikové často vnucují, příběh stárnoucího Američana (Marlon Brando) a mladé dívky (Maria Schneiderová), kteří [ONI] prožívají zoufalý sexuální poryv ve snaze o přežití.

Zcela náhodou se [ONI] setkají v prázdném bytě a my jako diváci můžeme sledovat, co se odehrává v Bran[dově] mysli.

Dochází mu, že když tu holku znásilní, ona odejde a byt připadne jemu.

Ve skutečnosti, zatímco se jí Marlon dravě zmocňuje, snaží se [ONA] pohledem odhadnout rozměry pokoje, aby věděla, jak velký koberec si má koupit.

Jeden z nich se odehraje blízko stěny pokoje a slečně Schneiderové dojde, že pokud byt někdy dostane, bude muset koupit spoustu tapet.

Poněkud poznamenána sexuálními prožitky vrací se slečna Schneiderová do bytu a přivádí s sebou snoubence, aby mu obydli ukázala.

[JÁ] Nevím, jestli Poslední tango v Paříži je dobrý film, ale jsem si jist, že režisér Bertolucci významným způsobem přispěl k zachycení jednoho z vážných problémů současnosti – nedostatku obytných prostor ve Francii.

Here, the summary provides the reader with the most important information. In my opinion, this summary would inform the reader in a better way than the English one resulting from the same criteria.

The differences between the two language versions are summarized in the following table.

	simple repetition	complex repetition	simple paraphrase	complex paraphrase	number of links	number of bonds
EN	121	7	102	626	856	62
CZ	165	28	77	544	814	55

Table 12: Summary of links and bonds – The Last Apartment In Paris (Poslední byt v Paříži)

The English version proves to be more cohesive than the Czech, establishing a higher amount of bonds, as recorded in the last column of the table. However, the previous columns are of interest here, as they show interesting findings. The total number of links is lower in Czech, as well as the number of links in the last two categories – simple and complex paraphrase, while in the first category, the simple repetition, is more frequent in Czech than English. This is in a complete opposition to the expectations. The category of complex repetition matches the expectations, as it is more frequent in Czech than English.

6.4 Television Moscow-Style (Moskevská televize)

This column has a slightly satirical undertone, commenting upon the way the Soviet television informs, or rather does not inform, its viewers about the real situation. Buchwald tells the story of his own visit to Moscow and his encounter with Soviet broadcasting.

The irregular gradation of adjectives or adverbs mentioned earlier occurs in both the Czech and the English version. The English irregularity “good – better” in sentences 6 and 13 was, in accordance with the principles explained earlier, treated as simple repetition, as both the expressions are of the same part of speech. In the following case, however, the change in the part of speech occurred:

(5) I know this because I had a television set in my room at the hotel, and I watched it constantly to see if the Soviets were doing any better with their electronics media than we were doing with ours.

(6) It would be unfair to say Soviet television is better than American television – it’s just different.

It is clear that the two instances of “better” do not share the same part of speech. This case was therefore treated as complex repetition, although the form of the expression was identical. In Czech, the situation was different due to the translation. Although Czech has an equivalent to “good-better” that is irregular as well, “dobrý – lepší”, this pair does not occur in the text. The Czech version contains only the second pair, “lepší” in sentence 6 and “lépe” in sentence 5. These are in analogy to the English version treated as a link in the category of complex repetition.

The numerals occurring in the text are treated as lexical items. With this part of speech, as with the adjectives described above, there are various forms, such as “one” and “first” that do not share any stem, but as with the adjectives, are treated as cases of simple repetition. Thus, for instance, “dvě” and “druhý” (sentences 10, 19 and 20) are classified as simple repetition. Treating numerals as lexical items entering the links also allows for including cases like “první” and “poprvé” (sentences 7 and 15) as complex repetition.

In English, “TV” or “television” can be used to refer both to the device and the broadcasting in general, so before the analysis, there was an assumption that it might be necessary to distinguish between these two. If there were instances of such ambiguity, Hoey’s (1991) condition of “synonymy within the context of a particular text” would be used to distinguish between the two senses. Here, however, this situation did not arise.

Substitution occurred in the text both in Czech and English, as the following example illustrates:

(3) I blame the Soviet television network for this.

(3) A může za to sovětská televize.

In both cases, the pronoun refers to the previous sentence describing the lack of information caused by the Soviet TV. Substitution is counted under the fourth category of complex paraphrase.

One of the categories that is problematic and influenced by the subjective judgement is the link triangle. In some cases it is difficult to decide whether a link is established or not. A very good example of the link triangle with an item not present in the text can be taken from the Czech version of the sample text: “rozhovor” in sentence 19 and “mluvil” in sentences 26 and 26 establish a link in the category of complex paraphrase via “hovořit”, an item not present in the text, as this item is synonymous to one of the two linked items and may paraphrase the other.

A curious situation arises in the Czech version due to the way the negative verbal prefix is counted, as illustrated previously. Here, the two forms of the verb “vysílali” in sentence 22 and “nevysílal” in sentence 15 are counted as simple repetition. The item of sentence 22 is synonymous with “dávali” in sentence 8. Therefore even the verb in sentence 15 must be treated as synonymous to sentence 8, even though their meaning, “nevysílat” and “dávat” would normally lead to including them into the category of antonymy, thus belonging to the fourth category of complex paraphrase.

In this text, the higher number of links in Czech within the fourth category of complex paraphrase may be said to be caused not only by the systemic tendency

of the Czech language to repeat things “in other words” instead of using the same wording, but by the translation as well. In the Czech version, the exact wording used by the translator led to the increased number of links with the fourth category, in particular the co-reference, as some expressions were translated in such a way that they established links not present in the English original. An example may be “sightseeing” translated as “po městě/na prohlídku města” in sentences 17 and 24 or “ven” in sentence 21, these items establishing links with “Moskva”, as they refer to the same place in this particular context.

I know this because I had a television set in my room at the hotel, and I watched it constantly to see if the Soviets were doing any better with their electronics media than we were doing with ours.

During a lull in the program, I decided to change channels (Moscow is supposed to have four), and lo and behold, there was a panel show with a moderator who looked just like David Susskind.

My Russian wasn't good enough to understand what they were saying, but, as in America, the fellow who looked like David Susskind was doing all the talking.

Two hours later I watched steel being made in a large foundry outside Moscow, and this program was followed by an interview with a professor of Siberian tree transplants.

I turned the channel and the panel was still in session, with the fellow who looked like David Susskind still doing the talking.

My wife kept insisting we leave the room but I decided to have one more go at the third channel, and I'm glad I did.

When we got back to the room eight hours later I immediately turned on the set.

This summary does not provide the reader with the author's opinion. A reader not familiar with the whole text would probably be confused by such a summary.

Vím to, protože jsem měl televizor ve svém hotelovém pokoji a televizi jsem soustavně sledoval, abych zjistil, jestli si vedou sovětská elektronická média lépe než naše.

Během přestávky jsem se rozhodl přepnout na jiný kanál (v Moskvě mají být čtyři) [a] [v]padl jsem doprostřed panelové diskuse, kterou řídil muž podobný Davidu Susskindovi.

Neumím rusky natolik, abych rozuměl tomu, o čem si povídali, ale podobně jako v Americe obstaral veškeré mluvení ten chlápek, co vypadal jako David Susskind.

O dvě hodiny později jsem se už díval na to, jak se v jedné huti u Moskvy vyrábí ocel [a] [p]ak následoval rozhovor s jakýmsi profesorem ze sibiřské ovocnářské šlechtitelské stanice

Přepnul jsem na druhý program, ale tam byla pořád ještě ta panelová diskuse a chlápek, co vypadal jako David Susskind, pořád ještě mluvil.

Manželka bez ustání naléhala, abychom šli ven, ale já se rozhodl, že se ještě mrknu na třetí program [a] [u]dělal jsem dobře.

Když jsme se o osm hodin později vrátili, okamžitě jsem se vrhl k televizoru.

This summary gives a similar amount of information as the English one. Again, it does not make clear what the author intended to say, but this version would be even more misleading than the English one due to the unclear reference in the second sentence (“vím to”), which seems to refer back to the first sentence, which is in fact not true.

The overview of the linkage and bondage in the two versions of this sample text is provided in the following table.

	simple repetition	complex repetition	simple paraphrase	complex paraphrase	number of links	number of bonds
EN	113	18	12	219	362	45
CZ	102	9	12	251	374	41

Table 15: Summary of links and bonds – Television Moscow-Style (Moskevská televize)

The English version of the sample text contains less links than the Czech version, but more bonds. Also, the number of links in the first two categories is higher in the English version, while the Czech version contains more links in the last category. These results correspond almost fully to the expectations. The reason for the prevalence of the fourth category of links in the Czech version may be the translation. As already suggested, there are more items that enter the fourth category links due to the wording of the Czech translation than in English.

