

Univerzita Karlova

Filozofická fakulta

Ústav anglického jazyka a didaktiky

Diplomová práce

Univerzita Karlova

Filozofická fakulta

Ústav anglického jazyka a didaktiky

Diplomová práce

Bc. Adéla Coufalová

**Vybraná sponová slovesa a jejich adjektivní doplnění v současné
psané a mluvené angličtině**

**Selected copular verbs and their adjectival complements in
contemporary written and spoken English**

Tímto děkuji Mgr. Lucii Lukešové, Ph.D. za odborné vedení a ochotu při konzultacích i zpracování této práce a za cenné rady týkající se fungování korpusů.

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 2. srpna 2019

Souhlasím se zapůjčením diplomové práce ke studijním účelům.

I have no objections to the MA thesis being borrowed and used for study purposes.

Abstract

The present thesis studies the copular verbs and their co-occurrence with adjectival complements in contemporary written and spoken English. The theoretical part of this work first describes the copular verbs in terms of their types and types of complements they take. A corpus-based approach to copular verbs follows. A part of the theory also provides an overview of the terminology that is used with copular verbs.

The work consists of two parts: the theoretical part which introduces the topic and describes the features of both the principal copula *be* and other copular verbs and the analytical part which analyses written language data excerpted from the Araneum Anglicum corpus and spoken language data from the Spoken BNC2014 corpus. The analytical part focuses on the following seven verbs: *feel*, *look*, *seem*, *remain*, *become*, *go* and *get* and their adjectival complements. These adjectives are then analysed in terms of their frequency and collocability with the verbs. The collocate lists of all verbs are then compared between the two corpora. Finally, the analysis also attempts to categorize the adjectives semantically. The analysis also focuses on adverbs which occur in these constructions and which function is to modify the adjectives.

The results of the analytical part are summarized in the conclusion.

Abstrakt

Tato práce se zaměřuje na sponová slovesa a jejich adjektivní doplnění v současné psané a mluvené angličtině. V části věnované teorii jsou sponová slovesa a terminologie s nimi spojená popsána tak, jak je vymezují různí autoři. Popis se zaměřuje na jednotlivé typy sponových sloves, možnosti jejich doplnění a sémantický popis sponové predikace. Následně jsou sponová slovesa popsána na základě výsledků výzkumů prováděných na korpusech.

Tato práce je rozdělena do dvou částí: první částí je teoretické pozadí sponových sloves. Druhá část se věnuje analýze jazykových dat psané angličtiny získaných z Araneum Anglicum korpusu a jazykových dat mluvené angličtiny získaných z Spoken BNC2014 korpusu. Analýza se zaměřuje na následujících sedm sloves *feel*, *look*, *seem*, *remain*, *become*, *go* a *get* a jejich adjektivní doplnění. Tato adjektiva jsou poté zkoumána z hlediska jejich

frekvence a kolokability s danými slovesy. Seznamy kolokátů jednotlivých sloves jsou následně porovnány mezi korpusy.

Snahou analýzy také je sémanticky rozřadit adjektiva do různých kategorií. Analýza se také zaměřuje na adverbia, která se v těchto konstrukcích často objevují modifikující daná adjektiva.

Závěr poskytuje shrnutí výsledků analýzy.

Key words

copular verb, subject complement, verbless adverbial clause, adjectival complement, frequency of occurrence, collocability, semantic categories of adjectives

Klíčová slova

sponové sloveso, jmenná část přísudku, doplněk podmětu, adjektivní doplnění, frekvence výskytu, kolokabilita, sémantické kategorie adjektiv

Table of Contents

1	Introduction	13
2	Theoretical background	14
2.1	Copular verbs in theory	15
2.1.1	Definition and terminology	15
2.1.2	The subject complement	16
2.1.3	The principal copular verb <i>be</i>	17
2.1.4	Types of copular verbs	18
2.1.5	Subject complement and its representation	21
2.2	A corpus-based approach to copular verbs	25
2.2.1	Biber and Conrad 2009	25
2.2.2	Malá 2014	26
3	Material and Method	29
3.1	Material	29
3.1.1	Corpora	29
3.1.2	Selected copular verbs	30
3.2	Method	30
4	Analysis	34
4.1	The verb <i>feel</i>	36
4.1.1	Comparing frequency and collocate lists within a single corpus	39
4.1.2	Comparing collocate lists across the corpora	41
4.1.3	Semantic categories of the adjectives	42
4.1.4	Adverbs occurring in the structure	43
4.2	The verb <i>look</i>	45
4.2.1	Comparing frequency and collocate lists within a single corpus	48
4.2.2	Comparing collocate lists across the corpora	49
4.2.3	Semantic categories of the adjectives	50
4.2.4	Adverbs occurring in the structure	51
4.3	The verb <i>seem</i>	53
4.3.1	Comparing frequency and collocate lists within a single corpus	56
4.3.2	Comparing collocate lists across the corpora	57
4.3.3	Semantic categories of the adjectives	59
4.3.4	Adverbs occurring in the structure	60

4.4	The verb <i>remain</i>	61
4.4.1	Comparing frequency and collocate lists within a single corpus.....	65
4.4.2	Comparing collocate lists across the corpora	66
4.4.3	Semantic categories of the adjectives	67
4.4.4	Adverbs occurring in the structure.....	68
4.5	The verb <i>become</i>	69
4.5.1	Comparing frequency and collocate lists within a single corpus.....	73
4.5.2	Comparing collocate lists across the corpora	74
4.5.3	Semantic categories of the adjectives	75
4.5.4	Adverbs occurring in the structure.....	76
4.6	The verb <i>go</i>	77
4.6.1	Comparing frequency and collocate lists within a single corpus.....	81
4.6.2	Comparing collocate lists across the corpora	83
4.6.3	Semantic categories of the adjectives	84
4.6.4	Adverbs occurring in the structure.....	85
4.7	The verb <i>get</i>	87
4.7.1	Comparing frequency and collocate lists within a single corpus.....	90
4.7.2	Comparing collocate lists across the corpora	92
4.7.3	Semantic categories of the adjectives	92
4.7.4	Adverbs occurring in the structure.....	93
5	Conclusion.....	95
6	Bibliography	97
7	Resumé	99

List of abbreviations

AAM Araneum Anglicum Minus corpus

SBNC Spoken BNC2014 corpus

List of Tables

Table 1: AAM: a frequency list of <i>feel</i>	36
Table 2: AAM: a collocate list of <i>feel</i>	37
Table 3: SBNC: a frequency list of <i>feel</i>	37
Table 4: SBNC: a collocate list of <i>feel</i>	38
Table 5: AAM: a frequency list of <i>look</i>	45
Table 6: AAM: a collocate list of <i>look</i>	46
Table 7: SBNC: a frequency list of <i>look</i>	46
Table 8: SBNC: a collocate list of <i>look</i>	47
Table 9: AAM: a frequency list of <i>seem</i>	53
Table 10: AAM: a collocate list of <i>seem</i>	54
Table 11: SBNC: a frequency list of <i>seem</i>	54
Table 12: SBNC: a collocate list of <i>seem</i>	55
Table 13: AAM: a frequency list of <i>remain</i>	62
Table 14: AAM: a collocate list of <i>remain</i>	62
Table 15: SBNC: a frequency list of <i>remain</i>	63
Table 16: SBNC: a collocate list of <i>remain</i>	64
Table 17: AAM: a frequency list of <i>become</i>	70
Table 18: AAM: a collocate list of <i>become</i>	70
Table 19: SBNC: a frequency list of <i>become</i>	71

Table 20: SBNC: a collocate list of <i>become</i>	72
Table 21: AAM: a frequency list of <i>go</i>	78
Table 22: AAM: a collocate list of <i>go</i>	79
Table 23: SBNC: a frequency list of <i>go</i>	80
Table 24: SBNC: a collocate list of <i>go</i>	80
Table 25: AAM: a frequency list of <i>get</i>	87
Table 26: AAM: a collocate list of <i>get</i>	88
Table 27: SBNC: a frequency list of <i>get</i>	89
Table 28: SBNC: a collocate list of <i>get</i>	90

1 Introduction

The aim of the present thesis is to describe and analyse selected copular verbs and their adjectival complements in contemporary written and spoken English based on the frequency of occurrence and the collocational strength of the combinations in contemporary language corpora. It is expected that the adjectival complements will differ depending on the verb, method used as well as the analysed registers (written or spoken).

Copular verbs are defined as verbs which express relations and attribute properties and which constitute the verbo-nominal category of predicates. Therefore, they are distinguished from a verbal category of predicates which are used to express actions and processes. Copular verbs are complemented by a subject complement which is represented by a noun or an adjective. Subject complement is the carrier of the actual content, thus copular verbs remain semantically empty.

The theoretical part focuses on the description of copular verbs from several theoretical points of view. It is divided into chapters covering differences in terminology, the division of copular verbs based on approaches of several authors, semantic categories of copular predication and types of complements copular verbs take. Also, the theoretical part provides a brief overview of the corpus-based approach to copular verbs which later, in the analysis, serves as an inspiration and a point of reference for our comparison.

In the analytical part, seven copular verbs are analysed in two corpora: *feel*, *look*, *seem*, *remain*, *become*, *go* and *get*. Araneum Anglicum Minus serves as the source of the up-to-date written language data, while Spoken BNC2014 serves as the source of contemporary spoken language data. The verbs are searched in both corpora with the focus on their adjectival complements. The results are then compared between the two corpora, i.e. between the written and spoken register, based on the frequency of co-occurrence and selected association measures. The analysis aims to reveal similarities and differences in the use of copular verbs complemented by adjectives in the two registers of contemporary English. The corpus-based analysis also explores the occurrence of adverbs which modify the adjectives and serves as a basis for an attempted semantic categorization of adjectives.

2 Theoretical background

There are two elements that are essential to constitute an English sentence: a subject and a predicate. While the subject can be represented by several means, the predicate is usually represented by a finite verb form or less frequently by an adverb or an interjection. Verbs, which function as predicates in a sentence, can be divided into two categories: a verbal category and a verbo-nominal category (*verbonominální predikace*). While the former category is represented by lexical verbs which carry the meaning on their own, the latter category is represented by copular verbs which are lexically empty and function solely as a linking device. On that account, meaning is expressed by the element following the verb, subject complement, and together they create a predicate phrase. In other words, a copular verb (or a copula) is a verb which connects a subject to the subject complement. (Dušková et al., 2012: 405). Whether a verb takes an object or a complement is specified by the verb's valency (Downing and Locke, 2006: 36). Copular verbs are also called *linking verbs* (Quirk et al., 1985) or *complex-intransitives* (Huddleston and Pullum, 2002).

In terms of concord, the grammatical number conveyed by the copula is in agreement with the subject number not with the number expressed by the complement. In contrast to the English language, in the Czech language the category of number is ruled by the complement. Lexically, the verbs are empty and their primary function is to ascribe a property to its subject. In the sentence, the actual content, i.e. the property, is carried by the complement. (Dušková et al. 2012: 402). The topic of copular verbs is treated differently in various grammars of the English language and the verbs are described there in varying detail.

This chapter aims at summarizing the information provided by well-known grammars and to offer an insight into the topic. Although the practical part of the present thesis explores only a selection of copular verbs it is necessary to also describe the verb *be* in the theoretical part because, as one of the chapters reveals, *be* constitutes a principal copula which serves as a model copular verb described frequently in grammar books. As Quirk et al. mention, the verb *be* is “overwhelmingly the most common” (Quirk et al., 1985: 1774). Phenomena generally concerning copular verbs are thus shown on this specific verb as they are applicable to the other copular verbs too.

2.1 Copular verbs in theory

2.1.1 Definition and terminology

As was already mentioned, copular verbs constitute the verbo-nominal category of predicates, i.e. verbs which express relations and attribute properties. By contrast, verbal predicates express actions and processes. While copular verbs carry information about their grammatical categories, i.e. person, number, tense and mood, the element which follows them, subject complement, is the carrier of the properties which are ascribed to the subject. Pustet (2005: 5) accentuates three syntactic functions of copulas: 1) a linker between a subject and a complement, 2) a syntactic hitching post to which verbal inflectional categories can be attached, and 3) a predicator which is added to lexemes that do not form predicates on their own.

In Czech grammars, the element following copulas is called *jmenná část přísudku* (*subject complement*) and is distinguished from *doplněk podmětu* (*verbless adverbial clause*). Dušková et al. (2012: 505) point out that the subject complement resembles the verbless adverbial clause in the postverbal position and it is also represented by a noun or an adjective. Often, the two elements are not differentiated, therefore, the postverbal element in the following example (*a begger*) can be identified as one or another. *He died a begger*. Dušková et al. suggest that the element is an instance of verbless adverbial clause because it follows a lexical verb which can stand on its own: *he died*.

On the other hand, the subject complement follows copular verbs which cannot stand alone as in: **it seems*; therefore, it constitutes an obligatory element of a sentence. In short, the verbless adverbial clause is a facultative element and the subject complement is an obligatory element.

Biber et al. (1999: 435) define the element as *a subject predicative*. Downing and Locke (2006: 63) characterize the element as *a subject complement* and explain that “the subject complement completes the predicate after a copular verb by specifying an attribute of the subject or its identity.” Huddleston and Pullum (2002: 217) call the element *a predicative complement* and distinguish between two types which are classified either as a depictive type or as a resultative type. While the latter occurs with verbs which express a change of a state

and the complement express the state at the end of the process the former type expresses the property of the subject without any such factor of change. Moreover, these constructions are defined as *complex-intransitive* and the subject is called *predicand* of the complement. Only constructions with the copula *be* are called *copular clauses* because it is a mere syntactic copula which is not true about the other verbs.

For the needs of the present thesis the element following the verb is called by its general term, i.e. the subject complement and the term copular verbs is applied on all verbs requiring subject-oriented complements.

2.1.2 The subject complement

The complement is usually represented by a noun or an adjective. The noun which follows the copula does not introduce a new referent as the referent remains the same, i.e. the subject and the complement denote the same referent. For example: *His mother is a widow*. The adjective in the function of the complement is specific for copulas and it denotes the property attributed to the subject, e.g. *His mother was very indulgent* (Dušková et al., 2012: 206).

On the contrary, Quirk et al. (1985: 1174) claim that the verbs remain copular even if the complement is represented by an adverbial. The complementing adverbials are mainly of the type space and time. They appear specifically with the verbs *be*, *keep* and *get*. These are considered copular because they cannot occur without the adjunct which makes it an obligatory element of the clause, e.g. *At last we got home.* - **At last we got.* Biber et al. (1999: 428) agree with Quirk et al. on using adverbials as complements of copular verbs. To be exact, they state that the copular verb *be* links the subject noun phrase with a subject complement or obligatory adverbial. “Many copular verbs can also be used to locate the subject of the clause in relation to time or place.” Therefore, *be* in the following example is also an instance of a copular verb. *Well that's how we got acquainted so well because she was in Olie's room a lot.* They express that if the verb is complemented by an obligatory adverbial it is usually a prepositional phrase. Huddleston and Pullum (2002: 257) propose that a locative expression following a copular verb is similar to the subject complement based on the fact that in both something is assigned to the subject, either a property or a location. The following examples demonstrate their claim *the letter is on the table* and *the letter is highly offensive*. While in the first example a location is assigned in the second example it is a

property. Also, both are oriented towards another element in the structure which is either the subject in complex-intransitive (copular) constructions or the object in complex-transitive constructions. Despite their semantic resemblance, the locative expressions still remain syntactically distinct hence they cannot be ranked among the subject complements and the verbs among the copulas.

Some verbs take both objects and complements. Huddleston and Pullum (2002: 253) provide three syntactic differences between a subject complement and an object. The first difference between these two elements concerns the admissibility of adjective and bare role noun phrases as complements and not objects. For example, it is possible to say *he seemed nice* or *she remained treasurer* but not **he met nice*. The possibility of having an adjective phrase as a complement and not an object reflects that complements tend to express properties. Similarly, bare noun phrases are excluded from the subject position because they are generally non-referential. Secondly, subject complements never correspond to subjects in related passive constructions. This is a property of objects which become subjects in passive voice. Thirdly, the case of a subject should correspond to the case of a subject complement. The third rule is mainly relevant in other languages with a richer case system which English as an analytical language does not have.

2.1.3 The principal copular verb *be*

Referring to my previous argument, it is necessary to define the verb *be* as it is perceived as the main representative of copular verbs. *Be* is a verb with several meanings and functions other than copular. The element which follows the verb indicates what function the verb has. Dušková et al. (2012: 175) distinguish three functions of the verb *be*, i.e. a lexical, copular and modal function. *Be* as a lexical verb carries the meanings of *to locate*, *to exist* and *to take place*. *Be* as a copular verb serves as a link between a subject and a complement which is represented by a noun or an adjective. *Be* followed by an infinitive functions as a modal verb. If the element is represented by a prepositional phrase the verb is lexical and the predication is locative, as in: *Will you be at home?* Similarly, if the verb is followed by an adverb either expressing a location or time the predication is again locative. As a consequence, the verb cannot be classified as copular. However, the verb can be copular even if it is followed by a prepositional phrase, e.g. *He was on his guard*. It is necessary to point out that in such cases it

is possible to paraphrase the verb + prepositional phrase using adjective, e.g. *He was watchful*. Moreover, verbs in sentences in which the subject is represented by *there* constitute existential or existential-locative predication subsumed under the semantic subtypes of verbo-nominal predication. The element which follows the verb is called a notional subject. Dušková et al. draw attention to the fact that in such sentences *be* alternates with *exist*, hence it can be perceived as a lexical verb too, e.g. *Contrary to popular belief, there are many more than five senses*. Quirk et al. (1985: 1403) also support the idea that the constructions with *there* as a subject are existential and the verb is lexical.

Huddleston and Pullum (2002: 113), on the other hand, distinguish six functions of the verb *be* and copular is one of them. The other functions are lexical, quasi-modal, motional, progressive, passive.

2.1.4 Types of copular verbs

Copular verbs can be divided into several different types, depending on the approach. In this section, we briefly summarize several important distinctions according to the selected grammar books and studies.

Dušková et al. (2012: 207) classifies copular verbs into two types: 1) “a type *be*” and 2) “a type *become*”. The verbs corresponding to the type *be* represent the meaning of “*to be someone, something*” and the verbs corresponding to the type *become* represent the meaning of “*to become someone, something*”. While the type *be* expresses a continuing state of existence, the type *become* expresses a change of a state. The other copulas can also be classified according to the two types.

The first type is further subdivided into 3 categories of verbs:

1a) The first category expresses the relation between the subject and the complement through senses. It ascribes the property to the subject but it is mediated by a verb of sensory perception. In other words, the factuality of utterances is relativized by the sensory verbs (see ex. 1). These verbs can be used as full lexical verbs too. The verbs are, for instance, *look, feel, smell, taste* or *sound*.

