Prepositions used with adjectives in English essays written by Czech secondary school students

Věra Sládková
Masaryk University in Brno

This article focuses on the frequency and accuracy of dependent prepositions which complement the adjectives in CZEMATELC 2017, a corpus consisting of 390 essays from the written part of the national school-leaving exam leading to certification of secondary education in the Czech Republic. The research findings reveal that the learners used adjectives from A1 to B2 level, according to the CEFR. A limited number of A1 adjective lemmas was considerably overused, but showed the lowest proportion of dependent prepositional complementation. As learners tended not to complement the adjectives at A2 – B2 proficiency levels either, adjective-preposition collocations frequently co-occurring in native speaker corpora were identified for further remedial work. In addition, corpus-based discovery-learning was proposed as a solution because it encourages awareness and gradually leads to learner autonomy.

   Key words: adjectives, dependent preposition, prepositional phrase, collocation, data-driven learning

1. Introduction

When teaching prepositions, English teachers should be aware of the influence of L1 on the pedagogical outcomes and, at the same time, be familiar with a range of strategies to diminish it. One way to deal with this problem, which would also address the issue regarding the polysemous nature of prepositions, is to teach them as collocations in combination with co-occurring words. The article attempts to address the issue of selecting adjective-preposition collocations for teaching by investigating a learner corpus and to highlight discovery-learning activities based on native speaker as well as learner corpora.

2. Prepositions co-occurring with adjectives

2.1 Dependent prepositions

Prepositions are "linking words that introduce prepositional phrases" (Biber, Conrad, & Leech, 2002, p. 28) and specify the relationship between two or more entities that they link, or express various other abstract relations. With the exception of stranded prepositions, they are inseparable from their complement, which can be a noun phrase, a gerund or an adverbial clause (Carter & McCarthy, 2006; Leech & Svartvik, 1993). Although many linguists find it hard to agree whether to categorise them as functional or lexical words, Biber, Johansson, Leech, Conrad and Finegan claim that prepositions

have "the ambiguous status of having borderline lexical membership while at the same time qualifying as functional words" (1999, p. 74) and argue that they can have both free and bound meanings. In English, bound prepositions can complement a verb, an adjective, a noun and an adjunct in a clause. The choice of the particular preposition is determined by the word the prepositional phrase complements. Bound prepositions contribute very little or no meaning, so they are largely called dependent prepositions in pedagogical grammar. They can act as one language unit with the preceding word, thus creating chunks of language with a high probability of co-occurrence.

2.2 Adjective complementation and CEFR levels

The adjective which determines what preposition must follow acts as subject predicative complementing a copular verb. Apart from a prepositional phrase, it can also be complemented with *to-infinitive*, or a *that-clause*. According to the *English Grammar Profile* (EGP), an online tool based on continuous research carried out on the *Cambridge Learner Corpus* (CLC), which provides detailed information about which language forms used with a particular meaning typically appear in learners' production tasks at a particular language level, both of these types of complementation are expected to be found in learners' production at B1 level, which is the required level for successful completion of the school-leaving exam in English in the Czech Republic. The EGP, however, does not offer a specific "can do statement" concerning prepositional phrases for this level, but it requires A2 learners to be able to form a very limited range of prepositional phrases and use them to complement adjectives.

2.3 Previous research

Previous research concerning Czech speakers and their use of adjective-preposition collocations is limited to Dušková's (1969) error analysis of texts written by Czech post-graduate students and Sparling's (1990) reference book aimed at helping Czech speakers of English to avoid typical errors.

In the international context, a large body of research has been aimed at collocations, but adjective-preposition collocations have been investigated mainly as collocational errors. They were found to be the second most problematic collocations in the oral production of Iranian learners by Sadeghi and Panahifar (2013). Other studies (Jafarpour & Koosha, 2006; Kulsitthiboon & Pongpairoj, 2018) compare various ways to teach them to data-driven learning.

2.4 Three approaches to teaching prepositions

The Prototype Approach (Lindstromberg, 1996) to teaching prepositions, which requires spatial or "prototypical" meaning as the starting point, recommends teachers to

look for examples of the most dominant meaning and to teach it first before showing the learners the figurative (i.e. psychologically related) meaning in other phrases with the same preposition. This approach favours the use of pictures and diagrams and claims to allow deeper learning than the so-called Traditional Approach (Lorincz & Gordon, 2012), which requires learners to focus on prepositions individually within a particular context and to create long lists to be learnt by heart. The shift from abstract definitions, which can be very difficult to comprehend, to concrete examples is apparent in the Collocation Approach (Sinclair, 1991). It encourages learners to pay attention to "the company [the prepositions] keep" (Kennedy, 1991, p. 215), i.e. looking for sequences of patterns containing prepositions, noticing recurring combinations, and learning prepositions in connection with words with a high probability of co-occurrence. This approach is based on the collocational principle (Sinclair, 1991) according to which people tend to process formulaic sequences of words as a single unit. Repeated exposure to these sequences is likely to facilitate learners' acquisition as Mueller's (2011) research has shown.