6.5 LBJ Betrayed in Portrait (Prezidentův portrét)

The last sample text satirizes the U.S. president Johnson on the basis of his refusal of a commissioned portrait of himself. Buchwald comments on this affair and describes how the portrait should have been done “properly” so that it would match the president’s personality.

As suggested previously, the problem was how to deal with the combination of a general classifier and the name of the president. In the English version, the text contained variants “LBJ” (sentence 20), “President Johnson” (sentence 1), “president” (sentences 2, 4, 6, 7, 8 etc.) and “Mr. Johnson” (sentences 3, 15, 22). The Czech equivalents were “prezident Johnson” (sentences 1, 15, 22), “pan Johnson” (sentence 3), “Johnson” (sentences 11, 12, 13) and “prezident” (sentences 4, 6, 7, 8 etc.). The expressions containing “president” were treated as instances of simple repetition. Thus even in sentence 1, where the combination of the general classifier “president” and the proper noun are combined, the common

noun was favored and the link was established within the first category. The remaining expressions with “Johnson” were treated, similarly to “Marlon Brando” in *The Last Apartment In Paris*, as complex paraphrase. The same principle applied to the name of the painter, “Mr. Hurd”. (These principles were applied in Czech as well.) The place names were counted as simple repetition, as they occurred in the same form on all occasions. In this sample text it would probably be possible to count each variant of the names separately as entering simple repetition links and then combining them in co-reference, as there are not so many of them. But as explained above, in accordance with the approach adopted in the sample text *The Last Apartment In Paris*, they were included in the fourth category. The possessive case was again included in the count of simple repetition with the respective item in the nominative case in both languages. Thus, for instance “prezidentovy” in sentence 14 was treated as simple repetition link with all the occurrences of “prezident”.

In the English version, a disambiguation in meaning had to be made, as there were three occurrences of the verb “look”, two of them conveying the meaning of “appear, seem” (sentences 10 and 11, in the expression “look as if”), while the third referred to the act of looking (sentence 8). Only the first two cases could thus be regarded as the case of simple repetition.

This text contains an example of meronymy (in the same place in both Czech and English), as both an expression referring to a part and a whole occur – these are “rty” in sentence 9, “tváře” in sentence 10 and “hlava” in sentence 14. Another part-whole relationship could be seen with “Senát”, “Sněmovna reprezentantů”, both in sentence 13 and the instances of “Kapitol”; however, it is not possible to include this case, as sentence 13 contains “Kapitol” as well, thus establishing simple repetition links with the other instances of “Kapitol”. Including the two items of the sentence 13 with them would then be double-accounting. The link established in the first category is therefore preferred over the other one, as the stronger type of link is favored in such cases.

A shift in meaning can be seen in the Czech version, where the translator decided to change the time reference in sentence 2 from “last week” to “před nedávnem”. The reason is obvious: while the original had to be published shortly after the portrait affair, there is a considerable gap between the time of publication

of the original and of the translation. The translator therefore deliberately decided to change the time reference, as if translated literally, it would no longer keep the correct time reference, which would possibly lead to the confusion of the recipient.

positioning of the figure and the general style were not consistent with other White House portraits.

But forgetting the portrait for a moment, anybody could see why the president would object to the way Mr. Hurd positioned the capitol in the background.

Anyone who knows the president's personality should have been aware that the only way to pose the president in such a picture was to have him standing in the capitol itself with one foot on the senate and the other on the house of representatives.

Also, if Mr. Hurd has studied his subject more closely, he would have painted the light shining over the capitol coming from the president's face and not from within the capitol itself.

Many museums are bidding for Mr. Hurd's portrait, but I think it would be unfair if it were publicly displayed, particularly when people know the president doesn't approve of it.

Here, sentence 3 alone would be sufficient to present the topic of the text. The summary above consisting of the six sentences works well this time, as it provides the most prominent information from the text and also presents the author's point.

Kdyby pan Hurd prostudoval svůj objekt podrobněji, musel by také onu zářiči namalovat tak, aby vycházela z prezidentovy hlavy, a ne z Kapitulu.

O portrét namalovaný panem Hurdem mají muzea veliký zájem, ale podle mě by nebylo správné ho vystavit veřejně, zejména když lidé vědí, že s tím prezident nesouhlasí.

From the point of view of the content, this seems to be an adequate summary of the text, as it provides the reader with key information as well as the author's point without containing any misleading or confusing part.

If the threshold was lowered to 5 bonds only, the summary would contain sentences 2, 13, 15, 16 and 20 as well.

Following is a table summarizing the numbers of links and bonds in the two versions of this sample text.

	simple repetition	complex repetition	simple paraphrase	complex paraphrase	number of links	number of bonds
EN	178	2	18	124	322	32
CZ	178	18	31	187	414	57

Table 18: Summary of links and bonds – LBJ Betrayed in Portrait (Prezidentův portrét)

It is obvious from the table that the Czech version of this sample text contains more links and bonds than the English version. The number of links in the first category is identical, while in the other three the amount of links is higher in Czech. The former finding is in opposition to the expectations, as English was supposed to be more cohesive than Czech. The latter finding corresponds to the expectations in that Czech creates more links in the latter categories (but it was expected that English would on the contrary establish more links in the first category). It was the wording of the translation that certainly contributed to the high amount of links in the Czech version.

6.6 Summary

The final part of this chapter is devoted to the summary of the results of the analysis of the four sample text by means of Hoey's method for analysis lexical cohesion in texts. This section provides concluding remarks and more general commentary on the findings.

The following table provides a comprehensive overview of the results in all categories of lexical cohesion analyzed in the sample texts, as well as the total amounts of links and bonds established in every text.

	simple repetition	complex repetition	simple paraphrase	complex paraphrase	number of links	number of bonds
Japanese Men Are Happy	253	19	44	341	657	101
Proč jsou v Japonsku muži šťastní	183	28	49	330	590	100
Last Apartment In Paris	121	7	102	626	856	62
Poslední byt v Paříži	165	28	77	544	814	55
Television Moscow-Style	113	18	12	219	362	45
Moskevská televize	102	9	12	251	374	41
LBJ Betrayed in Portrait	178	2	18	124	322	32
Prezidentův portrét	178	18	š1	187	414	57

Table 19: Summary of links and bonds in sample texts

The research aimed at answering three research questions concerning the analysis of lexical cohesion in Czech and English versions of four sample texts by means of Hoey's methodology.

The first research question dealt with the applicability of Hoey's method to Czech and English texts. As the previous sections of this chapter illustrate, the applicability was not completely unproblematic in either language. Even in English, certain categories and parts of speech, such as numerals or irregular gradation, posed problems and it had to be decided how to deal with them in this particular research. Another problematic area was the classification of proper names, as already suggested in the previous sections. In Czech, the same problems

applied, but apart from these, there were additional problems resulting from the systemic differences between the two languages: for instance, the negative forms of verbs, created by a negative prefix in Czech. In order to count the links analogically to the English version, the negative and the positive forms of the same verb were treated as simple repetition, which then lead to a paradoxical situation in case of synonymy – if a positive verb established synonymy with another verb, even the negative form of that particular verb had to be treated as synonymous, while in fact the meaning was opposite (as explained in greater detail above). In both languages it was necessary to treat every occurrence of particular words separately and assess individually its classification due to polysemy (if a word occurred in the text in two different meanings it could not be treated as simple repetition). From my point of view, based on the problems described earlier in this chapter, the applicability of Hoey's method to English versions of the text was in general less problematic than to the Czech ones.

The second research question posed a hypothesis that the English versions of the texts will be more cohesive than their Czech counterparts (strong claim) and that the Czech versions will establish more links but fewer bonds (weak claim). As illustrated by the overview table above, three out of the four sample texts are more cohesive in English than in Czech: in two cases, the number of bonds is slightly higher in English and in one case English is significantly more cohesive concerning the amount of bonds it contains. The strong claim may therefore be assumed to be confirmed.

The number of links in the Czech versions of the texts is higher than in English in two out of the four sample texts, in one case this difference is significant, and in the other the number of links in Czech is only slightly higher than in English. The weak claim cannot therefore be claimed confirmed or disproved. It needs to be emphasized that in order to provide generalizable results, the amount of sample texts would have to be significantly higher. The output of this research may serve only as a tentative statement of tendencies to be investigated further and in more depth.

The third research question formulated the expectation concerning the amount of links in each category of lexical cohesion. It was assumed that in the English texts the amount of links in the first category, simple repetition, will be higher than

in Czech, while in the remaining three categories, there will be a tendency of Czech texts establishing more links than the English ones. In the category of simple repetition, two out of the four sample texts correspond to the expectation, as they contain more links in English than in Czech; in one case the amount of the links is the same in both language versions. In the category of complex repetition, three out of the four texts contain more links in Czech. In the third category, simple paraphrase, two of the texts establish more links in Czech and in one case the number of links is equal in both language versions. In the last category, only two out of the four texts contain more links in Czech. From these results, a tentative conclusion may be drawn that the tendencies proposed by the third research question were proven, but again, a large-scale research would be needed to confirm the validity of this tentative conclusion.