1. *It sounds marvellous.*

1b) The second category refers also to the factuality of the utterance. The use of these verbs indicates a degree of factuality as means of epistemic modality (see ex. 2). The instances are *seem*, *appear*, *prove*, *turn out*. *Seem* and *appear* mitigate the level of factuality. On the contrary, *prove* and *turn out* confirm the factuality.

2. *Our effort seemed useless.*

1c) Finally, the third category expresses the duration of a property which remains unchanged. The following examples of verbs explicitly state that the property ascribed to the subject is still in duration (see ex. 3): *remain*, *stay*, *keep*, *continue*, *stand*, *rest*.

3. *He remained calm.*

2) The second type of copulas corresponding to the verb *become* is not further divided and includes verbs such as: *turn*, *grow*, *get*, *go*, *come*, *fall*, *make*.

4. *Her face turned purple.*

Dušková (2012: 408) also add that the copular predication can be divided according to which semantic subtype of predication the verbs and their complements express. The subtypes are qualifying, identifying, classifying, possessive and existential or existential-locative.

There is also a category of verbs of a type *have*, i.e. verbs such as *have*, *make*, *take*, *get*, which are sometimes considered copular too. They are distinct because they take an object, but the object is represented by a noun denoting action which is derived from a verb; therefore, it expresses the actual action in question (see p. 22).

However, Huddleston and Pullum (2002: 111) state that *have* in these constructions is a lexical dynamic verb and not copular. Additionally, they describe the verb *have* and other similar verbs mentioned above as *light verbs* which are *semantically light*, i.e. the content is carried by the nouns following these verbs. The noun in these constructions functions as head of the direct object, e.g. *We took a rest*.

The above mentioned verbs in these constructions are put in contrast with their ordinary use in which they convey full semantic content, e.g. *We took all we could find*. Such constructions allow for a greater range of modifications of the noun. As the following example shows, the use of a light verb with a noun is less awkward than the use of adverbial complementation. *She gave him an unusually passionate kiss.- She kissed him unusually passionately*. Also, there are often no adverbial counterparts, e.g. *We took a well-earned rest*. The two

constructions in contrast may often have a different meaning. If the following two examples are compared, it is understood that in the first example the person drank all of the milk, whereas in the second example with *have* the person possibly had just a sip. *He drank my milk.- He had a drink of my milk* (Huddleston and Pullum, 2002: 290).

Biber et al. (1999: 430) explain that, for example, *have a look* simply is an idiomatic multi-word phrase which occurs in conversation.

Quirk et al. do not mention the verb *have* in the category of copular verbs. In addition to the copular verbs already mentioned, Quirk et al. also list other verbs, which “have this function with severe restrictions on the words occurring in the complement.” Possible restrictions may be lexical, i.e. the verb occurs in already existing idiomatic sequences (e.g. *rest assured*), or semantic, i.e. the verb is restricted to a subset of words because of its semantics (e.g. *blush bright red*). These verbs are generally lexical verbs which function here almost as intransitive verbs and the complements are only optional.

Concerning the category of special verbs functioning as copular verbs, Malá (2014: 45) suggests that “there is hardly a clearcut boundary between copular and full lexical verbs” and proposes that there rather is a gradient which has the copular category of verbs in the centre and the verbs may depart from the centre. The directions in which the verbs may depart are described on the basis of semantic, syntactic and formal grounds. It leads us back to the general description of copular verbs, i.e. verbs which are semantically weak, intransitive with an obligatory subject-oriented complement typically expressed by an adjective phrase.

Quirk et al. (1985: 1171) distinguish two types of copular verbs: “current” and “resulting”. Accordingly, the complement has the role of current attribute or resulting attribute. Current copulas express a current state and staying in the state, while resulting copulas, on the other hand, express a change of a state. Based on the limited semantic content of the copular verbs, Quirk et al. (1985: 1174) define the copula *be* as the most neutral in meaning. This frequently occurring copula appears frequently in stative contexts. However, it is possible that it refers to activities and events too, e.g. *You’re being very helpful*. Also, the verb often resembles *become* as in *Cora was angry when she heard about the accident*.

From their point of view, other current copulas, i.e. apart from the verb *be*, are divided into the following categories: “verbs of seeming” (including verbs of sensory perception and verbs such as *seem*, *appear*) and “verbs of remaining” (e.g. *remain*, *stay*, *keep*). Resulting copulas

are called “verbs of becoming” and the representatives are for example *become*, *get*, *go*, *turn*, *grow*. While Dušková et al. classify the verb *go* as the verb which describes a change of a state, therefore it corresponds to the category of resulting copulas, Quirk et al. add that the verb may be understood as a current copula too, as in: *go hungry/ naked*.

Furthermore, Quirk et al. describe the verb *become* as a process verb which places emphasis on the duration of the change and contrasts it with the verb *get* which in turn places “more emphasis on the agency behind the event or on the result of the change.” *Go* and *turn* both tend to be used in contexts in which the changes happen in spite of human agency and something/ someone worsens as in *go mad*. *Turn*, especially, is used in contexts in which the changes happen naturally as in *turn green*. *Grow*, similar to *turn*, is used with natural changes with the emphasis on gradual development, therefore, it occurs with comparative adjectives, e.g. *grow tall*, *grow cooler*.

Biber et al. (1999: 436) distinguish the same types as Quirk et al., i.e. current and resulting copulas.

Pustet (2005: 5) suggest that the copulas may be differentiated based on the fact that some of them add meaning to the predicative phrases they are contained in. Such verbs are, for example, *become*, *feel*, *look* or *remain* as in *to become big*, *to feel good*, *to look nice*, *to remain intact*. These verbs are called semi-copulas and they resemble copulas because of their inability to create predicates on their own. On the other hand, they also resemble full lexical verbs because of the meaning they add to the predicate phrase. What is more, they cannot be left out without changing or affecting the meaning of the whole construction.

2.1.5 Subject complement and its representation

Even though an adjective is the most diagnostic form of the subject complement, the subject complement can take up various forms. The present chapter aims at summarising the types of complementation copular verbs take which is based on the types of complementation provided by Dušková et al. (2012: 206). They also state that, in terms of the semantic subtype of predication, qualifying and classifying subtypes prevail.

Dušková et al. (2012: 206) suggest that with the verbs of sensory perception, the complement is represented by a noun or an adjective (see ex. 5, 6). While the adjective is very common, the noun in such cases is often introduced by a preposition (*She felt like a different person*). However, Quirk et al. (1985: 1175) note that this use of *feel like* needs to be distinguished from its use as a monotransitive prepositional verb as in *I feel like a cup of coffee*, in which it alternates with the verb *want*. Example 5 is an instance of a qualification by false classification. Dušková et al. (2012: 413) point out that semantically the information which contains the noun in the complement is already included in the subject, therefore it is rather a qualifying predication.

5. *She felt a different person.- She felt different.*

6. *The wine tastes bitter.*

Adverbs and adjectives have sometimes the same form (see ex. 7), thus it is often ambiguous whether the verb should be treated as a copula or as a lexical verb. Example 7 can mean that he looked forceful or that he looked at something very intensely, depending on the context.

7. *He looked hard.*

The behaviour of sense verbs in a sentence is determined by the other elements in the sentence. The verbs *smell* and *taste* are followed by the prepositions *of* and *like* if the complement is represented by a noun, as in example 8, 9. The verbs *feel* and *look* take reflexive pronouns in the function of complements, as in example 10. The verb *feel* takes a progressive form only when the subject is animate (see ex. 11). Example 12 is an instance of mediopassive voice, what is meant is that *cigarettes are delicious when smoked*. Huddleston and Pullum (2002: 263) add that these verbs can take infinitival structures as their complements, however, the truth value of the proposition is stronger when adjectives are used, e.g. *She looked to be happy.- She looked happy*.

8. *It smells like a sherry.*

9. *The market hall smelt of fish.*

10. *You are looking yourself again.*

11. *I am feeling better.*

12. *Filter- tipped cigarettes smoke delicious.*

Copular verbs, such as *seem*, *appear*, *prove*, *turn out*, express speaker's attitude towards the degree of factuality of the proposition. This type of copular verbs can be followed by adjectives, nouns and infinitives. The verb *appear*, similarly as the verb *look*, is ambiguous when used with a word form which functions as an adjective and adverb at the same time (see ex. 13). Sentences with the verbs *seem* and *appear* allow for the person to whom the factuality is relativized to be expressed (see ex. 14). Especially with animate subjects, the verb *prove* appears with a reflexive pronoun, as in example 15. An additional construction *to be* can follow the reflexive pronoun but it is only obligatory with inanimate subjects.

13. *The train appeared fast.*

14. *What he said appeared quite sensible to me.*

15. *She proved herself an economical housewife.*

Verbs expressing duration of a state, another type of copular verbs, usually take adjectives as their complements. The verb *remain* appears with a noun too (see ex. 16). The prepositional phrase in the example overtly expresses the person for which the state is valid. A noun in the position of the complement is also feasible with the verbs *stand*, *keep* and *continue* (see ex. 17). However, with the verb *stand* it is only in fixed expressions such as *to stand godmother*. Prepositional phrases are also possible with this type of verbs, but it is necessary distinguish the uses which constitute verbal category and verbo-nominal category. Example 18 is an instance of the latter as the verb ascribes some quality to the subject.

16. *They remained for me the true heroes of my early life.*

17. *We continued friends.*

18. *He keeps in excellent health.*

In the vast majority of cases, verbs expressing a change of a state, the last type of copular verbs, are complemented by an adjective. Similarly, Malá (2014: 116) asserts that “*get*, *go* and *grow* are virtually limited to this type of complementation.”

Dušková et al. add that a noun is also a possible option, but it is restricted to some verbs only (*become*, *turn*, *fall*, *make*). The verb *come* occurs with past participles of verbs prefixed by –*un* the most (ex. 19), with *to be* (ex. 20) and with a prepositional phrase + noun (ex. 21). Constructions such as *the door blew open* can also be subsumed under this category because we can paraphrase them as *the door became open*.

19. *The knot came untied.*
20. *She came to be his admirer.*
21. *Every driver came to a halt.*

As was already mentioned in the chapter dedicated to the types of copular verbs (see p. 18), special type of predication occurs when a verb is complemented by an object. Formally, these verbs are divided into two types based on what relation between the two elements, i.e. the verb and the complement, is. The relation is based either on conversion (ex. 22) or derivation (ex. 23).

22. *I had another try.*
23. *She made a decision.*

Have is one representative of such verbs. Its primary meaning, i.e. *to possess*, is weakened and it is possible to replace *have* together with the noun complementing it with the verb from which the noun was derived (see ex. 24, 25). One advantage of it is that it is possible to modify the noun as in example 26. In the verbal predication it is not possible to modify the verb by an adverb (*I looked at him well*) or it has a different meaning (ex. 27). Nevertheless, the possessive meaning of the verb *have* is closely related to the meaning of the copula *be* (existential). When we compare examples 28 and 29 it is possible to observe that there is a change in the perspective. Example 28 describes the possessive relationship from the possessor's point of view and example 29 from the perspective of the thing that is possessed. It is significant for American English that *take* is used in place of *have*.

24. *We had a smoke.- We smoked.*
25. *We had a chat.- We chatted.*
26. *I had a good look at him.*
27. *She had a nice swim.- She swims nicely.*
28. *John has a new car.*
29. *The new car is John's.*

Similarly, the verbs *do* and *make* function as copular verbs. The preference of verbo-nominal predication over verbal is a sign of an academic style of writing which is known to be considerably nominal. Therefore, expressions such as *to do damage*, *to make a discovery* or *to make a conclusion* are found in academic texts more than in spoken discourse. Other

examples of verbs which are normally found in verbal contexts but perform here the role of copular verbs are:

30. *She gave a sudden shudder.*

31. *He took a bath.*

32. *We'll pay you a visit in summer.*

In conclusion, apart from the traditional copular verbs, a large variety of other verbs constitutes the verbo-nominal predication. A verb comprised in a verbonominal category (*to make a start*) can be transferred so it forms a verbal category (*to start*) if this transfer is not practicable the verb is not copular. This phenomenon is used when a speaker wishes to change the FSP structure as the complement can be modified (*He gave me a look which startled me.- He looked at me sternly.*) or if a verb cannot be used on its own because it requires an object (*I have made arrangements.- *I have arranged -*).

2.2 A corpus-based approach to copular verbs

In contrast with theoretical concepts, a corpus-based approach is based strictly on empirical corpus data, i.e. authentic language material. Corpus-based studies often aim to provide solid evidence for previous findings and to discover new regularities and tendencies that would otherwise remain hidden without the large amounts of data. For the purposes of this thesis, two corpus-based studies are especially relevant as a source of inspiration and a point of reference: Biber and Conrad's Real Grammar from 2009 and Malá's study from 2014.

2.2.1 Biber and Conrad 2009

Biber and Conrad (2009: 22, 23) use the latest corpus research and present a different way of introducing grammar to students, based on findings from real contexts. They observed the constructions with copular verbs in both writing and speech. They provide an analysis of seven verbs and assert that *feel*, *get*, *go* and *look* are more frequent in conversation, while *become*, *remain* and *seem* are more frequent in writing. For each verb there is a specific description of its meaning and a list of adjectives occurring with the verb with examples attached.

The verb *feel* describes physical sensations and mental sensations. In the former case it is commonly followed by the following set of adjectives: *better, good, tired, cold, sick, uncomfortable*. In the latter case it is followed by a different set of adjectives: *guilty, ashamed, uneasy, bad, sure*.

The verb *look* is defined as a verb describing positive and negative feelings about physical appearance. The adjectives *good, lovely, happy, nice* express the former feelings and the adjectives *awful, sad, terrible, pale, small, tired* express the latter.

The verb *seem* is said to express likelihood with the adjectives *clear, obvious, reasonable, likely, possible* and *unlikely*.

The verb *remain* expresses an absence of a change and its set of adjectives is consists of *closed, intact, unchanged, constant, uncertain* and *unknown*.

The verb *become* describes a change in understanding or importance and the adjectives are: *apparent, difficult, familiar, clear, evident* and *important*.

The verb *go* also describes a change to a negative state but with a different set of adjectives: *bad, deaf, mad, crazy, limp* and *wrong*.

The verb *get* expresses a change to a negative state with the set of following adjectives: *angry, mad, upset, bored, sick, wet, cold, tired* and *lost*.

2.2.2 Malá 2014

Malá (2014) uses the British National Corpus in her contrastive study of copular verbs for which she chooses nineteen copular verbs in order to show that the copular verbs differ not only in their meaning, frequency and distribution but also in the types of complements they allow for. Among the selected nineteen copular verbs there are also the ones presented by Biber and Conrad (2009) in their corpus-based study, i.e. *feel, look, seem, remain, become, go* and *get*. Malá (2014: 52-78) presents the copular verbs in great detail together with their frequency distribution across registers. The figures show how the frequency of verbs differs depending on the register. There are the most frequent adjectives listed for each verb, differentiated by the register.

The two sensory copulas from the group mentioned above, i.e. *feel* and *look*, are most frequent of the group of sensory verbs in the construction verb + adjective included in the study. *Feel* is the most common copula across all registers and it specifically outnumbers the other verbs of sensory perception in fiction. As for the frequent adjectives occurring with this verb in fiction, *sorry, guilty, sick, good, sure, tired, safe, uncomfortable, ill* and *strange* are listed. On the other hand, *look* is the most frequent in conversation and in newspapers. Among the most frequent adjectives, *nice, good, alright, lovely, different, right, bad, well, smart* and *big* are listed for conversation and *good, likely, set, certain, new, unlikely, dangerous, bleak, capable* and *great* for newspapers.

Seem is one of the verbs which in their meaning indicate a degree of factuality of what is being said. Malá points out that, according to her findings, *seem* is the second most frequent copula generally after the verb *be*. *Seem* in the combination with the adjectival complement is especially very common in fiction with adjectives *unlikely, good, likely, pleased, surprised, happy, interested, right, strange* and *able* and in academic prose with adjectives *likely, unlikely, reasonable, appropriate, clear, possible, probable, obvious, certain* and *plausible*.

The verb *remain* followed by an adjective is the most frequent representative of verbs of remaining across all registers. It significantly outnumbers the other verbs in academic prose and in newspapers. Malá points out that the verb appears with static adjectival complements and that it describes the continuation of a state. The most frequent adjectives occurring with *remain* are *unchanged, constant, silent, intact, unclear, strong, true, important, loyal* and *stable*.

The resulting copulas *become, get, and go* describe the result of a change of a state. When followed by an adjectival complement, the verbs *become* and *get* are the most frequent ones. In academic prose *become* is more frequent than the other verbs, on the other hand, *get* is more frequent in conversation. The most frequent adjectives occurring with the verb *become* are *clear, aware, apparent, involved, available, important, difficult, popular, common* and *familiar*. The verb *get* has a different set of frequently appearing adjectives *used, better, worse, involved, available, important, difficult, popular, common* and *familiar*. While she mentions that the verb *become* is often accompanied by an adjective in comparative or an adjective modified by a degree adverbial, the verb prefers nominal subject complements describing positions and functions performed by people. *Go* is generally less frequent in all registers than the other two verbs, nonetheless, it occurs frequently in conversation and

fiction. The adjectives accompanying the verb most frequently are different from the adjectives listed above for the other two resulting copulas. The adjectives are *wrong*, *mad*, *bust*, *bankrupt*, *crazy*, *wild*, *cold*, *berserk*, *white* and *free*.

3 Material and Method

3.1 Material

3.1.1 Corpora

To provide an analysis of contemporary use of copular verbs in the English language, the latest available corpora have been chosen for both registers- written and spoken. Spoken BNC2014 corpus has been selected as the source material of the spoken language data and Araneum Anglicum Minus corpus as the source material of the up-to-date written language data from the year 2013. Even though Written BNC2014 would make a better counterpart to its spoken version, this corpus is not yet available and the older version from 1994 would not provide data about the current use of copular verbs. Although Araneum Anglicum Minus is a webcorpus, which means that the available genres differ greatly from a traditional written corpus, it is the only evidence of a contemporary written English available for us at the moment.

The spoken component of the British National Corpus Spoken BNC2014 with more than 11.5 million words consists of transcripts of recorded spontaneous conversations. The conversations were collected from 672 people who recorded themselves using built-in recorders in their smartphones in informal settings in the UK between 2012 and 2016. The corpus can be accessed, after creating a free account and a following registration, via Lancaster University's CQPweb server¹.