3. Corpora and language teaching

The concept of using corpora in language learning was developed by Johns (1991), who proposed using concordance printouts to stimulate discovery learning by observing similarities and differences in authentic language samples taken from corpora, creating hypotheses and testing them. In his view, the learner assumes the role of a researcher and the teacher becomes more a facilitator of the learning process than the provider of language input. Instead of top-down processing whereby the learners are given the rules in a rather passive way and are required to apply them when using the language, data-driven learning (DDL) as this concept is also known, requires bottom-up processing of examples in context in order to formulate conscious, or even unconscious, generalisations concerning patterns of structure and meaning. This explicit approach to learning requires an active attitude from the learners and is cognitively demanding because the learners are presented with linguistic data and have to recognise patterns and regularities in the language use. Gabrielatos acknowledges that DDL can be compatible with various methodological approaches "that accept explicit focus on language structure and use" (2005, p. 25) and favour noticing and awareness-raising activities. Moreover, corpora-based discovery-learning can be exploited in different phases of a lesson, such as during presentation, revision and feedback stages, and in preparation for skill-based activities or during them. According to Gabrielatos, it can be incorporated in a wide spectrum of lessons whose aims can range from "totally teacher-centred to totally learner-centred" (2005, p. 12). Tan (2000) proposes *Investigative-oriented learning* (IOL) in which corpus-based work is integrated within the analytical stage of task-based learning during which the teacher usually highlights the language features that have been or should have been used during the task stage. She

distinguishes three skills (noticing, hypothesising and experimenting) that learners develop with the help of the corpora before proceeding with the last stage of a task-based activity.

Using corpora alongside course books can help teachers overcome a major problem they frequently face when they cannot find enough examples of language features they want to focus on. This is in part due to the fact that the latest course books tend to rely mostly on authentic texts which contain a natural density of language phenomena. Although corpora cannot replace out-of-class extensive reading, Gabrielatos claims that they "can offer condensed exposure to language patterns" (2005, p. 11) with the advantage of both extensive and intensive reading because the learners can observe a particular language feature taken from a large number of texts and at the same time concentrate on it. The learners have to be guided by the teacher (Bennett, 2010; Gabrielatos, 2005) until they acquire the necessary noticing skills which enable them to recognise patterns independently. Besides, language proficiency also needs to be taken into consideration when deciding how much guidance is needed. Gabrielatos (2005) holds that corpus-designed activities should direct learners away from the tendency to discover single correct answers and fixed rules and towards noticing alternatives and their contingency.

Corpus-designed activities are generally divided into hands-on and hands-off activities (Boulton, 2012). Computer-based hands-on activities, also known as hard version (Leech, 1997), require learners to have direct access to a corpus, whereas hands-off activities, or soft version, require the teacher to explore the corpus and create a set of activities, usually in paper form, for analysis in the classroom. These teacher-prepared activities may be more suitable for learners at low levels and those without any experience with DDL because the teachers can select sentences at the right level of difficulty and adapt them for a particular purpose (Gabrielatos, 2005). Many researchers (Ackerley, 2017; Bardovi-Harlig, Mossman & Su, 2017; Boulton & Cobb, 2017) are currently trying to obtain experimental evidence in order to compare the effectiveness of both approaches for particular groups of learners or language features.

Several studies (Barabadi & Khajavi, 2017; Boulton & Cobb, 2017) have found DDL to be more effective than using skills-based communicative approaches, probably because DDL is based on form-focused instruction, visual input, and repeated exposure to language features and expects active cognitive involvement from learners, which should gradually lead to their greater independence. The improved attitudes of students to learning after DDL was evidenced by Huei Lin (2016), who also noticed that non-native teachers felt more empowered by the use of corpora because they themselves developed a greater awareness of the language. Jafarpour and Koosha (2006) compared two approaches to teaching prepositions and their collocational patterns and found that a DDL approach based on concordancing outperformed conventional teaching. Vyatkina (2016) compared the hands-off DDL approach to traditional instruction when teaching low-intermediate L1 English learners of German and found it to be more effective for

learning new verb-preposition collocations, but equally effective for improving the knowledge of previously learned collocations.

Corpora in teaching writing have been largely associated with learner corpora, which enable a more precise description of learner language used in written communication in order to identify areas that need special attention in teaching. Comparative studies have focused mainly on the overuse and underuse of specific features of interlanguage in comparison to the language of native speakers, whereas error-analyses have tended to identify problematic language choices in the finished texts (i.e. Chuang & Nesi, 2006; Hinkel, 2005; Jaworska, Krummes, & Ensslin, 2015; Lee & Chen, 2009). However, recently, corpora have been used by learners to discover patterns when preparing for writing or during the writing itself, as well as for self-correction and remedial work after writing. These two different approaches to using corpora to improve writing have been termed pattern hunting and pattern refining by Kennedy and Miceli (2017), who present an account of a successful attempt to equip learners with the skills to both observe a corpus hands-on and to borrow chunks of language in order to enrich their writing and improve its accuracy. Tono, Satake and Miura (2014) reveal that learners are more likely to correct omission and addition errors than misformation errors when consulting a corpus.

4. Methodology

4.1 Research aim

If we want to improve teaching in order to facilitate the development of production abilities, it is necessary to identify the areas that require special attention and suggest teaching strategies which have been found to be effective through research. Gathering authentic samples of learner language from a particular exam situation and subjecting them to thorough analysis by means of corpus analytical tools is one way to do this. Moreover, learner data could be further exploited to create teaching materials and remedial activities informed by native speaker corpora in order to gradually prepare learners for corpus-based discovery-learning. The aims of this study are therefore as follows:

(1) To determine to what extent Czech secondary school students attempted to exploit the regularity of adjective-preposition co-occurrence in their school-leaving exam essays written in 2017.

(2) To identify the accurate and inaccurate uses of prepositions complementing adjectives.

(3) To ascertain if there is any relationship between the frequency and/or accuracy of adjective-preposition collocations in the studied essays and the CEFR level of the adjectives used.

(4) To propose how the results of the research and the data from the learner corpus CZEMATELC 2017 could be used to create remedial activities.