The previous sections presented also illustrative samples of abridgements of the sample texts that would be created with use of various thresholds. An interesting finding is that the highest possible threshold (15 bonds) was not applicable to the longest sample text. It is the aim of this thesis to test these summaries, their length of adequacy; however, from the examples it is obvious that (1) one of the methods proposed by Hoey – that is the omission of marginal sentences, could not be applied (if taking only a sentence with no bonds as marginal), as there are too few such sentences, so that a summary where only these sentences were omitted would be too long and (2) that the summaries created by means of Hoey's method do not on all occasions provide their reader with sufficient information and do not always reflect the point the author wishes to make in the text. The reason may be that the usual form of Buchwald's columns has two parts – there is usually a frame introducing a story and ending it by a witty or satirical point the author wants to make and the story itself. These two parts may not be directly connected with each other from the point of view of lexical cohesion. Taking into account that the story covers the bigger part of the text, it is clear why the frame does not usually contain the central sentences and subsequently is not represented in a summary. (This tendency is illustrated by the matrices showing the number of bonds established between the sentences.)

Following is the summary of the research question posed by the research and their status after the analysis has been performed.

1. The applicability of the methodology for analysis presented by Hoey is less problematic in English than in Czech.

Status: confirmed

Hoey's method is applicable to both Czech and English. As the examples in this chapter indicate, in Czech there are more problematic areas that need to be tackled than in English. These support the hypothesis. The conclusion is nevertheless only tentative, as it is inevitably subjective, resulting from my personal experience, as the results from such a small-scale analysis could not be generalized.

2. The English versions of the sample texts contain more instances of bonds than their Czech counterparts. (strong claim)

Status: confirmed

The Czech versions of the texts contain more links, but fewer bonds than the English versions. (weak claim)

Status: unresolvable

The results of the analysis confirmed that the English texts establish more bonds than their Czech counterparts, but they do not clearly confirm the hypothesis that the Czech versions establish more links than the English versions.

3. The English versions of the sample texts show more instances of simple repetition than their Czech counterparts. The Czech versions, on the contrary, will display more links in the other categories – that is, complex repetition and simple and complex paraphrase.

Status: confirmed

The results suggest that there is a tendency of English texts to establish more links in the category of simple repetition than their Czech counterparts and of the Czech texts to establish more links in the remaining three categories than the English ones.

7. CONCLUSION

This Master's thesis aimed at investigating the devices of lexical cohesion in Czech and English publicistic texts by means of Hoey's method for the analysis of lexical patterns in text.

Cohesion is one of the seven standards of textuality – seven parameters that must be met if a stretch of writing or speech is to be regarded as a text. Cohesion deals with surface-level connectedness of individual items of a text. Thus it may be defined as a surface-level linkage of text elements (Tárnyiková 2002) or interrelation between parts of a text (Halliday and Hasan 1976). An individual manifestation of such relations is called a cohesive tie. As these definitions illustrate, there are various approaches towards cohesion, resulting in many different classifications of cohesive devices. Halliday and Hasan (ibid.), for instance, distinguish five types of cohesion: reference, substitution, ellipsis, conjunction and lexical cohesion. The last sub-type is of particular interest of this thesis. Again, various approaches may be applied to this phenomenon; this thesis adopts the approach of Michael Hoey and his concept of lexical patterns in texts. Due to systemic differences among the languages, the devices contributing to cohesion may differ. In this research the differences in the devices of lexical cohesion are analyzed comparatively in Czech and English.

Hoey's approach is based on the claim that a majority of cohesive links in texts are created by means of lexical cohesion. (Apart from his own work, Hoey supports his claim even by referring to Halliday and Hasan's research.) Hoey characterizes four major categories of lexical cohesion based on multiple occurrences of lexical items in a text. In his work (1991), Hoey explains in detail the methodology of analysis of lexical cohesion in texts. The method is based on determining, classifying and counting the number of "links" the individual items in the text establish and the consequent number of "bonds" (these are created by three instances of a link) the sentences share with the other ones. The method is aimed to serve as a basis for creating abridgements of non-narrative texts.

The sample texts used were four columns from the collection *Have I ever lied to you?* by the American columnist Art Buchwald and their Czech translations by Jan Jiráček published under the name *Copak jsem vám někdy lhal?* The choice

of one author and one translator made it possible to eliminate variation in the lexical patterns due to the style of the author or the translator. The choice of texts posed a limitation upon Hoey's method: it was used as the basis for comparative analysis of lexical cohesion, but its original function – a means of creating abridgements – was not fulfilled, as it was not the aim of the thesis to deal with the summaries and their properties. (The summaries provided in the thesis serve as mere illustrative examples of the usage of Hoey's method. They are not dealt with in greater depth.)

Based on the previous research in this field (various projects applying Hoey's method are mentioned in the thesis, as well as their findings), three research questions were formulated:

1. The applicability of the methodology for analysis presented by Hoey is less problematic in English than in Czech.

2. The English versions of the sample texts contain more instances of bonds than their Czech counterparts. (strong claim)

The Czech versions of the texts contain more links, but fewer bonds than the English versions. (weak claim)

3. The English versions of the sample texts show more instances of simple repetition than their Czech counterparts. The Czech versions, on the contrary, will display more links in the other categories – that is, complex repetition and simple and complex paraphrase.

In order to answer these questions, an analysis of the four sample texts in their both language versions was performed. Hoey's method was used as a framework for analysis. It was adjusted for this particular purpose (for instance, Hoey lists four main categories of lexical cohesion, but apart from them he also mentions cases of lexical cohesion such as antonymy, link triangle, hyponymy or co-reference – for the purpose of this research all these were included in the fourth category of complex paraphrase).

Each version of every sample text was discussed separately with focus on the specific areas that posed a problem for the analysis. The results of the analysis were presented in the form of quantitative matrix of the text's bondage. The two language versions of a sample text were then assessed comparatively with regard

to the linkage in every category and the total number of links and bonds. All results were consequently compared in order to determine the validity of the three research questions.

The first question focused on the systemic differences between Czech and English and the problematic areas for the application of Hoey's method. The analysis revealed that there are problematic cases in both languages. It was necessary to deal with the classification of numerals or proper names for instance. In Czech, other areas needed to be tackled as well, such as the negative forms of verbs or affixes. The application of Hoey's method to the English texts was perceived as less problematic than to the Czech ones. The first research question is therefore assumed to have been confirmed.

The second research question posed a hypothesis that the Czech versions of the sample texts will prove to be less cohesive than their English counterparts (strong claim). The sub-question formulated the expectation of a higher number of links and a lower number of bonds to be established by the Czech texts than the English ones (weak claim). The strong claim was proven by the analysis, as three out of the four sample texts created a higher number of bonds in English than in Czech. The validity of the weak claim could not be stated, as the results corresponded to the expectations in two cases only.

The last research question focused on the prevalence of the individual categories of lexical cohesion in Czech and English. Due to the systemic differences between the two languages it was expected that the first category, simple repetition, will prevail in English, as it has a higher degree of tolerance towards the repetition of particular words multiple times, while Czech prefers to change the wording in order to avoid such repetition. There was therefore expected a tendency of Czech to establish more links in the latter three categories – complex repetition and simple and complex paraphrase. The results are not completely unambiguous here, but a tentative conclusion may be drawn from them that there is a tendency of English to form more links in the first category and of Czech to prefer the other three categories.

The research proved the applicability of Hoey's method of analysis to Czech and English columns as the means of analysis of lexical cohesion;

however, as the illustrative summaries created by Hoey's method show, the summaries do not always provide sufficient information in an adequately long stretch of text. It was not the aim of the thesis to test the acceptability of these summaries or investigating the parameters leading to the most acceptable versions, but one observation can be made from them: due to the structure of the columns (a frame introducing and ending the topic and a story itself), the point the author wants to make by the text is often not reflected by the summary.

In general, Hoey's method is applicable to Czech, but as already stated, adjustments to the method need to be done so that it fits the system of the Czech language.

An issue worth attention is the subjectivity of assessment. As emphasized several times in the thesis, the assessment of such phenomena as synonymy, antonymy or link triangle is inevitably subjective. Another problematic area is polysemy: a word may occur in a text twice conveying the same meaning, or with different meaning each time. It is therefore necessary to assess every item individually. This process is very demanding, as it requires full focus, as well as time-consuming, if performed manually (as on this occasion). The manual mode of analysis allows for mistakes in count. These can be eliminated by "double" record of each link and subsequent "cross-check" when transferring the bonds into the quantitative matrix, as well as by the simple fact that the number of links, if counted before the transfer, must be divisible by two, as every link is recorded twice. (This method is described in detail in Chapter 6).