As it was already said, the written language data have been excerpted from the Aranea Corpus Family, namely from the Araneum Anglicum Minus corpus which is a smaller version of the bigger corpus called Maius. It is a web corpus with a total of around 100 million words. The data were collected during the year 2013 via SpiderLing which is, as Benko (2014: 248) explains, “a specialized crawler for downloading textual data from the web.” As a result, users are provided with various texts from blogs, official websites and more. The Araneum

¹ The Spoken BNC2014 together with CQPweb server accessible at WWW: <https://cqpweb.lancs.ac.uk/>

Anglicum corpus can easily be accessed by KonText, which is a query interface for working with corpora².

Being a typical web corpus, the Araneum Anglicum Minus also does not correspond in size to the Spoken BNC2014 (it is approximately ten times larger), which is another issue in comparability we had to bear in mind in the analysis. However, as web corpora do not usually provide researchers with any linguistically relevant metadata that could be used to create a smaller comparable subcorpus, we decided to use the whole corpus as a source of data and apply such methodology that makes use of normalized data (such as relative frequency or collocation ranking based on comparable association measures, see 3.2).

For the needs of the analysis (see Chapter 4), the names of the two corpora are shortened to SBNC and AAM, respectively.

3.1.2 Selected copular verbs

Referring to the previous corpus-based studies (see 2.2.), the following seven verbs have been chosen for the analysis *feel, look, seem, remain, become, go* and *get*. These verbs are common representatives of copular verbs (yet not too common, such as the verb *be*, to prevent us from conducting the analysis within the scope of a thesis) and can be divided into two groups (cf. Dušková et al. described in 2.1.4). Four of the verbs correspond to the type *be*, i.e. *feel, look, seem, remain*, and the remaining three correspond to the type *become*, i.e. *become, go, get*. Therefore, there are representatives of both types of copular verbs presented by Dušková et al., i.e. copular verbs describing a continuing state of existence and copular verbs describing a change of a state. Also, the choice of these verbs is in full agreement with the set of analysed verbs by Biber and Conrad and Malá (see chapter 2.2).

3.2 Method

To obtain relevant corpus data for analysis, the adopted procedure had several steps, described in detail below.

² The Aranea Corpus together with KonText application accessible at WWW: <https://www.korpus.cz/>

1) The selected verbs were searched one by one in both corpora as lemmas together with the tag for adjectives immediately following them. An example of the CQL query for both corpora is provided below.

KonText- Araneum Anglicum Minus
<code>[lemma="feel"][tag="JJ"]</code>
CQPweb- Spoken BNC2014
<code>[lemma="feel"][pos="JJ"]</code>

2) Frequency lists of the copular verbs with their adjectival complements were obtained using appropriate functions in KonText and CQPWeb. The frequency information can be considered a useful first indicator of common recurring patterns in language, applicable e.g. in language learning at lower levels of proficiency.

3) In order to be able to also assess the collocational strength of the combinations, and not just their frequency (useful for more experienced users of language interested in prominent, albeit not the most frequent combinations), association measures were used to obtain comparable lists of copular verbs and their complements, based on the probability of co-occurrence.

- Specifically, the logDice association measure was used in KonText indicating which of the adjectives are more significant and which are rather coincidental.
- In CQPweb, a similar association measure called Dice coefficient was used as it is the closest available equivalent of the logDice measure, and the calculation system is roughly the same so the rankings of both lists can later be compared.

4) In both corpora, after the lists had been obtained, the results were scanned to validate whether they were truly the cases of a copular verb with an adjective or not. Very often, the adjectives constituted a part of a larger phrase in which it modified a noun. It specifically concerned the verb *get* as it has many different functions. It appears as a lexical verb with an object modified by the adjective or in passive constructions in which it is necessary to distinguish whether the element following the verb is a past participle or a participial adjective. Instances of past participles were excluded from the analysis.

To test whether the verb is used as a copular verb and not as an auxiliary, four rules can be used. These rules are defined by Quirk et al. (1985: 168) and are:

- the adjective can be modified by *quite, rather, very*, etc.;
- the possibility of replacing *get* by another copular verb (*look, seem, feel*, etc.);
- the adjective has a stative meaning;
- a rare possibility of adding a by-phrase expressing the agent of the passive clause.

5) After the results were filtered and only cases of copular predication were taken into consideration,

- a. the frequency list and list of collocation candidates were compared within the same corpus;
- b. the lists of collocation candidates were compared between the two corpora, i.e. the corpus of written language data and the corpus of spoken language data;
- c. the adjectives were semantically categorized where possible.

6) As a sub-objective, we also explored the use of adverbs which could modify the adjectives occurring with the copular verbs, specifically focusing on whether there tend to be more adverbs before one adjective (especially in the spoken register).

To obtain the data with adverbs the CQL queries had to be altered in the following way:

KonText- Araneum Anglicum Minus

[lemma="feel"][tag="RB"]{2,4}[tag="JJ"]

CQPweb- Spoken BNC2014

[lemma="feel"][pos="R. "]{2,4}[pos="JJ"]

4 Analysis

The analytical part of the thesis describes in detail the seven verbs mentioned above, i.e. *feel*, *look*, *seem*, *remain*, *become*, *go* and *get*, and their adjectival complements.

For each verb there are four tables, i.e. two tables of data excerpted from each corpus. The tables show the first fifteen adjectives the verb occurs the most frequently with in the given corpus and the first fifteen adjectives the verb collocates the most with. As was already mentioned, logDice and Dice coefficient measures were used to calculate the collocability of the verb. Occasionally, the verb does not have sufficient number of adjectives to include in the table, especially due to the limited number of material in the spoken language data, thus the table represents the number of adjectives possible. Tables in the present analysis are labelled whether they refer to the corpus of the written language data, i.e. the AAM corpus, or to the corpus of the spoken language data, i.e. the SBNC corpus.

It is divided into chapters according to the individual verbs. Each chapter is further divided into four subchapters.

- The first subchapter focuses on the frequency, i.e. which adjectives occur with the verbs repeatedly, and the collocational strength between the copula and the complement are observed. For each verb, there are always four tables, i.e. two tables of data excerpted from each corpus. The tables show the first fifteen adjectives the verb occurs the most frequently with in the given corpus and the first fifteen adjectives the verb collocates the most with. . Occasionally, the verb does not have sufficient number of adjectives to include in the table, especially due to the limited number of data in the spoken language data, thus the table represents the number of adjectives possible. The frequency lists are contrasted with the lists of collocation candidates of the same verb based on the ranking of the adjectives. It is expected that the two lists might differ.
- In the second subchapter, the results from the two corpora are contrasted highlighting the differences between the spoken language data and written language data.
- In the third subchapter, the adjectives are semantically categorized based on the lists of collocates, and the categories are compared with the sets of adjectives provided by Biber and Conrad and Malá (see 2.2). The analysis attempts to either propose different

sets of adjectives or affirm the already existing ones based on the presented research in the selected corpora.

- Ultimately, the final subchapter covers the use of adverbs which modify the adjectives occurring with the copular verbs with the tendency of the adverbs to appear several times with a particular adjective.

Not all adjectives that appeared in first positions in the initial lists the corpus shown were included in the analysis. One of the reasons was that the verbs could not be perceived as copular in these constructions. Several instances of the copular verb + adjective must have been excluded because the verb functioned as a transitive verb and the adjective only constituted a part of a larger phrase, e.g. *get amazing results*.

It needs to be pointed out that some of the adverbs were excluded from the analysis too. For example *as* which surely is an adverb in terms of parts of speech, however, it has a different function. It occurs in structures *as... as*, *quite as... as* or *just as... as*. The last two constructions mentioned also cannot be perceived as two adverbs which both equally modify the adjective. Such constructions are used to make comparisons and to say that something is equal in some property, e.g. *So clothes feel just as good as they look*. Also, the corpus of the spoken register reveals that the adverb *like* appears frequently in combinations with other adverbs with all verbs presented in the analysis. However, the classification of *like* as an adverb in such constructions is tricky. When we look at it closely we may see that in function it does not resemble the other adverbs in the construction. It may be concluded that it is used frequently in speech and it functions rather as a filling which is used when we make pauses in our speech, or we can use it as a way of approximation when we do not want to say the exact information. The corpus shows several examples in which *like* has this function. It appears before another adverb, e.g. *Cos I'm tired yeah I just feel like really lousy*, or after, e.g. *[...]I felt really like emotional cos I was like oh it reminds of being young*.

Very often, the adverbs the corpus returned are simply not modifiers of the adjectives, as they constitute a different type of adverbs, e.g. adverb of place. Therefore, the verbs themselves cannot be perceived as copulas. Malá (2014: 40) also points out in her study that some complex intransitive verbs take both types of complements, i.e. an adverbial complement (*remain outside*) and a subject complement (*remain silent*). In the present analysis, it specifically concerns the verbs *go* and *get*, as several instances of adverbs occurring with these verbs must have been excluded.

4.1 The verb *feel*

If we apply the division of types of verbs by Dušková et al. (2014), the verb *feel* ranks among the verbs of type *be*, namely the verbs of sensory perception. As Biber and Conrad (2009) point out, it specifically describes physical and mental sensations.

	Adjective	No.of occurrence	Instances per million
1	<i>good</i>	<i>1017</i>	<i>8,521543</i>
2	<i>comfortable</i>	<i>511</i>	<i>4,281719</i>
3	<i>bad</i>	<i>347</i>	<i>2,907547</i>
4	<i>guilty</i>	<i>305</i>	<i>2,555625</i>
5	<i>safe</i>	<i>299</i>	<i>2,50535</i>
6	<i>great</i>	<i>297</i>	<i>2,488592</i>
7	<i>confident</i>	<i>237</i>	<i>1,985846</i>
8	<i>uncomfortable</i>	<i>165</i>	<i>1,382551</i>
9	<i>happy</i>	<i>124</i>	<i>1,039008</i>
10	<i>sad</i>	<i>96</i>	<i>0,804393</i>
11	<i>tired</i>	<i>90</i>	<i>0,754119</i>
12	<i>sick</i>	<i>82</i>	<i>0,687086</i>
13	<i>ashamed</i>	<i>73</i>	<i>0,611674</i>
14	<i>secure</i>	<i>72</i>	<i>0,603295</i>
15	<i>proud</i>	<i>71</i>	<i>0,594916</i>
Total	<i>10 506 hits</i>		

Table 1: AAM: a frequency list of *feel*

	Adjective	LogDice measure
1	<i>comfortable</i>	10,08
2	<i>guilty</i>	9,54
3	<i>confident</i>	9,19
4	<i>uncomfortable</i>	8,82
5	<i>safe</i>	8,79
6	<i>bad</i>	8,25
7	<i>tired</i>	7,88
8	<i>sad</i>	7,80
9	<i>ashamed</i>	7,73
10	<i>good</i>	7,59
11	<i>sick</i>	7,54
12	<i>helpless</i>	7,45
13	<i>happy</i>	7,36
14	<i>anxious</i>	7,27
15	<i>lonely</i>	7,24

Table 2: AAM: a collocate list of *feel*

	Adjective	No.of occurrence	Instances per million
1	<i>bad</i>	159	13,82609
2	<i>sorry</i>	113	9,826087
3	<i>good</i>	86	7,478261
4	<i>sick</i>	86	7,478261

5	<i>guilty</i>	55	4,782609
6	<i>comfortable</i>	45	3,913043
7	<i>awful</i>	30	2,608696
8	<i>awkward</i>	22	1,913043
9	<i>weird</i>	22	1,913043
10	<i>safe</i>	21	1,826087
11	<i>terrible</i>	21	1,826087
12	<i>happy</i>	20	1,73913
13	<i>ill</i>	20	1,73913
14	<i>nice</i>	20	1,73913
15	<i>alright</i>	19	1,652174
Total	1 304 matches		

Table 3: SBNC: a frequency list of *feel*

	Adjective	Dice coefficient
1	<i>sick</i>	0,081
2	<i>guilty</i>	0,075
3	<i>comfortable</i>	0,054
4	<i>bad</i>	0,038
5	<i>sorry</i>	0,033
6	<i>awkward</i>	0,024
7	<i>safe</i>	0,023
8	<i>confident</i>	0,023
9	<i>ill</i>	0,022

10	<i>uncomfortable</i>	0,022
11	<i>awful</i>	0,022
12	<i>tired</i>	0,018
13	<i>terrible</i>	0,017
14	<i>embarrassed</i>	0,016
15	<i>dizzy</i>	0,015

Table 4: SBNC: a collocate list of *feel*

4.1.1 Comparing frequency and collocate lists within a single corpus

In both of the lists obtained from the AAM corpus the adjective *free* ranked first. However, this construction needed to be excluded from the analysis because it rather is a fixed expression which is stereotypically used to say that a person should not hesitate or be shy and it works as an invitation to do something.. If used in this way it is followed by a *to*-infinitive. The corpus shows that out of 1678 instances of *feel free* 1620 instances are instances of this phrase which means that only 58 instances are not followed by the infinitive. Examples below show both cases (see ex. 1, 2). *Free* also ranks first in the list of collocates.

1. *Please feel free to contact us to help you with the planning of your event.*
2. *Although there is a strong drive in humans to feel free and independent, [...]*

Table 1 shows the ranking of the adjectives based on frequency and Table 2 shows the ranking of the adjectives based on collocability. The adjectives ranked in top positions differ moderately in the two lists. There are some adjectives which appear in one list and not in the other, namely the adjectives *anxious*, *helpless*, *lonely*, *secure*, *great* and *proud*. The adjectives *anxious*, *helpless* and *lonely* occur among the first fifteen adjectives *feel* collocates the most with but not among the fifteen most frequently occurring adjectives. Conversely, the adjectives *secure*, *great* and *proud* are positioned among the first fifteen adjectives in the frequency list, but not among the collocations. As predicted, the two lists differ: While an adjective may be very frequent, another one may occur at lower frequency, but very often in the combination with the verb.

The two morphologically simple adjectives *good* (ex. 3) and *bad* (ex. 4) are frequently used to describe in a simple way how a person feels. There is a sense of opposition between the adjectives, therefore, we may regard them as antonyms. They appear in high positions, namely *good* ranks second and *bad* ranks fourth in the frequency list. However, in the list of collocates, the adjectives rank seventh and eleventh, respectively, suggesting that despite high frequency, their collocational strength with *feel* is not so prominent. Another pair of antonyms- *comfortable* (ex. 5) and *uncomfortable* (ex. 6), rank roughly similarly in the two lists, i.e. the positive member of the pair ranks higher than the negative member of the pair. The antonyms *happy* (ex. 7) and *sad* (ex. 8) also appear in both lists. In the frequency list, *happy* is ranked higher than *sad*. On the other hand, in the list of collocates, *sad* is ranked higher.

In addition to antonyms, we may observe synonyms among the adjectives included in individual lists. Two near-synonyms which describe that something is protected from harm or danger, *safe* (ex. 9) and *secure* (ex. 10), appear in the frequency list with *safe* being more frequent.

3. *You might not feel good about being a senior in a freshman packed class.*
4. *And he felt bad that he'd messed up the whole scenario [...]*
5. *We want to feel comfortable on our travels and stay at a place where we have all the comforts.*
6. *If you feel uncomfortable, just allow that discomfort to be there.*
7. *If you have success, you'll definitely feel happy.*
8. *I feel sad that I was not wiser about the nutrition of cereals.*
9. *Women need to feel safe to speak out about their experiences [...]*
10. *If she feels secure in their relationship it could be a very good one, indeed.*

Table 3 displays the frequency of adjectives occurring with the verb *feel* and Table 4 displays the collocates of *feel* based on the SBNC. The adjective *free* was excluded from the analysis for the same reason as mentioned at the beginning of this subchapter.

Bad (ex. 11) ranks high in both lists. Similarly, *sick* (ex. 12) ranks first in the collocate list and fourth in the frequency list. However, the two lists differ in following adjectives *good*, *weird*,

happy, nice and *alright* which appear in the frequency list only and *confident, uncomfortable, tired, embarrassed* and *dizzy* which appear in the list of collocates only.

11. *I just feel bad every time your parents come round it looks like I don't do anything.*

12. *[...]I noticed that when I eat rich food I will feel sick for a couple of days afterwards.*

Commonly, the adjectives which occur lower in the rankings in the frequency list, such as *ill* (ex. 13) on rank 14, appear in lower positions (ranked 9) among the collocates, too. Alternatively, it does not appear among the collocates at all, e.g. the adjective *weird* (ex. 14) which ranks tenth in terms of frequency.

13. *You don't feel ill, you just get told you've got cancer.*

14. *I feel weird too I was like that's okay we can feel weird together.*

4.1.2 Comparing collocate lists across the corpora

As was already mentioned, the comparison of the two registers, i.e. written and spoken, is based on the lists of collocates of the two corresponding corpora which are displayed at the beginning of chapter 4.1. (Table 2, 4).

The two collocate lists differ significantly with almost a half of the adjectives being different in both lists. The adjectives *sad, ashamed, good, helpless, happy, anxious* and *lonely* appear only in the written register, while the adjectives *sorry, awkward, ill, awful, terrible, embarrassed* and *dizzy* appear only in the spoken register.

Generally, both of the lists represent adjectives with negative connotation. However, very often it is difficult to state whether the connotation is negative, positive or rather neutral. If we consider *confident* to be a positive property, then the list of collocates of the AAM corpus includes five adjectives with positive connotation: *comfortable, confident, safe, good* and *happy*. The collocate list obtained from the SBNC shows even fewer positive adjectives: *comfortable, safe* and *confident*. The rest of the adjectives carry negative connotations.

Especially in the spoken register, the verb *feel* tends to co-occur with adjectives related to health condition, e.g. *sick*, *ill*, *dizzy* and to some extent even *tired*. What is more, in the same register, *sick* ranks first in the overall ranking (ex. 15). Conversely, in the written register, *feel* collocates the most with *comfortable* which suggests a positive state (ex. 16).

15. *It's my worst kind of hangover when I feel sick.*

16. *If I make those jokes, it means that I feel comfortable being myself around you.*

Guilty and *ashamed* may be used in some contexts as synonyms. In the written register both adjectives *guilty* (ex. 17) and *ashamed* (ex. 18) occur frequently with *guilty* being ranked higher. In the spoken register, we find only *guilty* (ex. 19). In both lists *guilty* ranks second, therefore, the verb *feel* strongly collocates with this adjective in both registers.

17. *I still feel guilty saying no to people [...]*

18. *I felt ashamed of how I don't pray for all of them.*

19. *Ah don't feel guilty I'm sure if he was in the same situation he'd do the same thing.*

4.1.3 Semantic categories of the adjectives

Generally, the adjectives occurring with *feel* may be categorized as follows

- a. a negative state connected with health and physical condition,
- b. a negative mental state,
- c. a sense of a pleasant feeling.