4.2 Research design

Before attempting to describe adjective-preposition collocations in learner language, it is worth clarifying that this specific lexico-grammatical feature is approached in the sense outlined by Halliday (1992), who views grammar and lexis as the notional ends of a lexicogrammatical continuum. It is analysed from the perspective of Pattern Grammar (Hunston & Francis, 2000), which allows grammar to be the starting point of the analysis, although lexis is its main focus. As the learner language is expected to be much more variable than native speaker language, it is important to avoid comparative fallacy, i.e. failure to acknowledge the unique features of interlanguage. Consequently, frequency analysis, which can also draw attention to language feature avoidance (Ellis & Barkhuizen 2005, p. 98), was chosen as the main research method. The identification of adjectives followed by dependent prepositions was carried out by means of a freely available online corpus analytical tool, AntConc 3.4.4w (Anthony, 2014). The manual frequency counts had to be accompanied by detailed qualitative analysis of the context because learner language is full of inconsistencies and therefore has to be checked to see if a particular form is used with the appropriate meaning. This mixed research paradigm is best interpreted in relation to an external model which can make the frequency data meaningful. As a result, the relationship between the frequency of correctly and incorrectly used dependent prepositions and the CEFR level of adjectives was ascertained by a simple statistical comparison of the percentages of adjectives assigned to individual CEFR levels according to the *English Vocabulary Profile* (EVP). This online tool was chosen despite the fact that CLC, a 50 million-word corpus on which EVP is based contains a relatively small share of language samples from native speakers of Slavonic languages, including Czech speakers of English (Proudfoot, 2010), hence it is questionable if the specified linguistic forms used to express meaning aligned to each CEFR level apply also to Czech speakers of English. However, Salamoura and Saville claim that the large amount of language samples across all major language families allows extensive research of the involvement of the mother tongue in "a learner's linguistic profile [and] cross-linguistic differences per CEFR level is one of the main premises under investigation" (2010, p. 109).

4.3 Context and participants

The essays were written by final-year students in Czech upper-secondary education (i.e. aged 19 and above), the majority of whom had studied English for 11 years. Czech was most likely their first language. In most cases, English would have been their L2, but it could also have been their L3 or L4, which would imply considerably less time spent on English language instruction (in extreme cases only 4 years). As well as English, the learners might have been also learning German, French, Spanish or Russian. However, the detailed information about the learners is unavailable for the legal reasons as confidentiality has to be strictly observed in the case of a high-stake exam.

4.4 Learner corpus

The CZEMATELC (Czech Maturita Exam Learner Corpus) 2017 consists of 390 essays which were written in May 2017 by 195 students and obtained by means of consent from the Centre of Educational Assessment (CERMAT). It is a random sample of essays because the sets of essays from each school were chosen randomly by a computer at CERMAT and allotted to individual assessors, one of whom is the author of this study. The analysed essays represent 0.455% of all essays based on the same assignment and written at the same time within 60 minutes. The raw corpus contains 44,044 tokens and 2,765 types.

The corpus consists of two different types of essays: a longer one (120–150 words) and a shorter one (60–70 words). The longer one was a story about an unexpected visitor and the shorter one required the students to ask a friend to lend them a bicycle. In both of them, the students were prompted in Czech about what to include in each paragraph. The students were allowed to use Czech-English or English-English dictionaries which can contain appendices with grammar explanations, but dictionaries with "essential descriptions of particular text types" (Centrum pro zjišťování výsledků ve vzdělávání (Centre of Educational Assessment), 2017, p. 1–2) were not allowed.

4.5 Procedure

Firstly, an alphabetical list of all types was created by means of AntConc 3.4.4w (Anthony, 2014). The types that could be identified as correctly spelt adjectives and the types that looked similar to adjectives (e.g. *affraid*) or correctly or incorrectly spelt adverbs were viewed in concordance lines to see if they occupied the attributive or predicative positions typical of adjectives. Those words whose form and/or function could be attributed to adjectives were categorised according to their position in the sentence and the way in which they were complemented (i.e. prepositional phrase, *that-clause*, *to-infinitive*, adverbial clauses, no complementation). The aim was to find out the frequency with which the identified adjectives were used in attributive and predicative positions, select those that frequently co-occur with prepositions if they

occupy predicative positions, and to pinpoint the successful and unsuccessful attempts to complement them with prepositional phrases.

For this reason, the accuracy of dependent prepositions was judged with the help of the *Cambridge Advanced Learner's Dictionary* (2008). In cases where the dependent preposition or the type of complementation used in CZEMATELC 2017 were not found in the dictionary, a relatively large reference corpus, the *Brown Family (C8 tags)* consisting of 5,748,130 tokens and 121,888 types, was used. It includes texts of a wide range of genres both from American and British English: the *Brown Corpus* (texts from 1961), the *Frown Corpus* (texts from 1992), the *LOB* (texts from 1961) and the *FLOB* (texts from 1991). If any instances of the investigated phenomena with the appropriate meaning were found in the reference corpus, they were considered accurate regardless of the frequency of their use. As the reference corpus did not contain current language, a university educated native speaker of British English (M.A. TESOL, DELTA) was consulted in case of any doubt and when no correspondence was found.

Finally, the adjectives found in CZEMATELC 2017 that tend to be complemented with prepositional phrases were assigned to the CEFR levels at which they are commonly used in productive tasks by learners using the *English Vocabulary Profile* and the relationship between the frequency and accuracy of dependent prepositions used with them at individual levels was investigated.