The last remark I would like to make is that on the validity of the results. It is obvious that the results of this research may be taken only as tentative, as the small number of sample texts does not allow for any generalizations. A large-scale research would be necessary for confirming or disproving the tentative conclusions of this work. Such a research is, however, not manageable in the fully manual mode and performed by a single researcher.

APPENDICES

Appendix 1: Last Apartment In Paris

1. It is incumbent on every columnist to see “Last Tango In Paris” and comment on it.
2. Some critics have called it the greatest movie of our time.
3. Others have written that it is one of the great rip-offs of the film industry.
4. But having seen the movie, I would like to advance the opinion that most critics have missed the point of the picture.
5. “Last Tango In Paris” is not, as has been described, the story of an aging American (Marlon Brando) and a young girl (Maria Schneider) in a desperate sexual battle for survival.
6. It is really a simple heart-warming film about two people trying to rent the same apartment In Paris.
7. Only those who have ever searched for an apartment In Paris can appreciate what Brando and Miss Schneider go through for this lovely flat near Seine.
8. In the film, Brando plays a washed-out American, whose wife has just committed suicide.
9. He wants the apartment in the worst way.
10. So does the young French girl.
11. They meet by chance in an empty flat and you see Brando’s mind working.
12. He figures if he rapes the girl, she’ll go away and he’ll get the apartment.
13. But Miss Schneider, a child of the French bourgeoisie, is made of sterner stuff, and she puts up little resistance to Brando’s assault.
14. As a matter of fact, while she’s being bounced around by Marlon she is really measuring the floor to see how much carpeting it will take.
15. The next day they are back at the apartment again.
16. Brando has bought a table, chairs and a bed to assert his claim to it.
17. But Miss Schneider is not impressed and walks about the place as if it were hers.
18. This infuriates Brando and he throws her down on the bed and keeps muttering, “It’s mine. It’s mine.”
19. Miss Schneider just laughs at him.

20. All the time they are making love she is looking at the window trying to figure what size curtains she'll need for the room.
21. Brando, exhausted and fearful that he'll lose the flat, visits his mother-in-law and his dead wife.
22. We see a tiny hotel he lives in and realize why Brando is so intent on getting the apartment.
23. Miss Schneider goes off with her fiancé and we discern why she wants a new place to live.
24. Back to the apartment.
25. Brando is now desperate.
26. He shows Miss Schneider a dead rat.
27. It shakes her up, but not enough to give up the place.
28. So Brando decides to humiliate her with several unnatural sex acts.
29. One takes place against the wall and Miss Schneider realizes if she ever gets the flat she's going to have to buy a lot of wallpaper.
30. Rather than being frightened by Brando's brutality, Miss Schneider becomes more determined than ever to wrest the key away from him.
31. The next time they meet she's in her wedding dress and Brando is so mad he throws her in the tub.
32. Miracle of all miracles, the plumbing works and Brando gives Miss Schneider a bath while she figures out what color scheme would go best with the white medicine cabinet.
33. By this time, Brando is worn out and figures the apartment isn't really worth it.
34. He leaves without telling Miss Schneider his name.
35. A little battered from the sexual encounters, Miss Schneider returns triumphantly with her fiancé to show him the flat.
36. But after all Miss Schneider's been through, the fiancé takes one look at the place and declares,
37. "It's too big."
38. This is when I started to cry.
39. I don't know if "Last Tango In Paris" is a great movie or not, but I believe that director Bertolucci has made an important social statement about one of

the real outrages of our time, which happens to be the housing shortage in France.

Appendix 2: Poslední byt v Paříži

1. Od každého sloupkaře se dnes čeká, že [ON] zajde na film Poslední tango v Paříži a napíše, co si o tom myslí.
2. Někteří kritici prohlašují, že je to nejlepší film současnosti.
3. Jiní prohlašují, že je to jeden z největších neúspěchů současné kinematografie.
4. Také jsem [JÁ] na film zašel a musím říct, že většině kritiků unikla hlavní myšlenka filmu.
5. Poslední tango není, jak nám kritikové často vnucují, příběh stárnoucího Američana (Marlon Brando) a mladé dívky (Maria Schneiderová), kteří [ONI] prožívají zoufalý sexuální poryv ve snaze o přežití.
6. Ve skutečnosti je to srdcervoucí příběh o dvou lidech, kteří se pokoušejí pronajmout si v Paříži tentýž byt.
7. Jenom člověk, který se někdy pokoušel najít si v Paříži bydlení, může pochopit, čím vším musel Brando a slečna Schneiderová projít, aby [ONI] získali vyhlídnutý byt nad Seinou.
8. Ve filmu hraje Brando neúspěšného Američana, jehož žena právě spáchala sebevraždu.
9. Za každou cenu se [ON] snaží najít bydlení.
10. Mladá Francouzka se snaží o totéž.
11. Zcela náhodou se [ONI] setkají v prázdném bytě a my jako diváci můžeme sledovat, co se odehrává v Bran[dově] mysli.
12. Dochází mu, že když tu holku znásilní, ona odejde a byt připadne jemu.
13. Jenže slečna Schneiderová je z francouzské měšťanské rodiny a není žádná měkkota [a] Bran[dovu] útoku klade jen nepatrný odpor.
14. Ve skutečnosti, zatímco se jí Marlon dravě zmocňuje, snaží se [ONA] pohledem odhadnout rozměry pokoje, aby věděla, jak velký koberec si má koupit.
15. Následujícího dne se v bytě [ONI] setkají znova.
16. Brando už zakoupil stůl, židle a postel, aby dal jasně najevo svoje práva.

17. Na slečnu Schneiderovou to ale neudělalo nejmenší dojem a prochází se bytem, jako by jí patřil.
18. Bran[da] to ale rozčílí natolik, že ji povalí na postel a začne mumlat:
„*Moje, moje...*“
19. Slečna Schneiderová se mu jen směje.
20. Po celou dobu milování si [ONA] prohlíží okna a snaží se přijít na to, jak dlouhé záclony si má koupit.
21. Brando, vyčerpaný a plný obav, že byt nezíská, navštěvuje svou tchyni a zesnulou manželku.
22. Vidíme nehostinný hotel a dochází nám, proč Brando tak strašně potřebuje dostat byt.
23. Slečna Schneiderová se schází se svým snoubencem a my se dovídáme, proč chce mít kde bydlet.
24. [MY] Jsme zpátky v bytě.
25. Brando je zoufalý.
26. [ON] Ukazuje slečně Schneiderové mrtvou krysou.
27. [ONA] Je otřesena, ale bytu se nevzdává.
28. Brando se tedy rozhodne ponížit ji několika pohlavními styky.
29. Jeden z nich se odehraje blízko stěny pokoje a slečně Schneiderové dojde, že pokud byt někdy dostane, bude muset koupit spoustu tapet.
30. Brandova brutalita slečnu Schneiderovou zdaleka tolik neděsí jako otázka, jak mu z ruky vykroutit klíč od bytu.
31. Když se [ONI] setkají příště, má ona na sobě svatební šaty a Branda nenapadne nic lepšího než ji hodit do vany.
32. Zázrak nad zázraky, vodovod funguje a Brando může slečnu Schneiderovou vykoupat, zatímco ona si promýšlí nejlepší barevné řešení koupelny.
33. Postupem času si Brando uvědomí, že byt nestojí za všechnu tu námahu.
34. [ON] Odejde, aniž slečně Schneiderové sdělí své pravé jméno.
35. Poněkud poznamenaná sexuálními prožitky vrací se slečna Schneiderová do bytu a přivádí s sebou snoubence, aby mu obydlí ukázala.
36. Po všem trápení, které si [ONA] prožila, musí čelit faktu, že snoubenec si byt prohlédne a řekne:
37. „[BYT] *Je příliš velký.*“

38. To je ta chvíle, kdy [ONA] začala plakat.
39. [JÁ] Nevím, jestli Poslední tango v Paříži je dobrý film, ale jsem si jist, že režisér Bertolucci významným způsobem přispěl k zachycení jednoho z vážných problémů současnosti – nedostatku obytných prostor ve Francii.

Appendix 3: Japanese Men Are Happy

1. American women could learn a lot from Japanese wives.
2. Despite subversive attempts to liberate her, the Japanese wife is still a slave to her husband, who is her “only master on earth”.
3. This has not only made for happiness in the Japanese household, but it also has kept the divorce rate down to 10 per cent, as opposed to in the United States where 25 per cent of all couples seem to find reasons to split up.
4. What are the major differences between the American and Japanese woman?
5. For one thing, the Japanese woman is much more concerned about the welfare of her husband.
6. On cold mornings, for example, I was told a good Japanese wife will prostrate herself on her husband’s side of the bed, so he won’t have to step down on a cold floor.
7. It’s small gestures like this that make for a solid marriage and happy home.
8. Another area in which Japanese women excel is giving their husbands baths.
9. There is an old Japanese proverb that a “family who washes together sloshes together,” and anyone who has been bathed by a Japanese woman will never want to take an American shower again.
10. Unlike the American woman, a Japanese wife looks forward to giving her husband a bath.
11. She stays home all day cooking the hot water so it will be just right when her husband arrives from a hard day at the office.
12. First the wife will bow to him, and then help him off with his clothes.
13. Then she’ll start scrubbing him down with soap, making sure not to get any in his eyes.