In the written register, *uncomfortable*, *tired* and *sick* belong to the first category. The second category comprises the adjectives *guilty*, *bad*, *sad*, *ashamed*, *helpless*, *anxious* and *lonely*. The third category comprises the positive adjectives of the list, i.e. *comfortable*, *confident*, *safe*, *good* and *happy*. Clearly, the second category, i.e. the category which describes a negative mental state, is the largest. Obviously, the categories are not clear-cut, and one adjective may belong to more categories depending on the context. For example, the adjective *comfortable* may be connected with a sense of physical comfort as well as mental comfort. However, it can be noted that the adjectives are mostly negative in terms of connotation.

As shown in the theoretical part, Biber and Conrad (2009) present two sets of adjectives, in which they also distinguish between physical sensations and mental sensations. In the former case *feel* is commonly followed by: *better, good, tired, cold, sick* and *uncomfortable*. In the latter case it is followed by: *guilty, ashamed, uneasy, bad, sure*. Not only are they constituted by mostly negative adjectives but also the adjective *uncomfortable* is categorized as a physical sensation.

In the spoken register, the adjectives constitute the same categories as in the written register even though the adjectives differ to some extent. The first category, i.e. the category of adjectives relating to physical and health conditions, includes adjectives *sick, uncomfortable, ill, tired* and *dizzy*. This category is, in opposition to the written register, enlarged by the adjectives *ill* and *dizzy*. The second category is constituted by *guilty, bad, sorry, awkward, awful, terrible* and *embarrassed*. The fact that this register does not include many positively oriented adjectives is also reflected in the third category comprising only the adjectives *comfortable, safe* and *confident*.

Malá (2014) asserts that *feel* is the most common copula across all registers and it specifically outnumbers the other verbs of sensory perception in fiction. As for the frequent adjectives occurring with this verb in fiction, *sorry, guilty, sick, good, sure, tired, safe, uncomfortable, ill* and *strange* are listed. In her findings we can also see that the adjectives are mostly negative and apart from the adjectives *sure* and *strange*, which do not appear in either of the lists of collocates (probably due to different genres included in the corpora), the adjectives roughly correspond.

4.1.4 Adverbs occurring in the structure

Firstly, single adverbs occurring with the verb were explored. In the written register, the verb *feel* co-occurs and the adjectives are modified frequently by the following set of adverbs: *so, very, really, pretty* and *too*. These adverbs comprise more than a half of the hits the corpus returned (2 198 in total). Examples below (20-24) illustrate their uses. Most of the adverbs are used for emphasis to express a high degree of the property, e.g. *very*, or a moderately high degree, e.g. *pretty*. The most frequent one is the adverb *so*.

20. *Why am I still feeling so tired?*

21. *I felt very happy about how the trip had gone, [...]*
22. *I felt really sad but realized that there was nothing I could do.*
23. *My co-workers are so smart, I usually feel pretty stupid.*
24. *But I feel too weak to speak any further.*

In the spoken register, the corpus returned 661 matches in which only one adverb, *really*, covers almost all instances in which one adverb modifying the adjective complementing the verb *feel* occurs. In the written register, the same adverb ranks third. The other adverbs which can be found frequently in the written register such as *so* and *very* commonly occur in the spoken register too, however, the adverb *really* clearly prevails. The other adverbs constitute a minority, e.g. *particularly*, *absolutely*, *slightly*.

25. *I feel really bad but I'll see I might just see how I get on with packing for holiday.*

We also wanted to know whether the verb *feel* takes adjectives which are complemented by more than one adverb. In the written register, there are not so many instances (99 hits in total). In many cases *so* is the first adverb and it is followed by *very* or *incredibly* (see ex. 26). In a few examples, we can observe the combination of *pretty damn* or *pretty darn*, the former is more frequent (see ex. 27). Also, the adverb *really* is followed by the same adverb which constitutes the repetitive *really really* (ex. 28).

26. *We feel so very fortunate to have met up with the team at Love 146.*
27. *So I imagine it feels pretty darn fantastic to be a few months away from the release.*
28. *Even though you are hearing "Don't worry about it, I'm fine" you just still feel really really bad that it happened at all.*

In the spoken register, the corpus returned 50 matches and we can observe that one combination of adverbs prevails, that is the combination of adverbs *really really* (ex. ____). This combination also occurs in the written register but in the spoken register it is more frequent. Also, it is in agreement with the occurrence of *really* in almost all cases of single adverb modifying the adjective of the copula *feel* in the spoken register. A similar combination of *very very* ranks second in frequency (ex. ____).

29. *[...] we all felt really really proud.*

30. *I mean you would feel very very awkward wouldn't you?*

4.2 The verb *look*

The verb *look* can be classified as a verb of type *be*, namely a verb of sensory perception, according to the classification of types of copular verbs by Dušková et al. (2014). Biber and Conrad (2009) express that it is a verb describing positive and negative feelings about physical appearance.

	Adjective	No.of occurrence	Instances per million
1	<i>good</i>	767	6,426768
2	<i>great</i>	545	4,566609
3	<i>bad</i>	144	1,20659
4	<i>different</i>	130	1,089283
5	<i>amazing</i>	124	1,039008
6	<i>beautiful</i>	100	0,83791
7	<i>nice</i>	99	0,829531
8	<i>delicious</i>	70	0,586537
9	<i>similar</i>	68	0,569779
10	<i>awesome</i>	65	0,544641
11	<i>fantastic</i>	59	0,494367
12	<i>cool</i>	58	0,485988
13	<i>familiar</i>	46	0,385439
14	<i>fabulous</i>	45	0,377059
15	<i>online</i>	43	0,360301

Total	5 423 hits		
-------	------------	--	--

Table 5: AAM: a frequency list of *look*

	Adjective	LogDice measure
1	<i>delicious</i>	8,13
2	<i>amazing</i>	8,11
3	<i>fabulous</i>	7,70
4	<i>awesome</i>	7,67
5	<i>great</i>	7,67
6	<i>fantastic</i>	7,66
7	<i>nice</i>	7,61
8	<i>gorgeous</i>	7,58
9	<i>beautiful</i>	7,53
10	<i>foolish</i>	7,43
11	<i>promising</i>	7,29
12	<i>good</i>	7,22
13	<i>bad</i>	7,20
14	<i>familiar</i>	7,16
15	<i>identical</i>	7,15

Table 6: AAM: a collocate list of *look*

	Adjective	No.of occurrence	Instances per million
1	<i>nice</i>	276	24

2	<i>good</i>	273	23,73913
3	<i>lovely</i>	86	7,478261
4	<i>alright</i>	68	5,913043
5	<i>amazing</i>	68	5,913043
6	<i>great</i>	48	4,173913
7	<i>beautiful</i>	36	3,130435
8	<i>okay</i>	36	3,130435
9	<i>fine</i>	35	3,043478
10	<i>awful</i>	32	2,782609
11	<i>cool</i>	31	2,695652
12	<i>different</i>	31	2,695652
13	<i>bad</i>	30	2,608696
14	<i>terrible</i>	30	2,608696
15	<i>old</i>	26	2,26087
Total	1 989 matches		

Table 7: SBNC: a frequency list of *look*

	Adjective	Dice coefficient
1	<i>nice</i>	0,034
2	<i>amazing</i>	0,033
3	<i>lovely</i>	0,029
4	<i>beautiful</i>	0,023
5	<i>terrible</i>	0,019
6	<i>awful</i>	0,019

7	<i>great</i>	0,016
8	<i>good</i>	0,016
9	<i>alright</i>	0,016
10	<i>gorgeous</i>	0,014
11	<i>horrible</i>	0,013
12	<i>stupid</i>	0,012
13	<i>cool</i>	0,012
14	<i>young</i>	0,011
15	<i>ridiculous</i>	0,01

Table 8: SBNC: a collocate list of *look*

4.2.1 Comparing frequency and collocate lists within a single corpus

Different, *similar*, *cool* and *online* are included only in the frequency list, while *identical*, *gorgeous*, *foolish* and *promising* are represented only in the list of collocates. The remaining adjectives are identical, however, they differ in their rankings. If we compare the frequency list (Table 5) with the list of collocates (Table 6) we can see that the adjective *good* (ex. 31) outnumbers as a top frequent word the other adjectives in the former list. On the contrary, *good* ranks twelfth in the latter. *Delicious* (ex. 32) is the adjective *look* collocates the most with and together with the other adjectives with positive connotation such as *amazing*, *fabulous*, *awesome* or *great* they make up the top five highest ranked adjectives. Predominantly, there are more positively oriented than negatively oriented adjectives in both lists. In the frequency list there are also a few adjectives which are rather neutral such as *different*, *similar*, *familiar* or *online*. The adjective *different* may be understood as negative in some contexts. As example 33 illustrates, we are not sure how the person towards whom the property is ascribed looks and whether *different* is a positive evaluation. Among the highest ranked adjectives of the first list there is one adjective with negative connotation- *bad* (ex. 34) which, which, similarly to *good*, ranks only thirteenth in the second list.

31. *Some hospitals look good on some measures and not on others.*
32. *These mushrooms look delicious and would be wonderful with grilled chicken.*
33. *She looks different somehow tonight.*
34. *It always looks bad to be late and can be detrimental to your business or personal relationships.*

Looking at spoken data, table 7 shows the frequency, and we can observe that the adjective *nice* ranks first. The same adjective, i.e. *nice* (ex. 35), ranks first also among the collocates, which confirms the collocational strength of this particular word combination. The set of adjectives in both lists is roughly the same, especially in higher positions, only the rankings differ. For example the adjective *terrible* (ex. 36) is more prominent in the collocate list, while the adjective *good* (ex. 37) ranks second in the frequency list and only eighth among the collocates. The adjectives *okay*, *fine*, *different*, *bad* and *old* appear in the frequency list only, while the adjectives *gorgeous*, *horrible*, *stupid*, *young* and *ridiculous* appear only in the list of collocates.

35. [...] *there are times when you want to look nice and attractive [...]*
36. *I mean I got a middle parting it done for my hair and it looks terrible [...]*
37. *Well, everybody looks good when they're young.*

The adjectives are mostly positive (e.g. *beautiful*, *lovely*, *amazing*) in both lists with a few negative ones (e.g. *terrible*). The frequency list also displays two rather neutral adjectives (*different*, *old*).

4.2.2 Comparing collocate lists across the corpora

Table 6 shows the collocate list reflecting the written language data, while Table 8 reflects on the collocates of the verb *look* obtained from the spoken language data corpus.

The two collocate lists differ greatly. Only six adjectives appear in both lists and the other nine adjectives appear in one of the lists only. *Delicious*, *fabulous*, *awesome*, *fantastic*, *foolish*, *promising*, *bad*, *familiar* and *identical* only occur in the written register. *Lovely*, *terrible*, *awful*, *alright*, *horrible*, *stupid*, *cool*, *young* and *ridiculous* occur among the highest ranked collocates in the spoken register.

Even though the adjectives in the lists differ, they are mostly positive in both registers. Especially in the written register there are eleven adjectives with a positive connotation, two with a negative connotation (*foolish, bad*) and two adjectives which can be considered neutral (*familiar, identical*). *Delicious* is the highest ranked collocate of the verb in the written register, *nice* is ranked first in the spoken register. While *delicious* is not included in the spoken list, *nice* ranks seventh in the written. In the spoken register there are slightly more negative adjectives (*terrible, awful, horrible, stupid, ridiculous*) and one neutral. The neutral adjective *young* may be considered also positive or negative in some contexts. The adjectives which are shared by both lists are solely positive: *amazing, great, nice, gorgeous, beautiful* and *good*. Undoubtedly, the positive adjectives may be used as synonyms in contexts in which a speaker evaluates something positively and the difference may depend on the level of enthusiasm. Example 38 is an instance of *amazing* used about someone's appearance, while example 39 shows its use when talking about food. Example 40 shows the use of *delicious*. These examples are from the written language data, example 41 shows how *amazing* is used in the spoken register.

38. *She looks amazing in these photos.*

39. *This looks amazing, can't wait to eat it.*

40. *That bread looks delicious.*

41. *Oh so now I have to go and look amazing.*

4.2.3 Semantic categories of the adjectives

The adjectives may be divided into the following categories:

- a. a positive evaluation of appearance,
- b. a negative evaluation of appearance,
- c. a relation to previous experience.

The categories are established based on the fact that the verb *look* refers to something that can be seen. The verb *look* collocates the most with positive adjectives. Examples 38, 39 illustrate that it is not only the appearance of animate subjects that is evaluated but also inanimate.

In the written register, the majority of adjectives belong to the first category: *delicious, amazing, fabulous, awesome, great, fantastic, nice, gorgeous, beautiful, promising* and *good*. The second category includes the adjectives *foolish* and *bad*. The third category may seem a little abstract but the adjectives *familiar* and *identical* are connected with our previous experience and whether we perceive it as familiar or identical to something else.

In the spoken register, we only have evidence of the first two categories. The adjectives *nice, amazing, lovely, beautiful, great, good, alright, gorgeous, cool* and *young* belong to the first category. The adjectives *terrible, awful, horrible, stupid* and *ridiculous* belong to the second category.

Biber and Conrad (2009) state the verb *look* describes positive and negative feelings about physical appearance and also mention a number of adjectives for both groups of feelings (see chapter 2.2.1.). In terms of the semantics, the categories correspond. As was already mentioned, in the collocate list of the written register, there are not so many negative adjectives and the ones stated by Biber and Conrad do not appear in the list. In terms of the positive adjectives provided by these authors, their set of adjectives appear among the collocates of both lists, except for the adjective *happy*. In the collocate list of the spoken register there are a few negative adjectives, however, only *awful* and *terrible* correspond to the adjectives mentioned by Biber and Conrad.

Malá (2014) states that *look* is most frequent in conversation and in newspapers. In both registers, the adjectives are also mostly positive or neutral. If we compare the adjectives provided in her lists (see chapter 2.2.2.), we can see that the adjectives correspond only remotely which may be the result of different genres in the data.

4.2.4 Adverbs occurring in the structure

Let us first comment on the use of single adverbs which modify the adjectives in the written register. In total, the corpus returned 1 887 hits after searching the construction in question, i.e. *feel+ adverb+ adjective*. The following four adverbs cover more than a half. *So* ranks first, *very* second, *pretty* third and *really* fourth. To be more specific, the first two adverbs, i.e. *so* and *very*, outnumber the other adverbs quite significantly. Apart from *pretty*, which usually

expresses the quality only to a certain degree, the other adverbs express rather an extremely high degree of the property. Examples below demonstrate their uses:

42. *This cake looks so yummy that I can't wait to try baking it.*

43. *I've seen some cool tribal tattoos that do this and they look very nice.*

44. *It looks pretty new so I'm thinking it must be post 1992, [...]*

45. *After finishing the shading and other elements the suit looked really good.*

In the spoken register, the corpus returned 1 144 matches. One adverb dominates and covers a great part of the instances the corpus returned. It is the adverb *really* (ex. 46) which expresses the meaning of “thoroughly” and which ranks fourth in the written register. *So*, *very* and *quite* appear also very frequently, however, still in fewer instances than *really*. Other adverbs appear only insignificantly and include, for example, *absolutely*, *completely*, *fairly*.

46. *God she looks really good in that photo.*

In the written register, the verb *look* does not occur very often with more than one adverb modifying the adjective (85 hits in total). When it does, the adverbs are usually *pretty darn/ damn* (ex. 47) or *so damn/ darn/ goddamn* (ex. 48).

47. *And I am looking pretty damn good!*

48. *CSI does this a lot and it's one of the reasons their shows look so damn good.*

In the spoken register, the combination of adverbs is different. The corpus returned 41 matches. The repetition of adverb *really*, i.e. *really really*, occurs in almost all cases (ex. 49). Example 50 shows that in one case the adverb is even repeated four times. Generally, there is a tendency of the adverbs to repeat as it is possible to spot *very very*, *so so*, *quite quite* or *pretty pretty pretty*. These, however, appear in lower frequency.

49. *[...]she changed she put her photo up and she looked really really pretty in it.*

50. *[...] they just look really really really really muddy.*

4.3 The verb *seem*

The verb *seem* is a representative of the type *be* copular verbs. The verb refers to the factuality of the utterance. Verbs of this type function as means of epistemic modality and are subsumed under the second category of type *be* copulas defined by Dušková et al. (2014).

	Adjective	No.of occurrence	Instances per million
1	<i>obvious</i>	127	1,064145
2	<i>strange</i>	125	1,047387
3	<i>reasonable</i>	108	0,904943
4	<i>clear</i>	108	0,904943
5	<i>impossible</i>	103	0,863047
6	<i>unlikely</i>	91	0,762498
7	<i>possible</i>	90	0,754119
8	<i>odd</i>	83	0,695465
9	<i>appropriate</i>	80	0,670328
10	<i>fair</i>	56	0,46923
11	<i>simple</i>	51	0,427334
12	<i>good</i>	47	0,393818
13	<i>important</i>	46	0,385439
14	<i>logical</i>	45	0,377059
15	<i>natural</i>	43	0,360301
Total	4 865 hits		

Table 9: AAM: a frequency list of *seem*

	Adjective	LogDice measure
1	<i>strange</i>	8,85
2	<i>unlikely</i>	8,65
3	<i>obvious</i>	8,59
4	<i>reasonable</i>	8,43
5	<i>impossible</i>	8,32
6	<i>odd</i>	8,29
7	<i>daunting</i>	7,91
8	<i>logical</i>	7,66
9	<i>absurd</i>	7,61
10	<i>trivial</i>	7,50
11	<i>plausible</i>	7,43
12	<i>overwhelming</i>	7,38
13	<i>appropriate</i>	7,35
14	<i>silly</i>	7,35
15	<i>inevitable</i>	7,29

Table 10: AAM: a collocate list of *seem*

	Adjective	No.of occurrence	Instances per million
1	<i>fine</i>	21	1,826087
2	<i>nice</i>	15	1,304348
3	<i>unlikely</i>	9	0,782609
4	<i>good</i>	8	0,695652

5	<i>odd</i>	8	0,695652
6	<i>ridiculous</i>	8	0,695652
7	<i>crazy</i>	7	0,608696
8	<i>fair</i>	7	0,608696
9	<i>happy</i>	6	0,521739
10	<i>stupid</i>	6	0,521739
11	<i>weird</i>	6	0,521739
12	<i>friendly</i>	5	0,434783
13	<i>logical</i>	5	0,434783
14	<i>natural</i>	5	0,434783
15	<i>pointless</i>	5	0,434783
Total	241 matches		

Table 11: SBNC: a frequency list of *seem*

	Adjective	Dice coefficient
1	<i>unlikely</i>	0,056
2	<i>logical</i>	0,031
3	<i>pointless</i>	0,023
4	<i>friendly</i>	0,018
5	<i>odd</i>	0,016
6	<i>natural</i>	0,014
7	<i>ridiculous</i>	0,012
8	<i>crazy</i>	0,01
9	<i>stupid</i>	0,008

10	<i>strange</i>	0,007
11	<i>fine</i>	0,007
12	<i>fair</i>	0,006
13	<i>happy</i>	0,005
14	<i>weird</i>	0,003
15	<i>nice</i>	0,002

Table 12: SBNC: a collocate list of *seem*

4.3.1 Comparing frequency and collocate lists within a single corpus

Table 9 shows the frequency of occurrence of adjectives with the verb *seem*, while table 10 shows the list of adjectival collocates of the same verb. *Obvious* ranks very high in both lists (ex. 51) and so does *strange*, first in the collocate list and second in the frequency list (ex. 52). Top positions are occupied by roughly the same adjectives from which we can conclude that *seem* appears with matching set of adjectives in terms of frequency and collocability. The lower positions display greater variation in the range of adjectives. The following seven adjectives appear in the frequency list only: *clear*, *possible*, *fair*, *simple*, *good*, *important* and *natural*, while the next set of adjectives appear among the collocates only: *daunting*, *absurd*, *trivial*, *plausible*, *overwhelming*, *silly* and *inevitable*.