4.6 Limitations

Several factors could have influenced the results of the research. Firstly, the corpus is relatively small, which could raise doubts concerning balance and representativeness (cf. McEnery et al., 2006). Balance was achieved by including an equal number of essays based on the same prompts. This ensures close comparability and reduces the importance of the need for a large corpus. Nevertheless, analysing essays based on the same assignment has a tremendous impact on the results because the task restricts the range of language features used and considerably influences the frequency of their use.

Secondly, the corpus represents merely examples taken from an exam situation, which may render the examples unrepresentative of Czech secondary school students' performance as a whole. However, analysing language samples from an exam situation will certainly help to inform the teaching in preparation for the exam.

Thirdly, the research does not attempt to describe the students' production purely in terms of accuracy. It focuses only on adjectives and the correct use of dependent prepositions that follow them regardless of the accuracy of the rest of the sentences.

Finally, the results can also be influenced by the analytical tool and the method used. Frequency analysis does not provide any explanations and reveals only the frequency data concerning the features the researcher decides to count, so a reliable analytical tool is very important. For this reason, AntConc 3.4.4w (Anthony, 2014) was

chosen because it can analyse a raw corpus, and the necessary manual analysis is relatively fast. However, human error should be taken into consideration.

5. Results and discussion

Based on the *Cambridge Advanced Learner's Dictionary* (2008), CZEMATELC 2017 contains 50 lemmas of adjectives within the A1 – B2 language proficiency levels on the CEFR that could potentially be complemented by prepositional phrases. Two of them *allergic* and *enraged* could not be assigned to a particular CEFR level, hence they were included within adjectives at B2 level for statistical purposes and one of them *(easy)* was complemented by a preposition in such a way that together they created an idiomatic expression which was out of the scope of this research. Therefore it was not analysed further. Within the studied adjectives, there are also 12 lemmas with some kind of spelling inconsistency, such as **tipical*, **alergic*, **affraid*, **carefull, *suprised* (n=22), **supprised*, **exietet*. Several adverbs, for example *badly* (n=2), *carefully* (n=1) and *gratefully* (n=1), carried the same functions in the sentences as adjectives, so they were included in the analysis as well as all the inconsistently spelt adjectives. The learners also confused adjectives with suffixes –ed/-ing, for example: *exciting* (n=1) and**suprising* (n=1), as well as the meaning of the following adjectives: *afraid* vs *worried* and *scared* vs *scary*.

   The raw frequency of the studied adjectives was 705. However, the learners attempted to complement them with prepositional phrases only in 53 cases. This means that only 7.5 percent of all the analysed adjectives were complemented by either a correct or incorrect preposition, or it was clear that a preposition was omitted. Moreover, the learners used a correct dependent preposition in only 34 instances. The data in Table 1 show the frequency information for the individual CEFR levels.

Table 1 Comparison of raw frequency of analysed adjectives at different CEFR Levels

CEFR level	Raw frequency	Lemmas	Percentage of attempts	Number of attempts	Correct prepositions	Success rate
A1	482	12	5.8%	28	19	68%
A2	115	13	14%	16	10	62%
B1	97	18	6.2%	6	4	67%
B2	11	4+2	27.3%	3	1	33%
Total	705	49	7.5%	53	34	62%

Source: CZEMATELC 2017

As Table 1 shows, the learners appear to have been reluctant to complement adjectives with prepositional phrases with dependent prepositions in order to create more complex syntactic structures. This could be attributed to the typological proximity between Czech and English because the learners' mother tongue can only offer an exact equivalent in exceptional cases for the simple reason that most English adjectives

followed by dependent prepositions would be translated using different structural patterns. When relating the lemmas to language proficiency levels it would seem that the learners tended to overuse a limited number of A1 lemmas and rarely attempted to complement these adjectives with prepositional phrases (only in 5.8 percent of instances). However, the success rate is comparatively high at 68 percent. This may imply that the learners relied on a limited number of well-known adjectives which they encountered at the beginning of their English language instruction in simple syntactic structures, but they might not have acquired them adequately in combination with prepositional phrases later. This assumption is based on Philip (2007) who complains that teaching of collocations and new meanings of items that have already been learnt tends to be neglected.

With the growing level of difficulty, the raw frequency of adjectives at individual CEFR levels decreased. However, the number of lemmas grew, with the exception of B2 level. This group also showed the highest proportion of attempts to complement adjectives with prepositional phrases, but the success rate in terms of the correct use of prepositions is rather low at 33 percent. One possible interpretation could be that the learners were reluctant to take risks in an exam situation and therefore avoided using vocabulary on the margins of their language proficiency or beyond. The average success rate of 62 percent for all levels might also imply that the learners opted to complement adjectives with prepositional phrases only when they were relatively certain.

Table 2 reveals that the learners tended to use adjectives predominantly in predicative positions without complementing them with prepositional phrases. A1 adjectives appear to be exceptional because the difference between the numbers of adjectives used in attributive (n=210) and predicative positions (n=236) was relatively small. The data also reveal that the number of inappropriately used prepositions is so low that it is very difficult to make a general observation about error patterns, although the influence of the mother-tongue may lie behind the incorrect use of the preposition *on* and several omissions.