14. Finally, she'll rinse him off.
15. Only then will she allow him into the bathtub where he will soak up to his chin, while she serves him a cold beer or a hot glass of sake.
16. After the bath the wife will then massage her husband's back and even walk on him if he's really tired.
17. Then she'll dry him off and dress him for dinner.
18. By this time the husband is in a good humor and willing to listen to what the kids did in school.
19. Contrast this to the average American home where the American wife not only refuses to bow to her husband when he comes home, but in some cases won't even give him a bath.
20. And when she does give him a bath it's a slam bam, thank you ma'am, and about as romantic as a TV dinner.
21. Most American wives will run the water, hit their husbands a couple of times with a washcloth and then hand him a towel and say,
22. "Dry yourself."
23. No wonder American husbands are irritable and hard to get along with at the end of the day.
24. You would be, too, if your wife refused to walk on your back.
25. American women are afraid that if they offer to bathe their husbands, they will be considered inferior.
26. This is ridiculous.
27. A wife who knows how to bathe her husband in the Japanese style is a superior person, and one whom any husband would be proud of stepping on when he gets out of bed in the morning.

Appendix 4: Proč jsou v Japonsku muži šťastní

1. Americké manželky by se od japonských mohly mnohému přiučit.
2. Japonská žena opovrhne rozvratnými pokusy o osvo[bo]zení ženy a je stále otrokem svého manžela, který je jejím „jediným pánem na zemi“.
3. Tím jednak vstupuje do japonských domácností štěstí a zároveň se udržuje hladina rozvodů kolem deseti procent z celkového počtu uzavřených

manželství, zatímco ve Spojených státech nachází důvod jít od sebe až dvacet pět procent párů.

4. Jaké jsou hlavní rozdíly mezi japonskými a americkými manželkami?
5. Japonská žena se daleko víc stará o manželovo blaho.
6. Slyšel jsem například, že když je chladné ráno, dobrá japonská manželka si lehne manželovi před postel, aby si nemusel stoupat na studenou podlahu.
7. Právě taková drobná gesta jsou základem pevného manželství a šťastného domova.
8. Jiná oblast, v které japonské ženy vynikají, je koupání manžela.
9. Jedno japonské přísloví říká: „*jen ve společné koupeli se rodina může postříkat navzájem*“, a nikdo, koho jednou myla japonská žena, se už nechce vrátit ve sprše.
10. Na rozdíl od americké ženy se japonská manželka na koupání manžela těší.
11. [ONA] Zůstane celý den doma a ohřívá vodu, aby dosáhla správné teploty, až se manžel vrátí po těžkém dni z kanceláře.
12. Manželka se mu nejprve ukloní a pomůže mu se svlékáním.
13. Pak ho [ONA] začne mydlit a dává pozor, aby se mu mýdlo nedostalo do očí.
14. Nakonec ho [ONA] opláchne.
15. Až teprve potom mu dovolí, aby si lehl do vany a ponořil se až po bradu, zatímco ona mu běží pro vychlazené pivo nebo sklenici saké.
16. Po koupeli namasíruje manželka manželovi záda [a] když je hodně unavený, dokonce se mu po nich projde.
17. Pak ho [ONA] osuší a obleče k večeri.
18. V té chvíli je už manžel v dobré náladě a má chuť si poslechnout, co dělaly děti ve škole.
19. V americké domácnosti se manželka nejen manželovi neukloní, když přijde domů, ale někdy ho dokonce i odmítne vykoupat.
20. A pokud ho [ONA] i vykoupe, je to takové hrc frc a nazdar, příjemné asi jako večer u televize.
21. Většina Američanek svého manžela jen pocáká vodou, párkrát ho plácne žínkou, pak mu podá ručník a houkne:

22. „*Utři se sám!*“

23. Není divu, že američtí manželé jsou většinou podráždění a na sklonku dne je těžké s nimi vyjít.

24. A jak by to šlo, kdyby jim jejich manželky neodmítaly chodit po zádech!

25. Americké manželky se bojí, že by se nabídkou koupele ponížily.

26. To je nesmysl.

27. Žena, která ví, jak vykoupat manžela v japonském stylu, je bytostí nadřazenou a každý muž bude pyšný na to, že si na ni může stoupnout, když ráno leze z postele.

Appendix 5: Television Moscow-Style

1. You would think that being in Moscow during the Soviet-Czech crisis I would be very informed about what was going on.
2. But the truth of the matter is I didn't even know there was a Soviet-Czech crisis until I left the country.
3. I blame the Soviet television network for this.
4. For some reason, which I'll never understand, Soviet TV just isn't doing the job when it comes to informing the public.
5. I know this because I had a television set in my room at the hotel, and I watched it constantly to see if the Soviets were doing any better with their electronics media than we were doing with ours.
6. It would be unfair to say Soviet television is better than American television – it's just different.
7. The first time I turned on the set I hit it lucky.
8. There was an exciting film on, showing how trucks were assembled in a Soviet factory.
9. This program was followed by a visit to a Soviet stone quarry.
10. For two hours I sat glued to the set watching giant machines tear into a mountain and chew up stones until they were turned into gravel.
11. During a lull in the program, I decided to change channels (Moscow is supposed to have four), and lo and behold, there was a panel show with a moderator who looked just like David Susskind.

12. Seated around the table were a welder, a dam builder, a woman lumberjack and the minister of canal barges.
13. My Russian wasn't good enough to understand what they were saying, but, as in America, the fellow who looked like David Susskind was doing all the talking.
14. I switched to a third channel and got a beautiful test pattern.
15. The fourth channel didn't seem to be working, so I went back to the first channel where the stone quarry program had just ended.
16. It was followed by a visit to a pipeline worker's congress being held in Byelorussia.
17. By this time my wife wanted to go sightseeing, but I was so grabbed by the speeches I told her to wait.
18. She could see the Kremlin any time, but how often could she see a pipeline worker's congress on television?
19. Two hours later I watched steel being made in a large foundry outside Moscow, and this program was followed by an interview with a professor of Siberian tree transplants.
20. I turned the channel and the panel was still in session, with the fellow who looked like David Susskind still doing the talking.
21. My wife kept insisting we leave the room but I decided to have one more go at the third channel, and I'm glad I did.
22. Instead of featuring a test pattern the third channel was now showing a feature film on sheepshearing in Mongolia.
23. Even in black and white it was the highlight of Soviet TV that week.
24. By this time my wife was getting pretty angry so I reluctantly turned off the set and went out sightseeing with her.
25. When we got back to the room eight hours later I immediately turned on the set.
26. The David Susskind-type panel was still going on, and the guy was still doing the talking.

Appendix 6: Moskevská televize

1. Řekli byste možná, že když jsem byl během československo-sovětské krize v Moskvě, byl jsem o vývoji situace informován velmi dobře.
2. Ve skutečnosti jsem se o tom, že nějaká krize je, dověděl, až když jsem zemi opustil.
3. A může za to sovětská televize.
4. Z jakéhosi důvodu, který nikdy nepochopím, nevěnuje se sovětská televize vůbec tomu, aby informovala veřejnost.
5. Vím to, protože jsem měl televizor ve svém hotelovém pokoji a televizi jsem soustavně sledoval, abych zjistil, jestli si vedou sovětská elektronická média lépe než naše.
6. Bylo by nespravedlivé tvrdit, že sovětská televize je lepší než americká – je prostě jiná.
7. Hned když sejm poprvé zapnul přístroj, měl jsem velké štěstí.
8. Dávali zrovna vzrušující film o tom, jak se v sovětské továrně montují nákladní automobily. ~~Film představoval nejen oddané a uvědomělé dělníky montující vůz, ale i to, jak se automobil maže a zkouší.~~
9. Dále byl na programu pořad ze sovětského kamenolomu.
10. Dvě hodiny jsem nebyl schopen odlepit se od židle a zíral jsem na obrovské stroje zakusující se do skály, vylamující kameny a drtící je na šterk.
11. Během přestávky jsem se rozhodl přepnout na jiný kanál (v Moskvě mají být čtyři) [a] [v]padl jsem doprostřed panelové diskuse, kterou řídil muž podobný Davidu Susskindovi.
12. Kolem stolu seděli svářeč, stavbař z přehrady, dřevorubkyně a ministr vlečných člunů.
13. Neumím rusky natolik, abych rozuměl tomu, o čem si povídali, ale podobně jako v Americe obstaral veškeré mluvení ten chlápek, co vypadal jako David Susskind.
14. Přepnul jsem na třetí kanál a našel tam monoskop.
15. Čtvrtý kanál nevysílal ani to, tak jsem se vrátil na první program, kde zrovna končil pořad o zpracování kamene.
16. Následovala reportáž o schůzi dělníků na stavbě ropovodu v Bělorusku.