51. *At first glance it seem obvious that trying 160 cases costs a great deal less than trying 2298 cases [...]*

52. *It seems strange that divinely inspired biblical writers seemed not to know this very fundamental principle of writing.*

In both lists the adjectives tend to be rather neutral, and they describe how the situation is perceived. In the frequency list there is one positively evaluating adjective- *good* (ex. 53). In the list of collocates there is none so purely positive as the adjectives are rather negative, e.g. *absurd* (ex. 54), or they are neutral, e.g. *logical* (ex. 55). Depending on the context, some adjectives presented in Table 10 may be considered neutral or positive too, e.g. *plausible* (ex. 56) or *appropriate* (ex. 57).

53. [...] *it has therefore seemed good to this holy and universal synod that nobody at all should do this from now on [...]*
54. *Such actions seem absurd in terms of military logic.*
55. *It seems logical they would have something similar for your iPad.*
56. *This seems plausible and perhaps holds for other perceptual preferences as well.*
57. *It seemed appropriate to take a break from our drive to capture it.*

Now the spoken language data are contrasted, i.e. Table 11 is compared with Table 12. In the frequency list, commonly used adjectives like *fine* (ex. 58) and *nice* (ex.59) rank first and second respectively as opposed to their rankings in the list of collocates in which *fine* ranks only eleventh and *nice* ranks fifteenth. In the collocate list, *unlikely* (ex. 60) ranks first and *logical* (ex. 61) second which are semantically completely different adjectives than the first two from the frequency list. Even though the rankings of the adjectives significantly differ the lists display almost the same adjectives. The only exception to it is the adjectives *good* and *strange* which occur only in the frequency list and collocate list, respectively.

58. [...] *then it'll be time for second dinner which seems fine.*
59. *But anyway [name] seems nice I've only met her the once really.*
60. [...] *it seems unlikely that they would allow someone like that.*
61. [...] *it seemed logical to me at that time.*"

4.3.2 Comparing collocate lists across the corpora

Now let us compare the two collocate lists of the verb *seem* to highlight similarities and differences between the two registers. Table 10 shows the collocate list of the written language data, Table 12 shows the collocate list of the spoken language data.

The two lists are very different as only four adjectives are shared by both of the lists: *unlikely*, *strange*, *odd* and *logical*. While *unlikely* ranks first in the spoken register and second in the written register, *strange* ranks first in the written register and tenth in the spoken register. Examples 62, 63 show uses of the adjective *unlikely* in the spoken and written register, respectively.

62. *It seems unlikely given that you've only got an hour to get ready.*

63. *It was a title triumph that seemed unlikely given various setbacks and struggles.*

Even though the adjectives in the list vary to a great extent, semantically they share some features. First of all, they are mostly negative, e.g. *absurd* (AAM), *pointless* (SBNC), or neutral adjectives, e.g. *obvious* (AAM), *natural* (SBNC), in terms of connotation. On the other hand, there are also positive adjectives, especially among the collocates of the spoken register, namely *friendly*, *fine*, *happy* and *nice* (ex. 64). Depending on the context, even in the written register there are some adjectives which may be considered positive such as *reasonable*, *plausible* or *appropriate* (ex. 65).

64. [...] *I don't know him that well but he seems nice.*

65. *The triangular shape of the logo seemed appropriate as it if with the mountain scheme [...]*

As there are several synonymous adjectives in the lists, we can observe which of the adjectives is preferred in individual register. The adjectives *strange* and *odd*, which describe something unusual or unexpected, rank first and sixth in the written register, respectively. However, in the spoken register, *odd* ranks higher than *strange*. In the same list, another synonymous adjective appear- *weird*, which ranks even lower than *strange*. We can conclude that *strange* is preferred in the written register and *odd* in the spoken register.

Other synonymous adjectives, which express a lack of sense or intelligence, are *absurd*, *trivial* and *silly* in the written register and *stupid*, *ridiculous* or *pointless* in the spoken register. In the list of collocates of the written register, *absurd* ranks higher than the other adjectives, while, in the spoken register, *pointless* ranks the highest of the three adjectives. Another set of synonymous adjectives comprises *logical*, *natural* and *reasonable* which express something expected and very often based on reason or good judgement. While *logical* appears in both lists, *natural* can be found only in the spoken register and *reasonable* only in the written register. The conclusion can be that out of the three adjectives *logical* is more common in the spoken register and *reasonable* in the written register.

4.3.3 Semantic categories of the adjectives

The verb *seem* in its meaning refers to a degree of factuality of what is being said. The adjectives with which the verb collocates the most also reflect this meaning. The adjectives are mostly neutral or negative in terms of connotation, however, positive connotation also occurs. They can be classified into the following categories:

- a. distinctiveness,
- b. a lack of sense,
- c. probability,
- d. obviousness,
- e. a sense of rationality,
- f. a positive impression,
- g. a negative impression.

In the written register, the first category comprises the adjectives *strange* and *odd*. The second category includes the adjectives *absurd*, *trivial* and *silly*. The representatives of the third category describe the extent to which something is likely to happen or be the case. It is represented by the adjectives *unlikely*, *impossible* and *inevitable*. The fourth category comprises the adjective *obvious* and describes that something is observable and, therefore, apparent. The adjective *reasonable*, *logical* and *plausible* are connected with rational judgement. It must be noted that some of the adjectives that were assigned one category may well belong to another category, depending on the context. For example, the adjective *overwhelming* describes both a positive and a negative thing. *Appropriate* is assigned the positive impression category, while *daunting* the negative impression category.

In the spoken register, the set of categories of adjectives mostly corresponds with the one established for the written register. However, the adjectives which constitute the categories are often different and different categories prevail. The largest group of adjectives comprises *friendly*, *fine*, *fair*, *happy* and *nice* and belongs to the sixth category. The adjectives of this category may be referred to as giving a positive impression. The second largest group of adjectives belongs to the second category, i.e. the category where the adjectives describe a lack of sense. The adjectives are *pointless*, *ridiculous*, *crazy* and *stupid*. The first category is represented by three adjectives: *odd*, *strange* and *weird*, while the fifth category is represented

by two adjectives: *logical* and *natural*. The third category is represented only by the adjective *unlikely*. The categories of adjectives describing obviousness and negative impression are not present in the spoken register.

Biber and Conrad (2009) express that the verb *seem* is connected with likelihood and provide a set of relating adjectives. From the adjectives provided by the authors only *obvious*, *reasonable* and *unlikely* appear in the collocate list of the written register. In the present analysis these adjectives are grouped with other adjectives comprising three different categories. Among the collocates of the verb *seem* in the spoken register only *unlikely* was found in our data.

Malá (2014) points out that *seem* in the combination with the adjectival complement is especially very common in fiction with adjectives *unlikely*, *good*, *likely*, *pleased*, *surprised*, *happy*, *interested*, *right*, *strange* and *able*. In the present analysis, especially in the spoken register, there are also many adjectives with positive connotation. She also mentions that the verb is frequent in academic prose with a following set of adjectives which resembles more the adjectives presented in Table 10: *likely*, *unlikely*, *reasonable*, *appropriate*, *clear*, *possible*, *probable*, *obvious*, *certain* and *plausible*.

4.3.4 Adverbs occurring in the structure

The co-occurring adjectives of the verb *seem* are modified by a number of adverbs in the written register. The following six adverbs comprise more than a half of the hits the corpus returned (2 047 in total). Similarly to previous findings, *so* ranks first. Then the following set of adverbs follow *very*, *too*, *pretty*, *quite* and *almost*. The first three highest ranked adverbs, i.e. *so*, *very* and *too*, describe a great degree of the property the adjectives ascribe. On the other hand, the three lowest ranked adverbs of the six, i.e. *pretty*, *quite* and *almost*, express a fairly significant extent. *Quite* also means “completely” in some contexts. Particular uses of these adverbs in the corpus are illustrated by examples below:

66. *This humble wish seemed so impossible.*

67. *Therefore, it seems very likely that Joseph was the actual, historical person who buried Jesus in the tomb.*

68. *If I seem too critical, it's because I care a lot about his movies.*

69. *Everything seemed pretty normal until the night he was hospitalised with a twisted bowel.*

In the spoken register, the verb *seem* does not occur very frequently with an adverb modifying the adjective complementing it. The corpus returned only 276 matches in which the position of the adverb is almost equally occupied by the adverb *very*, *really*, *quite*. Although the adverb *really* does not occur as significantly in the written register, it is very frequent in the spoken register (ex. 70). On the other hand, the adverb *so* which ranks first in the written register is less frequent in the spoken register. The other adverbs which occur only few times in the spoken register include *too*, *terribly*, *incredibly*.

70. *You seem really excited about this thing.*

The corpus returned 121 hits and it shows that in several instances the adjective is modified by the following two combinations *so very* and *almost too* (ex. 71, 72). However, in the spoken register the adjectives of the copula do not seem to be modified by more than one adverb very often, as the corpus returned only 16 matches. Among the few matches the corpus returned there are cases of one adverb which repeats itself: *too too*, *so so*, *very very* or *quite quite* (ex. 73). Example 74 a case in which the adjective is preceded by three adverbs, while only two of the adverbs actually modify the adjective (*really really*). The use of *like* has already been commented on at the beginning of the analysis.

71. *He seemed so very kind, [...]*

72. *It seems almost too obvious.*

73. *He hasn't got any children, always wanted children so it seems quite quite good.*

74. *[...]it just seems like really really easy [...]*

4.4 The verb *remain*

The verb is classified as the type *be* copula, namely the category of verbs of which ascribe a property which remains unchanged (Dušková, 2014). Biber and Conrad (2009) point out that there is an absence of a change in a sentence in which the verb *remain* functions as the predicate.

	Adjective	No.of occurrence	Instances per million
1	<i>silent</i>	196	1,642303
2	<i>open</i>	190	1,592029
3	<i>unchanged</i>	171	1,432826
4	<i>true</i>	122	1,02225
5	<i>active</i>	113	0,946838
6	<i>intact</i>	111	0,93008
7	<i>high</i>	110	0,921701
8	<i>unclear</i>	103	0,863047
9	<i>anonymous</i>	102	0,854668
10	<i>strong</i>	94	0,787635
11	<i>available</i>	76	0,636811
12	<i>constant</i>	75	0,628432
13	<i>free</i>	67	0,5614
14	<i>unknown</i>	64	0,536262
15	<i>stable</i>	64	0,536262
Total	6 106 hits		

Table 13: AAM: a frequency list of *remain*

	Adjective	LogDice measure
1	<i>unchanged</i>	9,72
2	<i>silent</i>	9,56
3	<i>intact</i>	8,99

4	<i>unclear</i>	8,89
5	<i>anonymous</i>	8,63
6	<i>unresolved</i>	8,06
7	<i>constant</i>	7,89
8	<i>stable</i>	7,87
9	<i>active</i>	7,80
10	<i>uncertain</i>	7,75
11	<i>elusive</i>	7,68
12	<i>unknown</i>	7,65
13	<i>competitive</i>	7,61
14	<i>valid</i>	7,60
15	<i>vigilant</i>	7,36

Table 14: AAM: a collocate list of *remain*

	Adjective	No.of occurrence	Instances per million
1	<i>anonymous</i>	2	0,173913
2	<i>nameless</i>	2	0,173913
3	<i>adorable</i>	1	0,086957
4	<i>constant</i>	1	0,086957
5	<i>faithful</i>	1	0,086957
6	<i>intact</i>	1	0,086957
7	<i>invisible</i>	1	0,086957
8	<i>nice</i>	1	0,086957

9	<i>sealed</i>	<i>1</i>	<i>0,086957</i>
10	<i>sensible</i>	<i>1</i>	<i>0,086957</i>
11	<i>silent</i>	<i>1</i>	<i>0,086957</i>
12	<i>sorry</i>	<i>1</i>	<i>0,086957</i>
13	<i>unchanged</i>	<i>1</i>	<i>0,086957</i>
Total	<i>15 matches</i>		

Table 15: SBNC: a frequency list of *remain*

	Adjective	Dice coefficient
1	<i>nameless</i>	<i>0,2353</i>
2	<i>unchanged</i>	<i>0,1053</i>
3	<i>faithful</i>	<i>0,08</i>
4	<i>intact</i>	<i>0,0606</i>
5	<i>anonymous</i>	<i>0,0588</i>
6	<i>sealed</i>	<i>0,0556</i>
7	<i>invisible</i>	<i>0,0444</i>
8	<i>adorable</i>	<i>0,0364</i>
9	<i>constant</i>	<i>0,0156</i>
10	<i>silent</i>	<i>0,0141</i>
11	<i>sensible</i>	<i>0,0074</i>
12	<i>sorry</i>	<i>0,0004</i>
13	<i>nice</i>	<i>0,0001</i>

Table 16: SBNC: a collocate list of *remain*

4.4.1 Comparing frequency and collocate lists within a single corpus

Table 13 shows the ranking of adjectives in terms of their frequency and is compared with Table 14 which shows the ranking of collocates in the written register. *Silent* (ex. 75) ranks first, *open* (ex. 76) second and *unchanged* (ex. 77) third in the frequency list, while *unchanged* ranks first, *silent* second and *intact* (ex. 78) third in the collocate list. While *open* appears high in the frequency list, Table 14 shows that it does not have sufficient collocational strength to rank among the first fifteen collocates of the verb.

75. *You have an absolute right to remain silent during questioning [...]*

76. *Serving lines will remain open for 15 minutes after the doors close.*

77. *Food prices remained unchanged from April to June [...]*

78. *The heart of the story remains intact for both versions.*

After comparing Table 13 and Table 14 we may observe that the lists differ in some of the adjectives they display. Namely, the adjectives *open*, *true*, *high*, *strong*, *available* and *free* occur in the frequency list only. Contrarily, the adjectives *unresolved*, *uncertain*, *elusive*, *competitive*, *valid* and *vigilant* occur among the collocates only.

The AAM frequency list includes adjectives which describe that something remains the same-*unchanged*, *constant* and *stable* and which, therefore, may be considered to be synonyms. *Unchanged* is the most frequent one of the three, as it ranks third in the overall frequency. *Constant* ranks twelfth, whereas *stable* ranks fifteenth. The same adjectives appear among the collocates in the same order. Their use is illustrated by examples 79-81.

79. *The fundamental principles remain unchanged.*

80. *My weight had remained constant for the last few days [...]*

81. *The results remain stable if further variables are added to the model.*

Another set of synonyms can be found in both the frequency and the collocate list. The adjectives *anonymous* (ex. 82) and *unknown* (ex. 83) which describe that something or someone is not known or identified. *Anonymous* ranks higher than *unknown* in both lists.

82. *Please do not leave your name if you wish to remain anonymous.*

83. *Judges will remain unknown until the day of the event.*

Generally, the adjectives are rather neutral in terms of connotation in both lists.

The SBNC corpus returned only 15 hits after searching the verb *remain* + adjective. Therefore, the frequency list and the list of collocates do not have all fifteen positions covered.

The fact that the rankings of adjectives in the frequency list and the list of collocates do not often correspond has already been proven and the verb *remain* is not an exception. As there were only thirteen adjectives found with *remain*, the same adjectives occur in both lists in a very limited number of occurrences. Table 15 shows that *anonymous* (ex. 84) ranks first, *nameless* (ex. 85) second and *adorable* (ex. 86) third. Table 16, which displays the spoken list of collocates, shows that these adjectives have different rankings there and that only the adjective *nameless* appear among the first three collocates. In the collocate list, *nameless* ranks first, *unchanged* (ex. 87) second and *faithful* (ex. 88) third.

84. [...] quite sensibly they've decided to remain anonymous.

85. Oh it's with the girl who shall remain nameless.

86. He still is adorable, remains adorable.

87. All remained unchanged in the plural [...]

88. It wouldn't've been fair for him to remain faithful [...]

4.4.2 Comparing collocate lists across the corpora

Table 14 shows the collocate list obtained from the AAM corpus, i.e. the corpus of written language data. Table 16 shows the collocate list obtained from the SBNC corpus, i.e. the corpus of spoken language data. Table 16 does not show all fifteen collocates as there were not enough data.

If we compare the two lists of collocates, we can see that they are very different in terms of the set of adjectives they display. In the written register, the following group of adjectives appear: *unclear*, *unresolved*, *stable*, *active*, *uncertain*, *elusive*, *unknown*, *competitive*, *valid*, *vigilant*. In the spoken register we can see *nameless*, *faithful*, *sealed*, *invisible*, *adorable*, *sensible*, *sorry* or *nice*.

Despite the different number and range of adjectives the two lists display, the adjectives are similar in terms of connotation and very often they are synonymous. For example, the adjectives *anonymous*, *unknown* and *nameless* describe someone not identified by a name. While the first two adjectives appear in the written register with *anonymous* ranked higher, *nameless* can be found only in the spoken register together with *anonymous*, with *nameless* ranking first in the list overall. The use of the adjective *anonymous* in both registers is illustrated by examples 89 (AAM) and 90 (SBNC). Similarly, the adjective *unchanged* may be considered synonymous to *constant* in one of its meanings, as both of them describe something unaltered. In both registers *unchanged* ranks higher than *constant*. Examples 91, 92 illustrate the use of *unchanged* in both registers, i.e. in the written and spoken register, respectively. However, it needs to be reminded that the number of occurrences of the verb *remain* + adjective in the SBNC corpus is very limited- only 15 matches in total.