Table 2 Adjectives in attributive and predicative positions at A1 – B2 Levels on the CEFR

CEFR level	Attributive position	Predicative position				
		Without prepositional complement	Correct prepositions	Incorrect prepositions	Omission	Confused meaning
A1	210	208	19	9	0	0
A2	2	97	10	2	2	2
B1	8	83	4	0	1	1
B2	0	8	1	2	0	0

Source: CZEMATELC 2017

5.1 A1 adjectives

The raw frequency of A1 adjectives is 482, but 36 of them were identified in greetings and incoherent sentences, so they were discarded from further analyses. The remaining ones are represented by 12 lemmas with very high frequency. The most frequent were: *good* (n=211), *happy* (n=144), *nice* (n=61), *bad* (n=14), *sorry* (n=13), *hard* (n=10), *tired* (n=7), *famous* (n=7). Some of them occupied predominantly attributive positions in sentences, namely the adjective *famous*, which is used only in this position as in (1).

> 1) **After the cinema we was at the famous restaurant Amigo in České Budějovice.* (1C-17-1.txt)

Most of the A1 adjectives in predicative positions assumed the role of subject predicative without being complemented further. Their number is relatively high as shown in Table 3. The table also shows that the learners complemented some of the adjectives with *to-infinitive* or various clauses as in (2) and (3).

> 2)* *I was so happy to see Diana after long time*. (1C-17-11.txt)
> 3) **I said, that I'm sorry, that I don't know, who is it*. (1P-17-1.txt)

Thorough analysis of both the clauses in which the learners used to complement the adjectives and the sentences without any complementation appear to indicate that merely a few sentences in the corpus would be considerably improved if prepositional phrases were used, largely because the learners would have avoided dealing with complicated grammar that leads to errors or misunderstandings. Moreover, a cursory look at the A1 adjectives in the reference corpus *Brown Family (C8 tags)*, reveals that even native speakers rarely complement these adjectives with prepositional phrases (i.e. *good at* – 0.7%, *good for* – 1.6%, *happy about* – 1.6%, *happy with* – 4%). A considerably higher density of these phenomena in learner texts would therefore be unnatural. The only A1 adjectives that co-occur with prepositions more frequently in the corpus were: *different from* (8.6%), *sorry for* (11.7%) and *tired of* (27%). Consequently, these collocations require more attention when teaching.

Table 3 A1 Adjectives in predicative position not complemented by prepositional phrases

Adjective	To-infinitive	That-clause	Reason clause	When-clause	If-clause	Without any complement	Total
happy	10	25	40	5	3	44	125
good	1					28	29
nice		2				17	19
bad						7	7
sorry	2	2			1	6	11
hard						5	5
tired		1				6	6
clever						3	3
bored						1	1
different						1	1
important						1	1

Source: CZEMATELC 2017

The number of correctly used dependent prepositions complementing A1 adjectives is relatively low and it can be illustrated by the following examples (4) and (5).

4) *Is Thursday good for you*? (2M-17-7.txt)
5) *I wasn't very happy about this visit.* (1M-17-3.txt)

No omissions were identified and the number of incorrect dependent prepositions was also very low. As can be seen in Table 4, complementing the adjective *happy* with the preposition *from* is the most frequent error in this group of adjectives. However, this error only appears three times as in (6) and could be attributed to mother tongue influence. The confusion between two prepositions that can complement one adjective can be seen in (7). The learner probably blended two constructions: people can be *good to* other people, but a thing or situation can be *good for* them. The cross-linguistic influence is more noticeable in the overuse of the preposition *on* that incorrectly complements several adjectives across all four language proficiency levels. In this group of adjectives, it complements the adjective *good* as in (8), especially as the whole sentence appears to be an exact translation from Czech.

6) **He was nice and so friendly. I was so happy from him.* (1L-17-11.txt)
7) **It will be really good to me, if you borrow me your bike.* (2H-17-12.txt)
8) **Now she is hospitalized and she is good on it.* (1J-17-4.txt)

Table 4 A1 Adjectives with dependent prepositions

Adjectives	Correct prepositions		Incorrect prepositions	
good	good **for** sb	4	good **on** sth	2
			good **to** sb	1
happy	happy **about** sth	1	happy **from** sb	3
	happy **for** sb	1	happy **of** sth	1
	happy **for** sth	6		
hard	hard **for** sb	1	hard **in** sth	1
important	important **for** sb	1		
nice	nice **to** sb	4	nice **from** sb	1
			nice **too** sb	1
sorry	sorry **for** sth	1		
Total		**19**		**10**

Source: CZEMATELC 2017

5.2 A2 adjectives

The A2 adjectives in the studied school-leaving essays seldom assumed attributive positions because they largely belong to a group of adjectives that have a stronger natural preference for predicative positions (cf. Biber et al., 2002). A2 adjectives in the corpus complemented the following verbs: *to be* (n=102), *to feel* (n=9), *to make* (n=3), *to look* (n=2), however, in most cases they were not complemented further. Apart from prepositional phrases, they were complemented with *that-clauses* merely five times. The adjective *surprised* was complemented with a *that-clause* four times (9) and the adjective *mad* once (10).

> 9) *I was very surprised that he was still speaking Czech very well…*(1D-17-2.txt)
> 10)*When I heard doorbell ringing I was pretty mad that I must leave my computer.* (1D-17-7.txt)

This group also shows the second highest proportion of attempts to complement the adjectives with prepositional phrases with a greater than average success rate. This may be explained by the fact that many of these adjectives (such as *afraid of, worried about, full of, interested in*) are already presented with dependent prepositions in elementary and pre-intermediate course books, which can be considered "one of the primary sources of [foreign language] input in the classroom" (Tono, 2004, p. 45). If extensive opportunities to practice them in different types of exercises is lacking, they are at least included in the input activities. However, the raw frequency of attempts (n=16) and the slightly greater than average success rate seem to be low, especially as the above mentioned adjective-preposition collocations (*interested in* – 58%, *worried about* – 35%, *afraid of* – 29%, *full of* – 24%) show a high percentage of instances of co-occurrence in the reference corpus. It can be argued that an attempt should be made to look for ways to enhance their acquisition. All the successful and unsuccessful attempts

to complement these adjectives in CZEMATELC 2017 are presented in Table 5. The adverb-adjective confusion and inappropriate spelling of the adjective *careful* can be illustrated by the two following examples. The first one uses an incorrect preposition (11), and the second one, which does not mention the bicycle directly, uses a correct dependent preposition (12).