17. Žena projevila přání jít se podívat po městě, ale já byl tak fascinován projevem na schůzi, že jsem ji požádal, aby počkala.
18. Kreml může vidět kdykoliv, ale kdy se člověku poštěstí vidět v televizi schůzi dělníků ze stavby ropovodu?
19. O dvě hodiny později jsem se už díval na to, jak se v jedné huti u Moskvy vyrábí ocel [a] [p]ak následoval rozhovor s jakýmsi profesorem ze sibiřské ovocnářské šlechtitelské stanice.
20. Přepnul jsem na druhý program, ale tam byla pořád ještě ta panelová diskuse a chlápek, co vypadal jako David Susskind, pořád ještě mluvil.
21. Manželka bez ustání naléhala, abychom šli ven, ale já se rozhodl, že se ještě mrknu na třetí program [a] [u]dělal jsem dobře.
22. Místo monoskopu tam vysílali hraný film o pěstování ovcí v Mongolsku.
23. Bylo to sice černobílé, ale i tak to byl vrchol všeho, co mi zatím sovětská televize nabídla.
24. Manželka už začínala být pěkně naštvaná, [a tak jsem ihned] vypnul televizor a vydal se na prohlídku města.
25. Když jsme se o osm hodin později vrátili, okamžitě jsem se vrhl k televizoru.
~~Dva z programů vysílaly monoskop, ale~~
26. [D]iskuse řízená chlápem podobným Davidu Susskindovi ještě běžela a chlápek pořád ještě mluvil.

Appendix 7: LBJ Betrayed in Portrait

1. It seems a pity that after all President Johnson has done for the arts, the least the artists could do is paint a decent portrait of him.
2. Last week it was revealed that the president angrily rejected artist Peter Hurd's commissioned portrait of him as "the ugliest thing I ever saw".
3. Mr. Johnson, whose taste in paintings leans more toward Norman Rockwell's magazine style, objected to the Hurd portrait on the grounds that it was too large, the capitol building in the background was too prominently lit and "inappropriate," and the positioning of the figure and the general style were not consistent with other White House portraits.

4. Many people, including art critics and museum directors, have come to the defense of Mr. Hurd, but as has been the case lately, very few people have come to the defense of the president.
5. I wish to be included in the minority.
6. I believe the president had every right to reject a portrait of himself if it didn't fit the image he has of himself.
7. The likeness of the president as I saw it in photographs leaves much to be desired.
8. In Mr. Hurd's portrait the president seems to be looking off dreamily into the distance as if searching for a consensus, of a friendly congressman.
9. His lips are set as though he has just had a conference with a group of Democratic governors.
10. His chin looks as if he's just heard a report from Ambassador Henry Cabot Lodge on Vietnam, and his nose looks as if he's just smelled another resignation from his White House staff.
11. The suit he is wearing is dark and conservative with a vest giving the impression the president might be a wealthy real estate man, or a successful television station owner, and the book he is holding in his hand looks as if it could have been written by Arthur Schlesinger.
12. But forgetting the portrait for a moment, anybody could see why the president would object to the way Mr. Hurd positioned the capitol in the background.
13. Anyone who knows the president's personality should have been aware that the only way to pose the president in such a picture was to have him standing in the capitol itself with one foot on the senate and the other on the house of representatives.
14. Also, if Mr. Hurd has studied his subject more closely, he would have painted the light shining over the capitol coming from the president's face and not from within the capitol itself.
15. In the commissioned portrait, Mr. Johnson objected to the size of the capitol, which is one-sixth the size of the president.
16. The implication can be clearly drawn from the portrait that the capitol plays a great part in the president's life, when anyone who lives here knows the exact opposite is true.
17. I think Mr. Hurd owes it to the president to try again.

18. He could start by showing the president smiling ~~as if he just heard that Bobby Kennedy has an accident while skiing.~~
19. Instead of a book, he could be holding a 1964 Lou Harris or George Gallup opinion poll.
20. And instead of the capitol, Mr. Hurd could paint in the LBJ ranch ~~or the Pedernales river.~~
21. Many museums are bidding for Mr. Hurd's portrait, but I think it would be unfair if it were publicly displayed, particularly when people know the president doesn't approve of it.
22. You can say anything you want to about Mr. Johnson's taste in art, but at least he knows what he likes.

Appendix 8: Prezidentův portrét

1. Je vážně škoda, že po všem, co prezident Johnson vykonal pro umění, nejsou mu umělci schopni ani namalovat pořádný portrét.
2. Před nedávnem dokonce musel portrét předložený Peterem Hurdem odmítnout slovy, že je to „*nejošklivější věc, jakou kdy viděl*“.
3. Pan Johnson, jehož uměleckému vkusu odpovídají spíš časopisecké ilustrace Normana Rockwella, Hurdovu dílu vytkl především to, že obraz je příliš veliký, že budova Kapitolu v pozadí je příliš nápadně osvětlena a „*nevhodná*“ a že výraz portrétované osoby a celkové vyznění portrétu neodpovídá jiným podobiznám v Bílém domě.
4. Na obranu pana Hurda vystoupilo mnoho lidí, včetně uměleckých kritiků a ředitelů muzeí, ale jak se stalo v poslední době zvykem, na pomoc prezidentovi přispěchal jen málokdo.
5. Rád bych se teď do této menšiny zařadil.
6. [JÁ] Jsem přesvědčen, že prezident má plné právo odmítnout vlastní portrét, neodpovídá-li jeho představám o sobě samém.
7. Prezidentově podobě, jak [JÁ] ji znám z fotografií, zůstal portrét rozhodně mnohé dlužen.
8. Na Hurdově portrétu se prezident zasněně dívá do dálky, jako kdyby tam hledal všeobecné porozumění či sympatizujícího kongresmana.

9. Jeho rty jako by napovídaly, že se právě zúčastnil porady se skupinou demokratických guvernérů.
10. Jeho tváře jako by ukazovaly, že mu zrovna velvyslanec Ellsworth Bunker¹⁷ přednesl zprávu z Vietnamu [a] [n]os působí dojmem, že právě zavěťřil další rezignaci v Bílém domě.
11. Prezident je oblečen do tmavého usedlého obleku s vestou, který vzbuzuje dojem, že Johnson je bohatým majitelem nemovitostí či úspěšným vlastníkem televizní stanice [a] kniha, kterou drží v ruce, vypadá, jako by ji napsal Arthur Schlesinger.
12. Necháme-li chvíli stranou podobu prezidenta a věnujeme-li se zbytku obrazu, je nám na první pohled jasné, proč Johnson nesouhlasil se způsobem, jakým je zachycen Kapitol.
13. Každý, kdo prezidenta aspoň trochu zná, musí bez zaváhání uznat, že Johnson může být zobrazen jedině tak, že bude na Kapitolu stát s jednou nohou na Senátu a druhou na Sněmovně reprezentantů.
14. Kdyby pan Hurd prostudoval svůj objekt podrobněji, musel by také onu zář namalovat tak, aby vycházela z prezidentovy hlavy, a ne z Kapitolu.
15. Prezident Johnson měl námitky i proti velikosti Kapitolu, vždyť na obrazu dosahuje budova výšky jedné šestiny jeho postavy.
16. Obraz tak napovídá, že Kapitol hraje v životě prezidenta důležitou úlohu [a] [p]řítom každý z nás ví, že opak je pravdou.
17. Myslím, že pan Hurd dluží prezidentovi ještě jeden pokus.
18. Mohl by [HURD] začít tím, že prezidenta zachytí s úsměvem ve tváři.
19. Místo knihy by přitom [PREZIDENT] mohl třímat výsledky průzkumu veřejného mínění Harrisova nebo Gallupova ústavu z roku 1964.
20. Místo Kapitolu by pan Hurd mohl v pozadí namalovat Johnsonův ranč.
21. O portrét namalovaný panem Hurdem mají muzea veliký zájem, ale podle mě by nebylo správné ho vystavit veřejně, zejména když lidé vědí, že s tím prezident nesouhlasí.
22. O uměleckém vkusu prezidenta Johnsona si můžete myslet, co libo – nelze však popřít, že aspoň ví, co chce.

¹⁷ Ambassador Henry Cabot Lodge in the ST

SUMMARY

Následující resumé představuje shrnutí nejdůležitějších informací prezentovaných v této diplomové práci na téma Lexikální koheze v publicistických textech.

Základními parametry, jež definují text, jsou dle Beaugranda a Dresslera (1981) tzv. parametry textovosti. Beaugrande a Dressler definují sedm těchto parametrů. Jsou jimi koheze, koherence, situacionalita, informativita, intertextualita, intencionalita a akceptabilita. Tato diplomová práce se zabývá prvním z těchto sedmi parametrů, tedy kohezí, konkrétně jednou z jejích kategorií, kterou je lexikální koheze. Základem pro analýzu lexikální koheze je Hoeyho (1991) metoda. Pro komparativní analýzu kohezivních prostředků v češtině a angličtině byly zvoleny texty publicistického funkčního stylu.