89. *The author, who chose to remain anonymous, insisted that people have no right to carry firearms [...]*

90. *You know they will remain anonymous if that's what they want [...]*

91. *[...] our position that we would not attend remains unchanged.*

92. *All remained unchanged in the plural [...]*

4.4.3 Semantic categories of the adjectives

The adjectives that the verb *remain* takes as complements can be divided into three categories. However, not all fifteen adjectives can be categorized, as there is a great variety of adjectives with individual meanings.

The adjectives are characterized by:

- a. persistency to a (negative) change,
- b. indefiniteness,
- c. remaining in a positive state.

In the written register, the first category comprises the adjectives *unchanged*, *intact*, *constant* and *stable*. These four adjectives are connected with the sense of something or someone remaining the same. In the case of *intact*, there is also the meaning of something not being

touched or ruined. The second category includes the adjectives *unclear*, *anonymous*, *uncertain* and *unknown*. The combination of these adjectives with the verb *remain* also refers to the absence of a change. To be exact, something or someone remains undefined. The adjectives *anonymous* and *unknown* are particularly connected with an absence of a name. The third category does not have any representative in the written register.

In the spoken register, all three categories mentioned have their representatives. The adjectives *unchanged*, *intact* and *constant* belong to the first category, while the adjectives *nameless* and *anonymous* belong to the second category. To some extent, the adjective *sealed* may also belong to the first category, as it describes something securely fastened, therefore, persistent to a change. In this register, we can also observe some positively evaluating adjectives, *adorable* and *nice*, which suggests remaining in a positive state. Similarly to the written register, some of the adjectives are not easy to categorize, as each of them has a distinct meaning.

Biber and Conrad (2009) also mention that the verb *remain* together with its complements describe an absence of a change and provide a set of particular adjectives (see chapter 2.2.1). Among the collocates in the written register all these adjectives occur too, except for the adjective *closed*. Among the collocates in the spoken register only *unchanged*, *constant* and *intact* occur in our data, too.

Malá (2014) points out that the verb appears with stative adjectival complements and that it describes the continuation of a state. In the findings of the present analysis we can see that the complements are also mostly stative and apart from the adjectives *strong*, *true*, *loyal* and *important* they appear in one or the other register. However, dynamic adjectives also occur in both registers, e.g. *active*.

4.4.4 Adverbs occurring in the structure

The adverbs modifying the adjectives co-occurring with the verb *remain* show a great variety in the written register. The corpus returned 949 hits and the highest ranked adverb *relatively* (ex. 93) covers only a small part of it (94 instances). In contrast to previous verbs, there is no one adverb that would be dominant in the list. The following adverbs cover a half of the hits: *largely*, *very*, *fairly*, *highly*, *completely*, *essentially*, *quite*, *so*, *too* or *almost*. Interestingly

enough, the adverbs appearing in first positions are different from the highest ranked adverbs occurring after the previous verbs, e.g. *so* which ranks first there but here it ranks ninth. The adverbs express a moderately high degree or that the property is valid only to some extent, e.g. *relatively, fairly, quite, almost*, or a high degree, e.g. *largely, very, highly, completely, so* and *too*. The adverb *essentially* refers to the basic or fundamental nature of something (ex. 94).

93. *The only penguins that remain relatively secure are those that breed in Antarctica, [...]*

94. *Many of them also remain essentially unchanged from their bare beginnings.*

In contrast to the corpus of the written register, the corpus of the spoken register returned only 2 matches and the adverbs are *completely* and *ultra*. Examples below show their use:

95. *[...] or to try and develop skills that would otherwise have remained completely dormant.*

96. *You have to remain ultra cool.*

In the written register, the corpus returned very little hits of the verb *remain* followed by more than one adverb complementing the adjective (30 hits in total). One such construction is illustrated by example 97 in which the adverb *almost* is followed by another adverb, *completely*.

97. *So far, they remained almost completely silent.*

The SBNC did not return any matches when searching for constructions with more adverbs modifying the adjective after *remain*. It is understandable when we look at the low number of matches of the construction with a single adverb.

4.5 The verb *become*

The verb *become* belongs to the second type of copular verbs which express a change of a state, if we use the division of copular verbs presented by Dušková et al. (2014). Biber and Conrad (2009) state that this verb describes a change in understanding or importance.

	Adjective	No.of occurrence	Instances per million
1	<i>aware</i>	638	5,345865
2	<i>available</i>	597	5,002322
3	<i>clear</i>	571	4,784465
4	<i>apparent</i>	331	2,773482
5	<i>pregnant</i>	242	2,027742
6	<i>interested</i>	217	1,818264
7	<i>effective</i>	208	1,742852
8	<i>popular</i>	205	1,717715
9	<i>familiar</i>	198	1,659061
10	<i>necessary</i>	193	1,617166
11	<i>obvious</i>	145	1,214969
12	<i>active</i>	140	1,173074
13	<i>evident</i>	134	1,122799
14	<i>famous</i>	129	1,080904
15	<i>possible</i>	128	1,072525
Total	14 303 hits		

Table 17: AAM: a frequency list of *become*

	Adjective	LogDice measure
1	<i>aware</i>	9,85
2	<i>apparent</i>	9,31
3	<i>clear</i>	9,04

4	<i>pregnant</i>	8,86
5	<i>available</i>	8,52
6	<i>familiar</i>	8,38
7	<i>evident</i>	8,09
8	<i>interested</i>	8,08
9	<i>popular</i>	7,99
10	<i>obvious</i>	7,87
11	<i>effective</i>	7,87
12	<i>necessary</i>	7,71
13	<i>famous</i>	7,69
14	<i>extinct</i>	7,61
15	<i>active</i>	7,53

Table 18: AAM: a collocate list of *become*

	Adjective	No.of occurrence	Instances per million
1	<i>famous</i>	10	0,869565
2	<i>clear</i>	6	0,521739
3	<i>difficult</i>	6	0,521739
4	<i>aware</i>	5	0,434783
5	<i>apparent</i>	4	0,347826
6	<i>cocky</i>	4	0,347826
7	<i>independent</i>	4	0,347826
8	<i>normal</i>	4	0,347826

9	<i>permanent</i>	4	0,347826
10	<i>pregnant</i>	4	0,347826
11	<i>automatic</i>	3	0,26087
12	<i>fluent</i>	3	0,26087
13	<i>illegal</i>	3	0,26087
14	<i>immune</i>	3	0,26087
15	<i>popular</i>	3	0,26087
Total	229 matches		

Table 19: SBNC: a frequency list of *become*

	Adjective	Dice coefficient
1	<i>apparent</i>	0,032
2	<i>cocky</i>	0,031
3	<i>famous</i>	0,026
4	<i>permanent</i>	0,022
5	<i>immune</i>	0,022
6	<i>independent</i>	0,022
7	<i>fluent</i>	0,021
8	<i>automatic</i>	0,018
9	<i>aware</i>	0,017
10	<i>illegal</i>	0,014
11	<i>pregnant</i>	0,014
12	<i>clear</i>	0,011
13	<i>popular</i>	0,009

14	<i>difficult</i>	0,006
15	<i>normal</i>	0,005

Table 20: SBNC: a collocate list of *become*

4.5.1 Comparing frequency and collocate lists within a single corpus

Table 17 shows the frequency list in which we can observe that the adjective *aware* ranks first. Interestingly, *aware* (ex. 98) also ranks first in the collocate list which is shown by Table 18. The verb *become* is the first verb from the group of analysed verbs which has the same adjectives in the first position in both lists.

98. *During those conversations he became aware that they wished to marry [...]*

Tables show that the adjectives listed do not differ greatly. To be precise, there is only one adjective in which the two lists differ. In the frequency list it is the adjective *possible* (ex. 99), which frequently combines with the verb *become* in *it*-clefts. In the list of collocates, it is the adjective *extinct* (ex. 100). As both adjectives mentioned appear towards the end of the list, we may conclude that the lists are roughly the same in terms of the adjectives they include.

99. *As we learn how to control stem cells it may become possible to correct these faults.*

100. *Davis points out that languages become extinct when people become extinct.*

We may observe a number of synonyms among the adjectives in both lists. The adjectives *clear* (ex. 101), *apparent* (ex. 102, 103), *obvious* (ex. 104) and *evident* (ex. 105), which describe that something is easily visible or understood, appear in both the frequency list and the collocate list. These adjectives are ranked differently in individual lists. In the frequency list, *clear* is the most common one followed by *apparent*, *obvious* and *evident*. Conversely, *apparent* collocates with *become* more strongly than the other three adjectives then *clear*, *evident* and *obvious* follow. These adjectives frequently appear in *it*-clefts which is illustrated by examples 105 and pseudo-clefts (ex. 103).

101. *It soon became clear that the service was down again.*

102.*It has become apparent that no one nation could successfully mount a serious exploration of Mars.*

103.*What becomes apparent during the energetic set is that Penate is a really nice guy [...]*

104.*Now his problem became obvious to everyone.*

105.*It has become evident that the crisis actually consists of four distinct crisis [...]*

The SBNC frequency list of adjectives occurring with the verb *become* is represented by Table 18 and the list of adjectival collocates of the same verb by Table 19. When we compare the two lists, we can see that in terms of the range of adjective the two lists correspond to each other. However, the rankings of the adjectives differ as expected, for example *famous* (ex. 106) is ranked first in the frequency list and third in the collocate list. Similarly, *apparent* (ex. 107) is ranked first in the collocate list and fifth in the frequency list. Correspondingly, the other adjectives also rank differently in the two lists, such as *clear* (ex. 108) but it does not occupy such a prominent position among the collocates.

106.*[...] Billy Bragg who didn't get any good qualifications but became famous by just being a singer [...]*

107.*[...] then it became apparent that everyone was like panicky about doing work for the final term [...]*

108.*It won't become clear until you know how it works [...]*

4.5.2 Comparing collocate lists across the corpora

The two tables are very different in terms of the range of adjectives they display. *Available, familiar, evident, interested, obvious, effective, necessary, extinct* and *active* are the adjectives which appear in the written register only. *Cocky, permanent, immune, independent, fluent, automatic, illegal, difficult* and *normal* appear solely in the spoken register.

Generally speaking, *apparent* ranks high in both lists, namely, it ranks first in the spoken register and second in the written register. Referring to the point made above, the two lists differ not only in the range of adjectives but also in the rankings of the co-occurring adjectives.

The four synonymous adjectives *apparent*, *clear*, *evident* and *obvious*, which describe something clearly visible or understood, occur in the written register. *Apparent* is the highest ranked of these adjectives, but it is also the highest ranked adjective among the collocates of the verb in the spoken register. In the spoken register, two of the adjectives mentioned are not included in the list, namely the adjectives *evident* and *obvious*. The ranking of *apparent* was provided above, *clear* is ranked third in the written register and twelfth in the spoken register, therefore, it is more significant for the former register. Examples 109 (AAM) and 110 (SBNC) illustrate the use of *clear* in both registers.

109. *However, what became clear was that I needed to exercise, [...]*

110. *It won't become clear until you know how it works [...]*

The adjectives *popular* and *famous* are also considered to be synonyms, as they both express that someone or something is well-known. In the spoken register, *famous* (ex. 111) is ranked third and overall it is ranked higher than *popular*. Conversely, in the written register *popular* (ex. 112) is ranked higher than *famous* but generally these two adjectives appear towards the end of the list of collocates.

111. *[...] and so she became famous because of the way she spoke [...]*

112. *[...], the are became popular for picnics and swimming.*

4.5.3 Semantic categories of the adjectives

The adjectives can be divided into the following categories. They are connected with:

- a. obviousness,
- b. attention,
- c. a personal or public act or its result,
- d. a change in operation/ understanding or necessity.

In the written register, there are adjectives *apparent*, *clear*, *evident* and *obvious* which refer to something that can be easily perceived or understood and which can be assigned the first category. The second category is connected with consciousness or the focus of attention on something and comprises the adjectives *aware* and *interested*. The third category is the largest and it comprises the adjectives *pregnant*, *available*, *familiar*, *popular*, *famous*, *extinct* and

active. As the title of the category suggests, it is connected with public or people in general and with their decisions or acts which bring about some results. The fourth category includes diverse adjectives which, however, share that there is a change in either necessity (*necessary*) or operation (*effective*). Both the first and second category have more representatives in the written register than in the spoken register.

In the spoken register, the third category is even larger and comprises *cocky*, *famous*, *immune*, *independent*, *fluent*, *illegal*, *pregnant* and *popular*. The adjectives *apparent* and *clear* belong to the first category, while the adjective *aware* belong to the second category. The fourth category comprises *permanent*, *automatic*, *difficult* and *normal*.

Biber and Conrad (2014) state that the verb *become* together with its set of adjectival complements constitute a category which reflects a change in understanding or importance. Their set of adjectives is similar to the two lists provided in the present analysis (Table 18 and 20), however, the adjectives are assigned different and more detailed categories, as some of the adjectives in both lists showed semantic resemblance on more levels.

Malá (2014) suggests that the verb *become* and its adjectives describe the result of a change of a state and provides a set of frequently occurring adjectives (see chapter 2.2.2). If we compare the lists provided in the present analysis with her set, we can see that the adjectives correspond to a great extent.

4.5.4 Adverbs occurring in the structure

Malá (2014) asserts that the verb *become* occurs frequently with an adjective modified by a degree adverbial. In the written register, we can also see this tendency, however, in the spoken register, there are fewer instances of such constructions.

In the written register, there are four dominant adverbs which modify the adjectives. They cover a half of the hits the corpus returned (3603 in total). The adverbs are *increasingly*, *very*, *so* and *too*. *Increasingly* ranks first. This adverb refers to progress and growth and can be paraphrased by “more and more” (ex. 113). The other three adverbs occurred frequently in previous chapters of the analysis. *Very*, *so* and *too* (ex. 114) when combined with an adjective all refer to a high degree of the property.

113.*Drink Naples is becoming increasingly popular with a younger generation of both Italians and foreigners.*

114.*Things have become too predictable.*

In the spoken register, the corpus did not return as many matches as in the written register (117 matches in total). The copula *become* occurs with an adverb modifying its complementing adjective only in 117 cases which is even less than the corpus returned in the case of the verb *seem*. Most of the instances of the adverbs are equally represented by the adverbs *really* (ex. 115), *quite*, *so* and *very*. Conversely, *really* and *quite* do not occur among the highest ranked adverbs of the written register. The other adverbs the corpus shows occur only infrequently, e.g. *terribly*, *completely*.

115.*So he was running this stall single-handed when it became really popular.*

In the written register, the corpus returned 121 hits. The adjectives of the verb *become* are frequently modified by the adverb *so* in combination with *very*, *deeply* and *incredibly* (ex. 116). Also the double appearance of one adverb is possible, cf. *very very* (ex. 117).

116.*It is interesting how your statements become so very qualitative when dealing with this.*

117.*Free markets allowed some people to become very very rich, but this was at the expense of the rest of the population.*

Conversely, the corpus of spoken language data returned only 7 matches when searching for constructions with more adverbs modifying the adjective. There is not a single combination prevailing or occurring more than once.

4.6 The verb *go*

The verb *go* is another representative of copular verbs of the type *become* which express a change of state (Dušková, 2004). Biber and Conrad (2009) add that this verb specifically describes a change to a negative state. It was necessary to exclude some constructions which were irrelevant for the analysis because either the second element was not an adjective or the verb *go* did not function as a copular verb expressing a change. Even though the irrelevant constructions appeared high in the lists, they had to be filtered out. Examples of excluded

items include: *outside, strong, back, great, first, next, full* and *deep*. Examples below show their use. Very often the verb has a different meaning and is followed by an adverbial (ex. 118) or the construction functions as a phrase in which the meaning does not involve a change of a state (ex. 119), the verb simply does not express a change of a state (ex. 120) or the verb expresses a process (ex. 121) . Such constructions appeared in both corpora and were left out of the analysis.

118.[...] *Jeanette has to have an escort when she goes outside.*

119.*A few days later, they are still going strong on our dining room table.*

120.*Where do we go next?*

121.*Running workouts have been going great.*

	Adjective	No. of occurrence	Instances per million
1	<i>wrong</i>	993	8,320445
2	<i>online</i>	254	2,128291
3	<i>live</i>	223	1,868539
4	<i>crazy</i>	178	1,49148
5	<i>viral</i>	152	1,273623
6	<i>bad</i>	148	1,240107
7	<i>public</i>	148	1,240107
8	<i>unnoticed</i>	146	1,223348
9	<i>bankrupt</i>	113	0,946838
10	<i>green</i>	88	0,737361
11	<i>wild</i>	85	0,712223
12	<i>mad</i>	83	0,695465
13	<i>hungry</i>	80	0,670328

14	<i>extinct</i>	58	0,485988
15	<i>free</i>	53	0,444092
Total	5 907 hits		

Table 21: AAM: a frequency list of *go*

	Adjective	LogDice measure
1	<i>wrong</i>	10,61
2	<i>unnoticed</i>	9,60
3	<i>viral</i>	9,49
4	<i>crazy</i>	9,27
5	<i>bankrupt</i>	9,18
6	<i>mad</i>	8,49
7	<i>hungry</i>	8,40
8	<i>extinct</i>	8,23
9	<i>wild</i>	8,06
10	<i>unpunished</i>	7,85
11	<i>online</i>	7,80
12	<i>unanswered</i>	7,73
13	<i>green</i>	7,49
14	<i>unchallenged</i>	7,49
15	<i>smoke-free</i>	7,41

Table 22: AAM: a collocate list of *go*

	Adjective	No. of occurrence	Instances per million
--	-----------	-------------------	-----------------------

1	<i>wrong</i>	67	5,826087
2	<i>crazy</i>	64	5,565217
.	<i>mad</i>	62	5,391304
4	<i>mental</i>	52	4,521739
5	<i>funny</i>	25	2,173913
6	<i>black</i>	22	1,913043
7	<i>wild</i>	20	1,73913
8	<i>bankrupt</i>	18	1,565217
9	<i>online</i>	18	1,565217
10	<i>blind</i>	17	1,478261
11	<i>red</i>	16	1,391304
12	<i>cold</i>	15	1,304348
13	<i>grey</i>	15	1,304348
14	<i>quiet</i>	15	1,304348
15	<i>bust</i>	14	1,217391
Total	1 386 matches		

Table 23: SBNC: a frequency list of *go*

	Adjective	Dice coefficient
1	<i>mad</i>	0,062
2	<i>mental</i>	0,055
3	<i>crazy</i>	0,05
4	<i>wrong</i>	0,029
5	<i>bankrupt</i>	0,026

6	<i>wild</i>	0,024
7	<i>blind</i>	0,021
8	<i>bust</i>	0,019
9	<i>online</i>	0,017
10	<i>bald</i>	0,016
11	<i>grey</i>	0,016
12	<i>quiet</i>	0,014
13	<i>mouldy</i>	0,0139
14	<i>black</i>	0,0128
15	<i>soft</i>	0,0125

Table 24: SBNC: a collocate list of *go*

4.6.1 Comparing frequency and collocate lists within a single corpus

As Table 21 and 22 extracted from the AAM show, *wrong* (ex. 122) ranks first in both lists. In the frequency list it exceedingly outnumbers the other adjectives in the number of occurrences. It also affirms Biber and Conrad's claim that the verb occurs with adjectives and together they constitute a change to a negative state. Other adjectives ranking high in the collocate list, which means they reflect a strong association with the verb, do not score high in the frequency list. For example the adjectives *unnoticed* and *viral* rank second and third, but they rank eighth and fifth in the frequency list, respectively. It is also true about the adjectives that appear at the end of the list of collocates but do not appear at all among the frequent adjectives. Consequently, the two lists differ to some extent as the frequency list displays some adjectives which does not occur in the collocate list and vice versa. The frequency adjectives are: *live*, *bad*, *public*, *free*, while the collocate list includes: *unpunished*, *unanswered*, *unchallenged*, *smoke-free*.