> 11) *I promise, i will be very carefully at your bike. Please answer me asap.* (2E-17-4.txt)
>
> 12) *… for that one day and I will be very carefull with it.* (2E-17-2.txt)

Dependent prepositions were omitted twice. The first example, (13), is in a collocation which does not tend to cause problems to Czech learners and the second example, (14), an adjective followed by a clause, might be attributed to the fact that the learner was trying to complement the adjective in a similar fashion to that possible in Czech by separating the clause with a comma.

> 13) *This invite was full* happy feeling, memories because…* (1S-17-5.txt)
>
> 14) *I didn't expect anyone, so I was quite interested*, who could it be.* (1I-17-3.txt)

The corpus also contains two examples (15) and (16) in which the dependent prepositions appear to be used correctly, but the adjective *afraid* seems to be used instead of the adjective *worried.* This is probably due to the influence of Czech, in which the exact equivalent for fear or being afraid is also commonly used when talking about worries.

> 15) *I gave him some piluls on sick. I afraid about him, but he was fine he had only diarrhoea.* (1S-17-6.txt)
>
> 16) *I am a little bit afraid about leaving me again.* (1F-17-12.txt)

Table 5 A2 Adjectives with dependent prepositions

Adjectives	Correct prepositions		Incorrect prepositions		Omission		Confused meaning	
afraid	afraid **of** sth	2					afraid **about** sth	**2**
busy	busy **with** sth	1						
careful	careful **with** sth	1	careful **at** sth	1				
full	full **of** sth	3			full *sth	1		
interested	interested **in** sth	1			interested * sth	1		
mad	mad **at** sb	1						
surprised			surprised **from** sb	1				
worried	worried **about** sb	1						
Total		10		2		2		2

Source: CZEMATELC 2017

5.3 B1 adjectives

The B1 adjectives in CZEMATELC 2017 have a lot in common. Most of them assume predicative positions without being complemented further, with the exception of seven adjectives that the learners tried to complement with prepositional phrases and six adjectives that were complemented with *to-infinitive* and/or different clauses. This is documented in Table 6. The only adjective that appears merely in attributive positions is the adjective *typical/tipical* (n=7) and the adjective *grateful* (n=16), which appears in this position only once.

Table 6 B1 Adjectives in predicative position not complemented by prepositional phrases

Adjective	To-infinitive	That-clause	Wh-clause	If-clause	Reason clause	Without any complement	Total
amazed			1			1	2
curious			1			0	1
excited	1	1				3	5
grateful		1	1	3		10	15
satisfied		1				1	2
scared	2	1			3	32	38

Source: CZEMATELC 2017

The relatively high frequency of the adjective *scared* (n=39) might have been affected by the task to a certain extent because many of the learners exploited the feeling of fear

in their narratives about an unexpected visitor. It appears eight times in the form of *scary*, which suggests that learners tend to confuse the meaning of these adjectives. This adjective is also complemented in several different ways. It is complemented with *to-infinitive* (17), with subordinate clauses starting with *because* (18) and with a *that-clause* (19), although in the last example the adjective *scared* more likely expresses the meaning of the adjective *worried*.

> 17) **I was wery surprised and a bit scared to open the doors but when I looked out…*(1P-17-6.txt)
> 18) **We was so scary because our favorite movie was scary movie…*(1P-17-11.txt)
> 19) **I always wanted to met my dad, but I was scared that we wont understand each other.* (1B-17-5.txt)

This confusion is similar to that in the only example in which the adjective is complemented by an inappropriate preposition (20).

> 20) **This moment I never won't to experience again, because I was so scared of my life and I was from this "an unexpected visitor" never home alone.* (1P-17-8.txt)

As evidenced in Table 7, four adjectives at this level appear only once, but they are always complemented with a prepositional phrase with a correct dependent preposition. These adjectives are *concentrated* (21), *frightened* (22), *proud* (23) and *suitable* (24).

> 21) *I was highly concentrated on studying for my school leaving exam* (1K-17-3.txt)
> 22) **Nothing was happend but we was really frightened of him.* (1F-17-13.txt)
> 23) *I'm very proud of my father and his outstanding work.* (1H-17-2.txt)
> 24)*… on Wednesday at 5 PM if it is suitable for you.* (2F-17-16.txt)

This may indicate that they were acquired together as collocations. However, their very low frequency may imply collocation avoidance which can be identified even in the written production of advanced learners (Nesselhauf, 2003; Philip, 2007). Interestingly though, three of these adjective-preposition collocations (*proud of* – 41%, *suitable for* – 28%, *satisfied with* – 27%) also tend to frequently co-occur in the reference corpus. The high percentage of co-occurring instances applies also to *concentrated on* (22%), which appears in CZEMATELC 2017 merely once, and *amazed at* (35%) and *typical of* (26%), which were not found at all.