Druhá kapitola této práce je věnována kohezi a koherenci, dvěma parametrům textovosti, jež spolu úzce souvisejí. Kohezi se rozumí propojení jednotlivých částí textu (Halliday a Hasanová 1976), a to na povrchové úrovni (Tárnyiková 2002), zatímco koherence označuje vzájemné propojení sémantických složek textu a jejich interakci s kontextem v závislosti na interpretaci čtenářem (ibid.). Koheze je tedy vlastností textu, kdežto koherence závisí na interpretaci, a může tedy být chápána jako subjektivní (Tanskanenová 2006). Kohezivní prostředky lze klasifikovat různými způsoby. Tato práce prezentuje klasifikace Hallidaye a Hasanové (1976), Beaugranda a Dresslera (1981) a Dooleyho a Levinsohna (2000) a poukazuje na jejich výhody a nevýhody. Vycházíme-li z první zmíněné klasifikace, můžeme prostředky koheze rozdělit do pěti skupin, a to na referenci, substituci, spojku, elipsu a lexikální kohezi. A právě této poslední kategorii se práce věnuje podrobněji.

Lexikální koheze je chápána jako prostředek propojující text pomocí lexikálních jednotek. Jinými slovy, výběr slovní zásoby napomáhá celkové soudržnosti textu (Halliday a Hasanová 1976). I v případě lexikální koheze práce prezentuje několik klasifikací – vychází z modelu Hallidaye a Hasanové (1976), který byl následně Hasanovou (1981) přepracován, a zmíněné modely doplňuje o klasifikaci Hoeyho (1991). Právě tato klasifikace je pro praktickou část práce stěžejní.

Třetí kapitola se věnuje Hoeyho metodě pro analýzu lexikální koheze. Hoey (1991) chápe lexikální kohezi jako základní a nejrozšířenější kohezivní

prostředek. Definuje čtyři základní typy lexikální koheze. Prvním typem je „jednoduché opakování“ (simple repetition), kdy se daná jednotka v textu opakuje. V této kategorii jsou povoleny změny v rámci gramatického paradigmatu. Druhou kategorií je „komplexní opakování“ (complex repetition), kdy se opakují jednotky se stejným slovním základem, které jsou nicméně realizovány různými slovními druhy. Do třetí kategorie, „jednoduchá parafráze“ (simple paraphrase), spadají synonyma. V rámci poslední, čtvrté kategorie, „komplexní parafráze“ (complex paraphrase) Hoey vymezuje antonymii a tzv. „link triangle“, kdy dva vztahy mezi jednotkami vytvářejí vztah třetí, přičemž jednotka, která druhé dvě propojuje, může, ale nemusí být explicitně přítomna v textu. Jinak řečeno, existuje-li vztah mezi A a B a zároveň mezi A a C, automaticky tím vzniká i třetí vztah mezi B a C, přičemž A je jednotkou, která nemusí být v textu explicitně přítomna, a přesto může fungovat jako pojítka mezi jednotkami B a C. Mimo to Hoey vymezuje ještě případy ko-reference, substituce, elipsy a hyponymie/hyperonymie. Výše zmíněné čtyři základní kategorie jsou uvedeny v pořadí od nejsilnějšího typu vazby po nejslabší. Opakování slova je tedy nejsilnější formou vazby.

Propojení jednotlivých lexikálních jednotek v textu vytváří vzorce. A právě analýza těchto vzorců je základem pro Hoeyho metodu sloužící ke tvorbě shrnutí textu. V rámci analýzy se na základě výše zmíněných kategorií identifikují vazby mezi jednotkami v textu. Každá taková vazba je označována jako „link“. Hoey podotýká, že jednotlivá vazba sama o sobě nemá v rámci celého textu zásadní důležitost, ale nabývá významu až v kombinaci s ostatními vazbami. Proto charakterizuje tzv. „bond“, který je tvořen třemi jednotlivými vazbami (linky). Základní principy analýzy jsou následující: (1) identifikují se pouze vztahy mezi plnovýznamovými lexikálními jednotkami, (2) do analýzy se nezapočítávají vztahy mezi jednotkami uvnitř téže věty a (3) tatáž jednotka v jedné větě nemůže sdílet vazbu se dvěma jednotkami jiné věty. Práce ale zároveň poukazuje na výjimky, které Hoey uvádí (např. s ohledem na první zmíněné pravidlo, kde poukazuje na to, že v konkrétním textu je možno započítat i slova neplnovýznamová). Následně se dle počtu vazeb (bondů) s ostatními větami určí „centrální“ a „marginální“ věty textu (přičemž počet vazeb nezbytný pro „centrální“ věty lze nastavit individuálně dle daného textu). Na základě této

klasifikace lze pak různými postupy tvořit shrnutí textu. Součástí popisu Hoeyho metody jsou i některé výzkumné projekty, které tuto metodu využily.

Jak již bylo zmíněno, cílem této práce je komparativní analýza prostředků lexikální koheze v publicistických textech v češtině a angličtině. Práce proto nabízí přehled kohezivních prostředků v češtině a angličtině dle Lukeše (2004).

Čtvrtá kapitola se věnuje charakteristice publicistického textu a představuje texty vybrané pro výzkumnou část práce. Jako texty pro analýzu sloužily čtyři sloupky amerického autora Arta Buchwalda ze souboru *Have I Ever Lied To You?* přeloženého do češtiny Janem Jirákem jako *Copak jsem vám někdy lhal?* Výběr publicistických textů zároveň umožnil testování využitelnosti Hoeyho metody pro analýzu lexikální koheze v tomto typu textu a v obou daných jazycích.

V páté a šesté kapitole je prezentován samotný výzkum. Na základě výše zmíněného byly vymezeny tři výzkumné otázky, jež měla práce za cíl zodpovědět. První hypotézou bylo, že Hoeyho metoda bude snáze aplikovatelná na texty v angličtině, vzhledem k tomu, že pro tento jazyk byla vytvořena. Horší využitelnost pro češtinu byla očekávána vzhledem k systemickým rozdílům mezi těmito dvěma jazyky. Druhá hypotéza se týkala četnosti výskytu linků a bondů v obou jazycích. V návaznosti na Molnárův (2012) výzkum byla formulována hypotéza, že vlivem rozdělení vět oproti anglickému originálu při překladu do češtiny bude anglická verze textů obsahovat více bondů, a naopak čeština bude vykazovat vyšší počet linků. Poslední hypotézou pak bylo, že vzhledem k vyšší toleranci angličtiny k opakování bude anglická verze textů obsahovat vyšší počet linků v první kategorii (simple repetition), kdežto v české verzi bude více linků v ostatních třech kategoriích.

Pro komparativní analýzu bylo nutné texty upravit tak, aby obsahovaly stejné informace. Vlivem překladu se totiž jazykové verze mírně lišily (překladatel na některých místech text rozšířil či informaci vypustil). Bylo tedy nutné vynechat ty části textu, které ve druhé verzi chyběly. Zároveň bylo nutno v češtině doplnit referenci v případech, kdy nebyla v textu explicitně uvedena, vzhledem k tomu, že čeština disponuje možností podmět vyjádřit implicitně pomocí verbální koncovky. Texty byly rozčleněny na jednotlivé věty. Ty byly očíslovány podle pořadí, v němž se v textu vyskytovaly. Pro účely výzkumu byla Hoeyho metoda omezena

na čtyři základní kategorie – jednoduché a komplexní opakování a jednoduchou a komplexní parafrázi (simple repetition, complex repetition, simple paraphrase a complex paraphrase). Případy elipsy, substituce, ko-reference a hyponymie/hyperonymie byly započítány do čtvrté kategorie. Zároveň bylo třeba upravit metodu pro aplikaci na český text. Problematickými byly zejména stupňování adjektiv a verbální negace, kde se projevíly systemické tendence angličtiny k analytickému vyjadřování oproti syntetickým formám češtiny. Další problematickou oblastí byla vlastní jména a možnosti jejich klasifikace v rámci Hoeyho čtyř kategorií.

Analýza textu probíhala manuálně. V první fázi došlo k úpravě textů. Následně byly analyzovány linky v jednotlivých kategoriích, od nejsilnější formy vazby, tedy jednoduchého opakování (simple repetition) po nejslabší, komplexní parafrázi (complex paraphrase). Tím se zamezilo zaznamenání slabého linku na místě, kde se zároveň vyskytoval i link silnější.

Linky byly zaznamenávány „dvakrát“, tedy k oběma větám, které danou vazbu sdílely. Následně byly identifikovány bondy (tvořené třemi linky mezi danými větami). Ty byly přeneseny do kvantitativního matrixu. Dvojí zaznamenání linků a bondů v této fázi sloužilo jako kontrolní mechanismus, protože daný link nebo bond musel být vždy zaznamenán u obou dvou vět, které ho tvořily.