122.*If something goes wrong, the system as a whole functions very poorly.*

In both lists there are adjectives which are connected with media and internet and which are neutral in terms of connotation, e.g. *online*, *live*, *viral* or *public*. Examples 123-126 illustrate their use in particular.

123.*I went online looking for sources of help.*

124.*It was intended to go live yesterday [...]*

125.*Ideas can go viral in the real world too.*

126.*The project intends to go public on the practice of censorship worldwide.*

In both lists and especially in the list of collocates, there are negatively oriented adjectives among the fifteen highest ranked adjectives, such as: *wrong*, *mad*, *hungry*, *extinct* (ex. 127-130). The verb *go* in *go hungry* may also be perceived as a copula which expresses being in a state or staying in a state, e.g. *In spite of the relief effort, thousands of people continue to go hungry*. In example 129 the verb expresses rather a change of a state.

127.*This can help protect you if such a thing goes wrong.*

128.*I'm so worried I am going mad.*

129.*I was quickly learning that one does not go hungry while visiting Korea.*

130.*Humans will go extinct someday.*

As we have mentioned at the beginning of this subchapter, the verb *go* has many functions. Some adjectives had to be excluded from the lists extracted from the SBNC as they did not constitute copular predication. For example, *straight* as in *we could just go straight home* or *slow* as in *mum wants to go slow so they can't cycle together*. Given by the nature of the spoken language data, there are very often not clear boundaries between clauses and the verb often belongs to the end of a sentence and the adjective to the beginning of another one. For example: *Here we go (.) good old Google*.

The two SNBC lists differ to some extent as expected: the frequency list shows the common adjectives *funny*, *red* and *cold*, which did not make it to the list of collocates. On the other hand, the collocate list includes the adjectives *bald*, *mouldy* and *soft*.

If we compare the two SBNC tables, i.e. Table 23 and 24, we can see that the four highest positions are occupied by the same adjectives but with different rankings. The adjectives are

wrong (ex. 131), *crazy* (ex. 132), *mad* (ex. 133) and *mental* (ex. 134). In the frequency list *wrong* ranks first, *crazy* second, *mad* third and *mental* fourth. Among the collocates, *mad* ranks first, *mental* second, *crazy* third and *wrong* fourth.

131. *I mean that is when you know that fashion has just gone wrong.*

132. *He can't have too much sugar he goes crazy.*

133. *Oh my god the white people would go mad if you said that.*

134. *Well she goes mental when she comes round here doesn't she?*

Both of the lists represent adjectives which are rather negative (e.g. *crazy*, *mad*, *bankrupt*) or they relate to colours (e.g. *red*, *black*, *grey*).

4.6.2 Comparing collocate lists across the corpora

Collocate tables 22 and 24 equally display some of the adjectives mentioned by Biber and Conrad (2009) above, namely *mad*, *crazy* and *wrong*. All of them are ranked high in the lists with *wrong* being ranked first in the written register and *mad* in the spoken register (ex. 135, 136).

135. *Don't give up when things go wrong.*

136. *My brother did go mad once when we were driving a big old asphalt lorry [...]*

Table 22 represents the written register and compared with the spoken register, we can observe that the two lists differ to a great extent in terms of the adjectives they display. In AAM, we can see adjectives, such as *unnoticed*, *viral*, *hungry*, *extinct*, *unpunished*, *unanswered*, *green*, *unchallenged* and *smoke-free* which do not occur among the collocates of the spoken register. On the other hand, *mental*, *blind*, *bust*, *bald*, *grey*, *quiet*, *mouldy*, *black* and *soft* only occur among the collocates of the spoken register. Nevertheless, we are able to observe similarities, too.

First of all, the adjectives in both lists tend to relate to mental health or emotions (see ex. 137). Especially in the spoken register, the highest three positions are occupied by adjectives *mad*, *mental* and *crazy*. *Crazy* and *mad* appear in the written register too but the former is ranked

higher. Apparently, in the written register the verb collocates more significantly with *crazy*, while in the spoken register with *mad*.

137. *I am about to go crazy from sleep deprivation.*

Secondly, the adjectives relate to lifestyle, social media or public life in general. For example, in the written register there are adjectives such as *viral* (ranked third) or *online* which are related to the internet connection. Similarly, *bankrupt* or *green* may be seen in situations which happen in public but from a different point of view and with a different connotation. *Bankrupt* is perceived as negative, whereas *green* as positive because it is connected with environmental awareness. In the spoken register, *bankrupt* appears together with its synonym *bust*, however, *bankrupt* ranks higher and so it is more significant in terms of collocational strength. Also, there are adjectives which relate to appearance such as *bald* or *grey*, connected with hair (see ex. 138). On the contrary, there are no adjectives concerning appearance among the fifteen highest ranked adjectives of the written register.

138. *I said you say this now but imagine when I go grey and the all the hair's even thicker and wirier.*

4.6.3 Semantic categories of the adjectives

Biber and Conrad (2009) explain that the verb *go* together with the provided set of adjectives describes a change to a negative state. The adjectives are: *bad*, *deaf*, *mad*, *crazy*, *limp* and *wrong*. In the present thesis, it is also observed that the adjectives have mostly negative connotation, however, we attempted to divide the adjectives into more detailed categories. The constructions describe:

- a. a change to a negative state or behaviour,
- b. a change connected with public behaviour,
- c. an observable change in appearance,
- d. something that can be reversed.

In the written register, the first category is represented by the adjectives *wrong*, *crazy*, *mad* and *wild*. Consequently, *crazy* and *mad* can be further specified as negative mental states. The most adjectives the verb *go* takes are assigned to the second category: *viral*, *bankrupt*, *extinct*,

online, green and smoke-free. This category is connected closely or remotely with public behaviour. For example, *viral* and *online* with the internet, *green* and *smoke-free* with the environment. The fourth category is largely represented in the spoken register, but in the written register it only includes the adjective *hungry*. The fifth category comprises adjectives with the prefix *-un*: *unnoticed, unpunished, unanswered* and *unchallenged*.

In the spoken register, the first category has more representatives than in AAM: *mad, mental, crazy, wrong* and *wild*. Conversely, the second category is smaller than in the written register and comprises the adjectives *bankrupt, bust* and *online*. The former two are connected with a financial situation of the subject. The third category is the largest and it significantly outnumbers the number of adjectives in the written register. It is covered by the adjectives *blind, bald, grey, quiet, mouldy, black* and *soft*. The fifth category is not present in the spoken register at all.

Malá (2014) states that the verb *go* is less frequent than the other two resulting copulas across all registers. She provides a set of adjectives which is similar to the adjectives occurring in the spoken register (see chapter 2.2.2).

4.6.4 Adverbs occurring in the structure

In the written register, the adjectives of the verb *go* are not modified by adverbs to a great extent, as the corpus returned 629 hits only. Also, the corpus shows that among the hits there are cases of adverbs with a different function, as in *go back next year* or *go home sick*. Such adverbs and constructions were excluded from the analysis which left us with a minimum number of relevant adverbs.

Completely ranks first (ex. 139) and the following set of adverbs with even a lower number of occurrences follows: *horribly, too, so, very, largely, totally, really, terribly*. While some of the adverbs occurred in the previous chapters such as *too* or *so*, other adverbs occur among the highest ranked adverbs solely following this verb such as *terribly, horribly*. These two adverbs mean “extremely” but very often in a very bad way. Also, both *terribly* and *horribly* modify the adjective *wrong* (see ex. 140, 141) almost in all cases:

139. *It's not that I've gone completely silent!*

140. *My uncle had heart surgery and it went terribly wrong.*

141. *But when the procedure goes horribly wrong, Quaid becomes a hunted man.*

In the spoken register, the corpus returned 409 matches, however, many examples must have been excluded, similar to the written register. It has already been shown that the adverb *really* is most dominant in the spoken register when it occurs before an adjective complementing the copulas chosen for the analysis. With the verb *go* it is also the case. The other adverbs which appear relatively frequently are *absolutely*, *too*, *completely*, and *horribly*. It must be noted that *horribly* in all cases modifies the adjective *wrong* (ex. 142). Additionally, *absolutely* appears frequently with the adjectives *mental* and *ballistic* which both are connected with one's loss of self-control and anger (see ex. 143, 144).

142. [...] *you can see how it could go horribly wrong couldn't you?*

143. *So this morning she went absolutely mental [...]*

144. [...] *and when her fella found out he went absolutely ballistic [...]*

In the written register, the corpus returned 84 hits. If the adjectives of the verb *go* are modified by more than one adverb, the combination of adverbs usually presents negative connotation. *So* frequently co-occurs with *horribly*, *dreadfully*, *catastrophically* and *badly* which together modify the adjective *wrong* (ex. 145). Many of the hits the corpus returned must have been excluded because they included different types of adverbs, e.g. adverb of place (ex. 146).

145. *I don't understand why it has all gone so horribly wrong.*

146. *They can't go back home unchanged.*

In the spoken register, the combinations of adverbs are different. The corpus returned 113 matches, however, many of them were excluded for the same reasons as mentioned above. We can observe the adverb *really* gets repeated so it becomes *really really* (ex. 147) and even *really really really really* (ex. 148). Therefore, this adverb is not only dominant as a single adverb but also in the cases where more than one adverb modifies the adjective of the verb *go*. Also the adjective *wrong* tends to be modified by adverbs which themselves have negative connotation: *very badly*, *so badly* or *so horribly horribly* (ex. 149).

147. [...] *it's just gone really really quiet and it probably is the good weather.*

148. *So it goes really really really really thick until there's just like a really thick paste on the meat [...]*
149. *It all went so horribly horribly wrong over the years [...]*

4.7 The verb *get*

The verb *get* is classified as type become copular verb and, similarly as the previous two verbs, it expresses a change of a state. Biber and Conrad (2009) claim that this change tend to be to a negative state. As described in chapter 3.2., *get* is a verb of various meanings and functions. It was necessary to go through the results manually and exclude the irrelevant constructions. These included *get* as a transitive verb, *get*-passives and *get* in fixed expressions in which we are no longer able to see *get* as a copula followed by an adjective. It may be argued that there are other adjectives which constitute a fixed expression and still they are included in the analysis. The reason for why *get ready* remains in the analysis and *get rid* is excluded is that in the former case it is possible to see *get* as involving a change and may be replaced by *become*. Also, the elements which follow the combination vary. However, in the latter case, the preposition *of* follows the construction in all cases which makes it a fixed expression *to get rid of*. Especially in the lists obtained from the SBNC, there are cases of *get* as a transitive verb as in *we got good grades*. Another example of a construction which had to be excluded, even though we are able to see a change of a state, is *get used* which is followed by the preposition *to* and constitutes a single fixed expression *get used to*.

	Adjective	No. of occurrence	Instances per million
1	<i>ready</i>	992	8,312066
2	<i>married</i>	640	5,362623
3	<i>sick</i>	426	3,569496
4	<i>pregnant</i>	354	2,966201
5	<i>tired</i>	274	2,295873
6	<i>angry</i>	212	1,776369

7	<i>old</i>	176	1,474721
8	<i>rich</i>	163	1,365793
9	<i>excited</i>	144	1,20659
10	<i>real</i>	140	1,173074
11	<i>creative</i>	138	1,156316
12	<i>involved</i>	132	1,106041
13	<i>drunk</i>	130	1,089283
14	<i>mad</i>	129	1,080904
15	<i>lucky</i>	113	0,946838
Total	16 012 hits		

Table 25: AAM: a frequency list of *get*

	Adjective	LogDice measure
1	<i>ready</i>	10,23
2	<i>married</i>	10,18
3	<i>sick</i>	9,45
4	<i>pregnant</i>	9,27
5	<i>tired</i>	8,96
6	<i>angry</i>	8,53
7	<i>excited</i>	8,11
8	<i>involved</i>	8,03
9	<i>drunk</i>	7,96
10	<i>mad</i>	7,90
11	<i>rich</i>	7,74

12	<i>lucky</i>	7,58
13	<i>creative</i>	7,52
14	<i>wet</i>	7,496
15	<i>comfortable</i>	7,318

Table 26: AAM: a collocate list of *get*

	Adjective	No. of occurrence	Instances per million
1	<i>ready</i>	166	14,43478
2	<i>drunk</i>	133	11,56522
3	<i>stuck</i>	125	10,86957
4	<i>lost</i>	124	10,78261
5	<i>bored</i>	101	8,782609
6	<i>old</i>	86	7,478261
7	<i>involved</i>	63	5,478261
8	<i>sick</i>	60	5,217391
9	<i>tired</i>	54	4,695652
10	<i>cold</i>	53	4,608696
11	<i>annoyed</i>	52	4,521739
12	<i>big</i>	52	4,521739
13	<i>angry</i>	50	4,347826
14	<i>pissed</i>	49	4,26087
15	<i>pregnant</i>	48	4,173913
Total	4 753 matches		

Table 27: SBNC: a frequency list of *get*

	Adjective	Dice coefficient
1	<i>ready</i>	0,055
2	<i>stuck</i>	0,051
3	<i>lost</i>	0,05
4	<i>drunk</i>	0,049
5	<i>bored</i>	0,04
6	<i>involved</i>	0,025
7	<i>sick</i>	0,022
8	<i>annoyed</i>	0,021
9	<i>pissed</i>	0,02
10	<i>tired</i>	0,02
11	<i>angry</i>	0,019
12	<i>confused</i>	0,019
13	<i>pregnant</i>	0,019
14	<i>married</i>	0,018
15	<i>cold</i>	0,017

Table 28: SBNC: a collocate list of *get*

4.7.1 Comparing frequency and collocate lists within a single corpus

Table 25 shows the list of frequency in the AAM corpus, whereas Table 26 shows the list of collocates in the same corpus. If we compare the two lists we notice that the first six positions are occupied by the same adjectives which is rather rare when compared with the lists of other

verbs. *Ready* ranks first, *married* second, *sick* third, *pregnant* fourth, *tired* fifth and *angry* sixth in both lists. Examples of how these adjectives are used are provided below:

150. *A child's home environment can help them get ready to read.*

151. *They fell in love and got married.*

152. *But it didn't matter because my mom got sick and didn't end up accompanying us after all.*

153. *We didn't get pregnant after seven years into our marriage.*

154. *I like that because I get tired of eating out all the time.*

The frequency list differs from the collocate list in adjectives *old* and *real* which are ranked seventh and tenth, respectively. The collocate list differs from the frequency list in adjectives *wet* and *comfortable* which are ranked fourteenth and fifteenth, respectively.

Table 27 and Table 28 were obtained from SBNC as our source of the spoken language data. The two lists differ only moderately. The frequency list includes the adjectives *old* and *big* which are not prominent enough to be mentioned in the collocate list. On the other hand, *confused* and *married* are included only in the collocate list.

The same five adjectives (ex. 155-159) appear in the first five positions of both lists but the exact rankings of the adjectives differ. *Ready* ranks first, *drunk* second, *stuck* third, *lost* fourth and *bored* fifth in the frequency list. *Ready* ranks first also in the list of collocates but *stuck* ranks second, *lost* third, *drunk* fourth and *bored* fifth.

155. *They went out for a walk and I got ready for work [...]*

156. *Can we do that after cocktails so that I can get drunk and then not have to pay attention?*

157. *I guess it's probably an old fashioned you kind of get stuck in your habits don't you?*

158. *But often the trouble is they get lost in their own little reality.*

159. *I bet if she was living here he'd get bored of her very quickly.*

4.7.2 Comparing collocate lists across the corpora

When we compare Table 26 from AAM and Table 28 from SNBC, we can see that the two lists do not correspond, even though more than a half of the highest ranked adjectives are identical. However, in the written register, the distinct adjectives are *excited*, *mad*, *rich*, *lucky*, *creative*, *wet* and *comfortable*, whereas *stuck*, *lost*, *bored*, *annoyed*, *pissed*, *confused* and *cold* are to be found only in the spoken register.

In both lists, we can observe adjectives reflecting negative emotions often connected with anger. In the written register, *angry* and *mad* occur and the former adjective is ranked higher (ex. 160). On the other hand, in the spoken register, other adjectives rank higher than *angry*, namely *annoyed* (ex. 161) and *pissed*.

Biber and Conrad (2009) state that the verb expresses a change to a negative state. In opposition to their claim, in the written register, we are able to observe a positive emotion too, e.g. *excited* (ex. 162). In the spoken register, there is no such positive emotion involved. However, we are able to see the adjectives *pregnant* and *married* which also describe a change to a positive state rather than negative. These two adjectives appear in both lists, although only in the written register they are ranked in top positions. In conclusion, the collocates in the spoken register are in agreement with Biber and Conrad's claim. On the other hand, in the written register, the adjectives tend to have a positive connotation too. Therefore, the change to a positive state is also possible.

160. *This is perhaps a reason why he gets angry when he is rushed.*

161. *And they get annoyed with you when you do stupid things [...]*

162. *I am one of those people who gets excited about every aspect of food.*

4.7.3 Semantic categories of the adjectives

Biber and Conrad (2009) state that the verb *become* together with its adjectival complements describe a change to a negative state. In the present analysis, the majority of adjectives also describes a negative state. Additionally, the adjectives can be divided into negative physical and mental states. However, as opposed to Biber and Conrad's claim, some of the observed

adjectives can also have a positive connotation, i.e. they can describe a change to a positive state.

The fifteen adjectives of both written and spoken register can be divided in the following categories:

- a. a positive state,
- b. a negative physical state,
- c. a negative mental state.

In the written register, the first category is the largest. It comprises the positive adjectives: *ready, married, pregnant, excited, involved, rich, lucky, creative* and *comfortable*. There are still some semantic differences among the adjectives; however, they all share the property of being rather positively oriented. Very often they involve some activity which led to the positive state, e.g. *ready, married, rich*. The second category comprises the adjectives *sick, tired, drunk* and *wet* connected with a particular negative feeling which is possible to perceive physically. Lastly, the third category is represented by the adjectives *angry* and *mad*.