Table 7 B1Adjectives with dependent prepositions

Adjectives	Correct prepositions	Incorrect prepositions	Omission	Confused meaning
concentrated	concentrated **on** sth 1			
curious			curious *sth 1	
frightened	frightened **of** sth 1			
proud	proud **of** sb 1			
scared				scared **of** sth 1
suitable	suitable **for** sb 1			
Total	4	0	1	1

Source: CZEMATELC 2017

5.4 B2 adjectives

Only individual instances of B2 adjectives, with the exception of **alergic* (n=2) and *enthusiastic* (n=5), were found. Interestingly, these two adjectives (*allergic to* – 27%, *enthusiastic about* – 13.5%) frequently co-occur with the prepositions in the reference corpus. The latter is also the only adjective at this level that is complemented once with a *that-clause*. As evidenced in Table 8, the only adjective correctly complemented with a preposition in this group is *exhausted* (25).

> 25) *I came home exhausted from work and wanted to pour a glass of… (1L-17-2.txt)*

The two following examples of incorrect complementation (26) and (27) come from one essay.

> 26) **But John is alergic on cheese and milk. I forgot it.* (1S-17-6.txt)
> 27) **Now I never forget on his alergic on cheese and milk. (1S-17-6.txt)*

They illustrate incorrect spelling of the adjective and also possibly the influence of the mother-tongue both on the spelling of the adjective and on the choice of the preposition because the exact equivalent of this preposition would follow this adjective in Czech. The low frequencies of B2 adjectives and the low success rate with regards to complementing them with dependent prepositions may imply that the difficulty of B2 vocabulary was beyond the learners' abilities. Moreover, it would seem that they were unwilling to take risks with language and therefore relied on simple syntactic structures because they knew that accuracy was an important criterion that would affect their grade.

Table 8 B2 Adjectives with Dependent Prepositions

Adjectives	Correct prepositions		Incorrect prepositions	
allergic			*allergic **on** sth	2
exhausted	exhausted **from** sth	1		
Total		2		2

Source: CZEMATELC 2017

5.5 Adjective-preposition collocations selected for teaching

In common with other studies (see e.g., Nesselhauf, 2003; Philip, 2007), this research attempted to determine which collocations need to be taught explicitly by identifying several adjective-preposition collocations at relevant CEFR levels which appear both in the learners' production and also tend to co-occur strongly in the reference corpus *Brown Family (C8 tags)* (see Table 9).

Table 9 Collocations selected for teaching

CEFR level	Collocations
A1	tired of, sorry for, different from
A2	interested in, worried about, afraid of, full of
B1	proud of , amazed at, suitable for, satisfied with, typical of, concentrated on
B2	allergic to, enthusiastic about

6. Conclusion and implications for teaching

The study aimed to investigate the frequency and accuracy of adjective-preposition collocations in CZEMATELC 2017 to see which collocations and how successfully were acquired and to select collocations for teaching. Corpus-based discovery-learning based on data not only from a native speaker corpus, but also from CZEMATELC 2017 is proposed.

The adjectives that could be potentially complemented with prepositional phrases range between A1 to B2 levels. However, the majority of learners opted not to complement most of them with prepositional phrases, which may imply that they failed to acquire adjectives with dependent prepositions as collocations. This raises serious doubts about the overall proficiency of Czech secondary school students because collocational knowledge was found by Williams (2000) to correlate strongly with the general proficiency of EFL learners. The number of correctly and incorrectly used adjective-preposition collocations was so low, that identifying error patterns and patterns of appropriate use was very difficult. However, the L1 influence as well as the tendency to complement correctly B1 adjectives with very low frequency was noticeable.

The collocations that have been selected on the basis of this research as requiring special attention in teaching should be taught as one unit and observed both in texts and

hands-off activities derived from native speaker corpora. Secondary schools are among those institutions where hands-on activities would be difficult to apply and where the relatively low language proficiency and inexperience of learners with corpus tools would create further barriers to adopting data-driven learning. However, hands-off activities that stimulate observation of adjective complementation in simulated concordance lines may aid input enhancement by emphasising the target structure. Repeated exposure through several activities may also provide input enrichment. The first phase, when learners work in groups and share their discoveries and support each other, should be followed by a clarification from the teacher that enables the confirmation or correction of hypotheses. As an additional tool, it is suggested to observe selected sentences and/or paragraphs from CZEMATELC 2017 because highlighting the features of learner language on their own could make some writing problems seem more obvious. In addition, being able to improve those sentences using appropriate collocations might be an important step towards consolidating collocation knowledge and developing writing skills.

References

Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning & Technology, 21*(3), 195–216.

Anthony, L. (2014). *AntConc*. Tokio: Waseda University. Retrieved from www.laurenceanthony.net/software/anconc/

Barabadi, E., & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English vocabulary. *Cogent education, 4,* 1–13. Retrieved from: https://www.cogentoa.com/article/10.1080/2331186X.2017.1283876.pdf

Bardovi-Harlig, K., Mossman, S., & Su, Y. (2017). The effect of corpus-based instruction on pragmatic routines. *Language Learning & Technology*, *21*(3), 76–103.

Bennett, G.R. (2010). *Using corpora in the language learning classroom - corpus linguistics for teachers*. Ann Arbor: University of Michigan Press.

Biber, D., Conrad, S., & Leech, G. (2002). *Longman student grammar of spoken and written English.* Harlow: Pearson.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* Harlow: Pearson.

Boulton, A. (2012). Hands-on/hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora,* (pp. 152–168). Brno: Masaryk University Press.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348–393.