Součástí práce je i podrobný popis jednotlivých textů a problému při aplikaci Hoeyho metody. Vzhledem k tomu, že Hoeyho metoda, jak již bylo zmíněno, slouží pro tvorbu shrnutí textu, součástí prezentace výsledků analýzy jsou i ilustrativní shrnutí.

Na základě výstupů z analýzy jsou v sedmé kapitole formulovány tyto závěry: (1) Hoeyho metoda je snáze aplikovatelná na angličtinu, (2) anglická verze textů obsahovala větší počet bondů než česká a (3) anglická verze textů obsahovala vyšší počet bondů v první kategorii, zatímco v české verzi byl vyšší počet bondů v ostatních kategoriích. Nepotvrdila se však druhá část hypotézy 2, a to že česká verze textů bude obsahovat více linků, ale méně bondů než verze anglická. V tomto případě byly výsledky nejednoznačné (očekávaná tendence se projevila pouze u dvou ze čtyř analyzovaných textů).

Vzhledem k rozsahu výzkumu jsou nicméně tyto výsledky pouze prozatímní a k jejich všeobecnému potvrzení či vyvrácení by bylo zapotřebí rozsáhlého výzkumu, který však dalece překračuje rozsah této práce i možnosti manuální analýzy prováděné jednotlivcem.

LIST OF REFERENCES

Primary sources:

- Buchwald, Art. 1969. *Have I Ever Lied to You?* Greenwich: Fawcett Crest.
- Buchwald, Art. 1992. *Copak jsem vám někdy lhal?* Translated by Jan Jiráček. Praha: Nakladatelství Lidové noviny.

Secondary sources:

- Ahmad, Khurshid and Mohamed Benbrahim. 1995. "Text Summarisation: The Role of Lexical Cohesion Analysis". In *The New Review of Document and Text Management*, Vol 1, 321-336. Accessed April, 1, 2016.
https://www.researchgate.net/profile/Khurshid_Ahmad/publication/278967303_Text_Summarisation_The_Role_of_Lexical_Cohesion_Analysis/links/5588a92a08ae8c4f34065178.pdf
- de Beaugrande, Robert-Alain and Wolfgang Dressler. 1981. *Introduction to Text Linguistics*. London: Longman.
- Berber Sardinha, Tony. 1999. "Looking at discourse in a corpus: The role of lexical cohesion." Paper presented at the 12th World Congress of Applied Linguistics (AILA), Waseda University, Japan. Accessed April 22, 2015.
<http://arxiv.org/ftp/cs/papers/0004/0004016.pdf>
- Čechová, Marie, Krčmová, Marie, and Eva Minářová. 2008. *Současná stylistika*. Praha: Nakladatelství Lidové noviny.
- Chloupek, Jan. 1993. Pochybnosti o funkčních stylech? In *Sborník prací Filozofické fakulty Brněnské univerzity*. 61-71.
- Christiansen, Thomas. 2011. *Cohesion: a discourse perspective*. New York: Peter Lang.
- Crystal, David and Derek Davy. 1969. *Investigating English Style*. Indiana University Press.
- Daneš, František. 1985. *Věta a text: studie ze syntaxe spisovné češtiny*. Praha: Academia.

- Dooley, Robert and Stephen H. Levinsohn. 2001. *Analyzing Discourse: A Manual of Basic Concepts*. Dallas: SIL International.
- Firbas, Jan. 1971. "On the Concept of Communicative Dynamism in the Theory of Functional Sentence Perspective." In *Sborník prací Filosofické fakulty Brněnské univerzity*, 135-144.
- Halliday, M. A. K. and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.
- Hasan, Ruqaiya. 1984. "Coherence and cohesive harmony." In *Understanding Reading Comprehension*, edited by J. Flood. Delaware: International Reading Association.
- Hoey, Michael. 1991. *Patterns of Lexis in Text*. Oxford: Oxford University Press.
- Hoey, Michael. 2001. *Textual Interaction: An introduction to written discourse analysis*. New York: Routledge.
- Knittlová, Dagmar, et al. 2010. *Překlad a překládání*. 1st edition. Olomouc: UP.
- Károly, Krisztina. 2010. "Shifts in repetition vs. shifts in text meaning: A study of the textual role of lexical repetition in non-literary translation." In *Target 22.1*: 40-70.
- Lukeš, Dominik. 2004. *Cohesion in Czech and English: A corpus-based overview of connective devices*. Unpublished paper. Accessed April 22, 2015.
http://www.academia.edu/7485954/Cohesion_in_Czech_and_English_A_corpus-based_overview_of_connective_devices
- McNair, Brian. 2008. "I, Columnist." In *Pulling newspapers apart: analyzing print journalism*, edited by B. Franklin. 106-114. New York: Routledge.
- Molnár, Ondřej. 2012. "Lexical Patterns in the Texts of the European Union." In *Text as a Dynamic Interplay of Text Parameters*. 31-45. Olomouc: VUP.
- Morley, John. 2009. "Lexical cohesion and rhetorical structure." In *Lexical Cohesion and Corpus Linguistics*, edited by J. Flowerdew and M. Mahlberg, 5-22. Philadelphia: John Benjamins.
- Oakland, A. P. and Ludmila Urbanová. 2002. *Úvod do anglické stylistiky*. Brno: Barrister & Principal Brno 2002

de Oliveira, P. C.F., Ahmad, Khurshid and Lee Gillam. 2002. "A financial news summarization system based on lexical cohesion." In *Proceedings of the International Conference on Terminology and Knowledge Engineering, Nancy, France*. Accessed April 22, 2015.

http://antigo.univille.br/arquivos/4694_Paper3.pdf

Richards, Jack C. and Richard Schmidt. 2010. *Longman Dictionary of Language Teaching and Applied Linguistics*. Longman Publishing Group.

Sanders, T. J. M. and Henk Pander Maat. 2006. "Cohesion and coherence: Linguistic approaches." In *Encyclopedia of Language and Linguistics*, edited by Beown, K. et al. 2nd edition, Volume 2. 591-595. London: Elsevier.

Sinclair, John. 2004. *Trust the text: language, corpus and discourse*. London: Routledge.

Tanskanen, Sanna-Kaisa. 2006. *Collaborating towards Coherence: Lexical cohesion in English Discourse*. Philadelphia: John Benjamins.

Tárnyiková, Jarmila. 2002. *From Text to Texture*. Olomouc: Univerzita Palackého v Olomouci.

Widdowson, H. G. 2004. *Text, Context, Pretext: Critical Issues in Discourse Analysis*. Oxford: Blackwell.

ANNOTATION

Autor: Ivana Tůmová

Katedra: Katedra anglistiky a amerikanistiky Filozofické fakulty Univerzity Palackého v Olomouci

Název česky: Lexikální koheze v publicistických textech: komparativní analýza

Název anglicky: Lexical Cohesion in Publicistic Texts: a Comparative Analysis

Vedoucí práce: Mgr. Ondřej Molnár

Počet stran (úvod – závěr): 95

Počet stran (celkově): 127

Počet znaků (úvod – závěr): 168 559

Počet znaků (celkově): 214 753

Počet titulů použité literatury: 30

Klíčová slova v ČJ: komparativní analýza, lexikální koheze, Hoey, sloupek, čeština, angličtina

Klíčová slova v AJ: comparative analysis, lexical cohesion, Hoey, column, Czech, English

Abstract:

This Master's thesis aimed at investigating comparatively the devices of lexical cohesion in Czech and English sample texts of publicistic style by means of Hoey's methodology for the analysis of lexical patterns in texts. Cohesion, one of the seven standards of textuality, refers to the linkage among the individual elements of a text. One type of these surface-level linkage devices is lexical cohesion, established through the choice of lexicon. Hoey's method of analysis of lexical patterns is primarily aimed as a means for creating abridgements or non-narrative texts. The research used this method as the basis for a comparative analysis of the devices of lexical cohesion in Czech and English columns by Art Buchwald, translated into Czech by Jan Jiráček.

Abstrakt:

Tato magisterská diplomová práce se zabývá komparativním výzkumem prostředků lexikální koheze v českých a anglických textech publicistického stylu pomocí Hoeyho metody analýzy lexikálních vzorců v textu. Koheze je jedním ze sedmi parametrů textovosti. Jedná se o vzájemnou propojenost mezi jednotlivými částmi textu. Jedním typem těchto prostředků propojujících text je lexikální koheze, která je realizována pomocí výěru slovní zásoby v textu. Hoeyho metoda analýzy lexikálních vzorců je primárně metodou pro tvorbu výtahů nenarativních textů. V tomto výzkumu je tato metoda využita jako základ pro komparativní analýzu prostředků lexikální koheze v českých a anglických verzích sloupků Arta Buchwalda přeložených do češtiny Janem Jirákem.