In the spoken register, there are fewer adjectives with positive connotation, therefore, the category is smaller than in the written register: *ready, involved, pregnant* and *married*. Conversely, the second category is the largest and it comprises the adjectives *stuck, lost, drunk, bored, sick, tired* and *cold*. The third category is covered by the adjectives *annoyed, pissed, angry* and *confused*.

4.7.4 Adverbs occurring in the structure

In the written register, the verb *get* is followed by the following set of adverbs which modify the adjectives complementing the verb: *too, so, very, really* and *pretty*. The five adverbs mentioned are the highest ranked adverbs of the list which are distributed in more than a half of the hits the corpus returned (2194 in total). *Too* ranks first in the list (ex. 163). The adverbs *too, so, very* and *really* are used for emphasis to express a high degree of the property the adverbs modify. *Pretty* is used, depending on the context, to express both a high degree and a moderately high degree (ex.164).

163. *I think Mom just got too soft in her later years of parenting.*

164. *Sometimes life can get pretty serious.*

In the spoken register, the corpus returned 1280 matches and *really* outnumbers the other adverbs significantly again (ex. 165). The other adverbs which appear quite frequently are *so*, *too* and *quite* (ex. 166). While *really*, *so* and *too* appear in the written register too, *quite* does not occur frequently there.

165. *I too Monday and Tuesday off last week so and then came back just because to be fair I get really bored at home.*

166. *We got quite drunk [...]*

In the written register, the corpus returned 107 hits. However, most of the instances must have been excluded because the verb did not function as a copular verb or the adverbs had different functions. The following combinations appeared several times: *so damn* (ex. 167) and the adverbs *too*, *very* and *really* which co-occurs with other adverbs or with itself (ex. 168).

167. *I apologize if this comes off harshly but I get so damn tired of that evo-psych bogeyman.*

168. *Sometimes I just get too too busy to respond right away, [...]*

However, in the spoken register, *get* appears more often with more than one adverb modifying the adjective (118 matches in total). Mostly it is the repetition of *really* which creates *really really* (ex. 169) or *really really really* (ex. 170) and once also *really really really really* (ex. 171). Less frequently we can observe *very very* (ex. 172).

169. *That's when it starts to get really really bad.*

170. *I've got a black belt in karate and he was getting really really really violent.*

171. *[...] I've been on the work away website and the help X website I'm getting really really really really excited.*

172. *It's a long way to drive and he gets very very tired [...]*

5 Conclusion

The present thesis focused on the description and analysis of the following seven copular verbs: *feel*, *look*, *seem*, *remain*, *become*, *go* and *get*. In the analytical part, copular verbs were searched in two corpora in order to examine their use with adjectival complements. The two corpora, i.e. Araneum Anglicum Minus and Spoken BNC2014, provided data from both written and spoken register, respectively. The aim of this descriptive work was to analyse the behaviour of copular verbs complemented by adjectives in the two registers.

In the analysis, the adjectives of the seven copular verbs were searched in both corpora and their frequency of occurrence and their collocability with the verbs were observed. While the frequency of occurrence of those adjectives was only compared with the list of collocates within one corpus to monitor both the most common combinations and the most statistically prominent, the collocate lists, based on comparable association measures, were then compared between the two corpora. Therefore, it was possible to see the differences between the two registers: written and spoken. Also, the adjectives occurring in the collocate lists were semantically categorized. As a sub-objective, the analysis also observed the occurrence of adverbs modifying the adjectives the verbs combine with.

It was expected that the frequency lists and collocate lists would differ which was confirmed in the analysis. Often, the same adjective appears in both lists, but the ranking is clearly different. For example, the adjective *terrible*, which occurs with the verb *look*, ranks quite high (fifth) among the collocates, but occupies only the fourteenth position in the frequency list, suggesting that it may not be the most common combination in terms of frequency, but it definitely forms a meaningful unit in terms of collocability. On the other hand, the adjective *good*, which often combines with the verb *feel*, ranks first in the frequency list as a very common combination, but only tenth in the collocate list. However, it can also happen (albeit more rarely) that the first ranked adjectives are the same for both lists of one register. For example, the adjective *aware*, which occurs with the verb *become*, ranks first in both lists in the written register suggesting a fairly stable preference of the verb for this adjective.

The collocate lists of each verb were then used to contrast the two registers. It was shown that the registers generally differ in the range of adjectives they use. For example, the two collocate lists of the verb *seem* share only four adjectives, the rest is unique for each register. Even though the particular adjectives mostly vary in the lists, they still share some semantic

properties that can be categorized. The adjectives were, therefore, divided into general semantic categories. The analysis showed that the verbs mostly share the same semantic categories in both registers. However, the prominence of the categories and their representatives often differed. The sets of adjectives in the categories were also partly compared with the sets of adjectives provided by the authors mentioned in the theoretical part (cf. 2.2).

Finally, the use of adverbs which modify the adjectives was also taken into account. As indicated above, the behaviour of the verbs differed in the registers, and so was the case with the adverbs. In the written register, the adverb *so* outnumbered the other adverbs with the verbs *feel*, *look* and *seem*. The other four verbs display the use of various adverbs in first positions. In the spoken register, the adverb *really* occurs frequently with most of the verbs except for the verb *remain* which does not have its adjectival complements modified by adverbs. Regarding the occurrence of more than one adverb modifying the adjectives, the combinations were various. Especially in the spoken register, there was a tendency of adverbs to repeat to create combinations such as *really really* or *very very*. In the written register, the combinations often include the adverb *so* in first place, as in *so very*, *so horribly*, *so damn*. It was also shown that, in the spoken register, the adjectives are very often modified by a single adverb but the adverb frequently occurs in combination with *like* which has its specific function for spoken register.

To conclude, we hope that the present thesis has provided a corpus-based description and a systematic analysis of the selected copular verbs and their adjectival complements. The results of the analysis as well as the description in the theoretical part may be used for further didactic purposes, principally when the focus is on the differences in the use of adjectives between the written and spoken register.

6 Bibliography

Benko, Vladimír (2014). *Aranea: Yet Another Family of (Comparable) Web Corpora*. In Petr Sojka, Aleš Horák, Ivan Kopeček and Karel Pala (Eds.): *Text, Speech and Dialogue. 17th International Conference, TSD 2014, Brno, Czech Republic, September 8-12, 2014. Proceedings. LNCS 8655*. Springer International Publishing Switzerland. Available at: <https://kontext.korpus.cz/> (last accessed: 01. 08. 2019)

Benko, V. (2015). *Aranea- comparable web corpora*. Praha: Ústav Českého národního korpusu FF UK. Available at: <http://www.korpus.cz> (last accessed: 01. 08. 2019)

Biber, D. et al. (1999). *The Longman Grammar of Spoken and Written English*. Harlow: Longman.

“go” in Cambridge Dictionary. Cambridge: CUP. Available at: <https://dictionary.cambridge.org/dictionary/english/go> (last accessed: 06. 07. 2019)

Conrad, Susan., Biber, Douglas (2009). *Real Grammar: Corpus-Based Approach to English*. London: Longman.

Dušková, Libuše. (2012). *Mluvnice současné angličtiny na pozadí češtiny*. Praha: Academia.

Downing, A., Locke, P. (2006). *A university course in English grammar* (2nd ed.). Oxon: Routledge.

Huddleston, R., G. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Love, R., Dembry, C., Hardie, A., Brezina, V. and McEnery, T. (2017). *The Spoken BNC2014: designing and building a spoken corpus of everyday conversations*. In *International Journal of Corpus Linguistics*, 22(3), pp. 319-344.

Malá, M. (2014). *English Copular Verbs. A contrastive corpus-supported view*. Praha: FF UK.

Pustet, R. (2005). *Copulas. Universals in the Categorization of the Lexicon*. Oxford: OUP.

Quirk, R. et al. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

7 Resumé

Tématem této diplomové práce je analýza vybraných sponových sloves a jejich adjektivních doplnění v současné psané a mluvené angličtině. Úvodní kapitola poskytuje vhled jak do obsahu teoretické části, tak představení části praktické. Stručně nastíněna je metoda výzkumu i materiál, ze kterého práce čerpá data. Zároveň uvádíme základní popis sponových sloves. Sponová slovesa jsou vymezena jako kategorie synsémantických sloves, která ve větě pouze připisují vlastnosti subjektu. Tyto vlastnosti jsou poté vyjádřeny nutným doplněním, které má nejčastěji formu substantiva nebo adjektiva. Cílem práce je popsat chování sponových sloves *feel, look, seem, remain, become, go* a *get* se zaměřením na jejich adjektivní doplnění v současné psané a mluvené angličtině.

Teoretická část je rozdělena na několik kapitol, z nichž každá se věnuje popisu sponových sloves z určitého pohledu. Nejdříve je nutné vysvětlit terminologii spojenou se sponovými slovesy. To se týká zejména termínu pro doplnění sponových sloves. V českých mluvnicích se využívá termínu *jmenná část přísudku*, který se do angličtiny překládá jako *subject complemen*, a který je obligatorním doplněním sponových sloves, tj. sponové sloveso nemůže stát samo o sobě. Jmenná část přísudku se odlišuje od *doplňku podmětu* (*verbless adverbial clause*), který je popisován jako volitelné doplnění plnovýznamových sloves. Ukazuje se, že autoři mluvnic se neshodují v tom, co lze považovat za sponové sloveso a co nikoliv. Dva protichůdné přístupy zaujímají Dušková a kol. (2012) a Quirk a kol. (1985). Dušková a kol. tvrdí, že sloveso, jehož doplnění má formu adverbia, již není sponové sloveso. Naopak Quirk a kol. zastávají názor, že i adverbium může tvořit doplnění sponových sloves.

Další z kapitol se věnuje popisu slovesa *be*, které je chápáno jako základní sponové sloveso, na kterém je možno jevy spojené s tímto druhem predikace vysvětlit. Následující kapitola popisuje rozdělení sponových sloves. Sponová slovesa lze rozdělit na typ *be*, který se dále dělí na další tři podtypy, a typ *become*. Zatímco první typ sloves popisuje současný stav a lze ho dále dělit na slovesa smyslového vnímání, slovesa popisující stupeň faktivity sdělovaného obsahu a slovesa popisující trvání příslušného příznaku, druhý typ sloves popisuje změnu stavu. Následně je tento typ predikace popisován ze sémantického pohledu, kde je opět možno vydělit několik typů. Jedná se o typ predikace kvalifikující, identifikující, klasifikující, posesivní a existenciální či existenciálně-lokativní. Předposlední část se věnuje popisu sponové predikace z pohledu možností jejího doplnění. Tato část se zaměřuje především na sponová slovesa jiná než *be*, jelikož ta jsou hlavním předmětem zkoumání v části analýzy.

Poslední kapitola je věnována korpusovému přístupu ke sponovým slovesům se zaměřením na adjektivní doplnění. Zde je popsáno, jaká adjektiva se vyskytují s danými slovesy nejčastěji na základě vybraných korpusových studií.

Třetí část této práce popisuje metodu získání materiálu a materiál samotný. V praktické části je zkoumáno sedm sloves *feel*, *look*, *seem*, *remain*, *become*, *go*, *get* se zaměřením na jejich adjektivní doplnění. Zdrojem dat pro každé z těchto sloves byly dva korpusy, jeden pro angličtinu psanou (Araneum Anglicum Minus korpus) a jeden pro angličtinu mluvenou (Spoken BNC2014). Kromě popisu korpusů je ve třetí kapitole podrobně popsán postup, který zahrnuje následující kroky. Všechna výše zmíněná slovesa byla nejprve vyhledána v obou korpusech a následně se jejich adjektivní doplnění zkoumala a porovnávala. Pro každé sloveso byly vytvořeny čtyři seznamy, tj. frekvenční a kolokační seznam reflektující psaný jazyk a stejné dva seznamy reflektující mluvený jazyk, obsahující patnáct nejvýše postavených adjektiv, které dohromady pokrývaly alespoň polovinu výskytů dané konstrukce v korpusu. Pro vytvoření kolokačních seznamů byly využity asociační míry logDice a Dice coefficient, které jsou založeny na podobném výpočtu a jsou tedy mezi sebou zhruba porovnatelné. Nejdříve byly porovnány frekvenční seznamy adjektiv s kolokačními uvnitř jednoho korpusu. Následně jsme kolokační seznamy porovnávali mezi oběma korpusy, tudíž bylo možné sledovat rozdíly mezi psanou a mluvenou angličtinou. Adjektiva vyskytující se jako doplnění daných sloves se poté rozřadila do několika sémantických skupin. Jako dílčí cíl byl zkoumán i výskyt adverbii modifikujících adjektivní doplnění.

Kapitola 4 věnovaná samotné analýze, se dělí do sedmi podkapitol podle jednotlivých sloves, z nichž se ještě každá skládá z dílčích částí analýzy (porovnání frekvenčních a kolokačních seznamů, porovnání kolokačních seznamů v psané a mluvené angličtině, sémantická kategorizace adjektiv a použití adverbii ve zkoumaných spojeních). Cílem analýzy bylo popsat, jaká adjektiva se s danými slovesy vyskytují, zda se liší v psaném a mluveném projevu a zda je možné je dělit do kategorií podle jejich významu.

Analýza ukázala, že frekvenční seznamy a kolokační seznamy se z velké většiny liší a mnoho adjektiv se vyskytuje z pohledu frekvence často, ale nejsou silnými kolokáty. Pro ilustraci je možné se podívat výsledky výskytů adjektiv na prvních místech frekvenčních a kolokačních seznamů slovesa *feel*. V psaném projevu se sloveso *feel* na základě frekvence nejčastěji pojí s adjektivem *good*, ale nejsilnějším kolokátem je *comfortable*. V mluveném projevu se frekventovaně vyskytuje na prvním místě s adjektivem *bad*, kdežto *sick* je statisticky

nejsilnější kolokát. Byly však pozorovány i případy, kdy se slova ve frekvenčním a kolokačním seznamu na předních příčkách shodovala, viz *aware* u slovesa *become* v psaném projevu.

Pro porovnání mezi korpusy byly využity kolokační seznamy, získané pomocí dvou srovnatelných měr. Očekávalo se, že adjektiva v psaném a mluveném projevu se liší, což analýza potvrdila. Většina adjektiv vyskytujících se na prvních patnácti místech v jednom korpusu se nevyskytovala ve druhém a naopak. Například sloveso *feel* se v psaném projevu silně pojí s následujícími třemi adjektivy *comfortable*, *guilty*, *confident*. V mluveném projevu první tři příčky obsazují tato adjektiva *sick*, *guilty*, *comfortable*. Sloveso *look* se v psaném projevu silně pojí s následujícími třemi adjektivy *delicious*, *amazing*, *fabulous*. V mluveném projevu první tři příčky obsazují tato adjektiva *nice*, *amazing*, *lovely*. Sloveso *seem* se v psaném projevu silně pojí s následujícími třemi adjektivy *strange*, *unlikely*, *obvious*. V mluveném projevu první tři příčky obsazují tato adjektiva *unlikely*, *logical* *pointless*. Sloveso *remain* se v psaném projevu silně pojí s následujícími třemi adjektivy *unchanged*, *silent*, *intact*. V mluveném projevu první tři příčky obsazují tato adjektiva *nameless*, *unchanged*, *faithful*. Sloveso *become* se v psaném projevu silně pojí s následujícími třemi adjektivy *aware*, *apparent*, *clear*. V mluveném projevu první tři příčky obsazují tato adjektiva *apparent*, *cocky*, *famous*. Sloveso *go* se v psaném projevu silně pojí s následujícími třemi adjektivy *wrong*, *unnoticed*, *viral*. V mluveném projevu první tři příčky obsazují tato adjektiva *mad*, *mental*, *crazy*. Sloveso *get* se v psaném projevu silně pojí s následujícími třemi adjektivy *ready*, *married*, *sick*. V mluveném projevu první tři příčky obsazují tato adjektiva *ready*, *stuck*, *lost*.

Součástí analýzy bylo i vytvoření všeobecných kategorií na základě významu pro každé sloveso, do kterých by bylo možné adjektiva zařadit. Ukázalo se, že kategorie se příliš neliší pro psanou a mluvenou angličtinu, ale že velikost obsazení kategorií a konkrétní adjektiva jsou jiná. Adjektiva některých sloves bylo snadné rozdělit do tří kategorií (např. adjektiva slovesa *look*), u jiných sloves bylo zapotřebí kategorií více (např. sloveso *seem*). V této části byla adjektiva taktéž porovnána se seznamy adjektiv poskytnutými autory v teoretické části věnované korpusovému přístupu k sponovým slovesům.

Podkapitoly analýzy věnující se adverbiiím ve funkci modifikátorů adjektiv odhalily, že mezi psaným a mluveným projevem jsou rozdíly v užití konkrétních adverbii. V mluveném projevu adverbium *really* frekvenčně výrazně převyšuje jiná adverbia u všech sloves kromě *remain*,

které se ale i s jinými adverbii vyskytuje jen ve dvou případech. U slovesa *become* je *really* nejfrekventovanější, avšak adverbia *quite*, *so* a *very* se také objevují velmi často. V psaném projevu je velmi časté adverbium *so*, které frekvencí převyšuje jiná adverbia u sloves *feel*, *look*, *seem*. U dalších sloves je ale taktéž velmi časté. Adjektiva slovesa *remain* se v mluveném projevu s adverbii takřka nevyskytují, avšak v psaném projevu je tomu naopak, přičemž nejčastější je adverbium *relatively*. Adjektiva slovesa *become* jsou nejčastěji modifikována adverbiem *increasingly*, u slovesa *go* je to adverbium *completely* a u slovesa *get* je to adverbium *too*. Výskyt několika adverbií za sebou je méně častý než výskyt samostatných adverbií. *Really* má tendenci se nejčastěji kombinovat samo se sebou, tudíž konstrukce jako *really really* jsou velmi časté. V souvislosti s tím se ukázalo, že adverbia vyskytující se v mluveném projevu mají obecně tendenci se opakovat a další konstrukce jsou například *too too*, *very very*, *so so*. V psaném projevu jsou nejčastější kombinace s adverbiem *so* na prvním místě, např. *so very*, *so damn*, *so horribly*.

Výsledky všech kapitol analýzy jsou poté stručně shrnuty v závěru v takovém pořadí, jak šly kapitoly po sobě i v analýze.

Po závěru následuje kapitola věnovaná seznamu literatury, ze které bylo čerpáno v teoretické části. Seznam literatury zahrnuje jak základní mluvnice angličtiny, tak zdroje dat a studie na dané téma.