Cambridge University Press. (2015). English Grammar Profile. Retrieved from http://englishprofile.org/english-grammar-profile/egp-online

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Centrum pro zjišťování výsledků ve vzdělávání (Centre of Educational Assessment). (2014). *Katalog požadavků zkoušek společné části maturitní zkoušky* (Catalogue of requirements for the common part of the school-leaving exam). Praha: Author.

Centrum pro zjišťování výsledků ve vzdělávání (Centre of Educational Assessment). (2017). *Písemná práce z cizích jazyků* (Written part of the school-leaving exam in foreign languages). Praha: Author.

Chuang, F-Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of L2 English produced by Chinese students. *Corpora, 1*(2), 251–271.

CQPweb v3.2.29. (2008–2016). Brown Family (C8 tags). Retrieved from https://cqpweb.lancs.ac.uk/.

Dušková, L. (1969). On sources of errors in foreign language learning. *International Review of Applied Linguistics in Language Teaching, 7*(1), 11–36.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

English Profile. (2012). English Vocabulary Profile. Retrieved from http://vocabulary.englishprofile.org/staticfiles/about.html

Gabrielatos, C. (2005). Corpora and language teaching: Just a fling of wedding bells? *TESL-EJ, 8*(4), 1–39.

Halliday, M. A. K. (1992). Language as system and language as instance: the corpus as a theoretical construct. In J. Svartvik (Ed*.), Directions in corpus linguistics: proceedings of the Nobel Symposium 82, Stockholm, 4 – 8 August 1991,* (pp. 61–77). Berlin: Mouton de Gruyter.

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning, 15*(1, 2), 29–53.

Huei Lin, M. (2016). Effect of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly, 50*, 871–893.

Hunston, S., & Francis, G. (2000). *Pattern Grammar: A corpus-driven approach to the lexical grammar of English.* Amsterdam: John Benjamins.

Jafarpour, A., & Koosha, M. (2006). Data-driven learning and teaching of collocation of prepositions: the case study of Iranian EFL students. Research on foreign languages. *Journal of Faculty of Letters and Humanities, 49*(200), 1–31.

Jaworska, S., Krummes, C., & Ensslin, A. (2015). Formulaic sequences in native and non-native argumentative writing in German. *International Journal of Corpus Linguistics, 20*(4), 500–525.

Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *ERL Journal, 4*, 27–45.

Kennedy, C., & Miceli, T. (2017). Cultivating effective corpus use by language learners. *Computer Assisted Language Learning, 30*(1-2), 91–114.

Kennedy, G. D. (1991). Collocations: Where grammar and vocabulary teaching meet. In S. Anivan (Ed.), *Language teaching methodology for the nineties,* (pp. 212–229). Singapore: RELC.

Kulsitthiboon, S., & Pongpairoj, N. (2018). Cooperative corpus consultation for acquisition of adjective + preposition collocations. *Journal of Language Studies, 18*(3), 57–72.

Lee, D. Y. W., & Chen, S. X. (2009). Making a bigger deal of the smaller words: Function words and other key items in research writing by Chinese learners. *Journal of Second Language Writing, 18,* 281–296.

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fliegelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora,* (pp. 11–23). Harlow: Longman.

Leech, G., & Svartvik, J. (1993). *A communicative grammar of English*. Harlow: Longman.

Lindstromberg, S. (1996). Prepositions: Meaning and method. *ELT Journal, 50*(3), 225–237.

Lorincz, K. & Gordon, R. (2012). Difficulties in learning prepositions and possible solutions. *Linguistic Portfolios, 1*(1), 1–5.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book.* London: Routledge.

Mueller, C. M. (2011). English learners' knowledge of prepositions: Collocational knowledge or knowledge based on meaning? *System: An International Journal of Educational Technology and Applied Linguistics, 39*(4), 480–490.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223–242.

Philip, G. (2007). Decomposition and delexicalisation in learners' collocational (mis)behaviour. In *Online Proceedings of Corpus Linguistics 2007*. Retrieved from http://ucrel.lancs.ac.uk/publications /CL2007/paper/170_Paper.pdf

Proudfoot, S. (2010). A corpus-led exploration of lexical verb use in Main Suite Writing papers. In H. Khalifa & I. Vidaković (Eds.), *Research notes 41/2010*, (pp. 26–31). Cambridge: University of Cambridge ESOL Examinations.

Sadeghi, K., & Panahifar, F. (2013). A corpus-based analysis of collocational errors in the Iranian EFL learners' oral production. *The Journal of Teaching Language Skills, 4*(4), 53–78.

Salamoura, A., & Saville, N. (2010). Exemplifying the CEFR: criterial features of written learner English from the English Profile Programme. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research,* (pp. 101–132). Eurosla.org: Eurosla.

Sinclair, J. (1991). *Corpus, concordance and collocation*. Oxford: Oxford University Press.

Sparling, D. (1990). *English or Czenglish? Jak se vyhnout čechismům v angličtině*. Praha: SPN.

Tan, M. M. L. (2000). *Prepositional clusters: investigative-oriented learning and English language teaching*. PhD thesis. Nottingham: University of Nottingham.

Tono, Y. (2004). Multiple comparisons of IL, L1 and TL corpora: The case of L2 acquisition of verb subcategorization patterns by Japanese learners of English. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and Language Learners* Vol. 17, (pp. 45–66). Amsterdam: John Benjamins.

Tono, Y., Satake, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL, 26*(2), 147–162.

Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, *20*(3), 159–179.

Walter, E. (Ed.). (2008). *Cambridge* a*dvanced learner's dictionary*. Cambridge: Cambridge University Press.

Williams, B. J. (2000). Testing ESL learners' knowledge of collocations. *ELT Journal, 35*, 115–122.