

Dragana Vuković Vojnović

University of Novi Sad, Serbia

2021, Vol. 18 (2), 51-68(230)

revije.ff.uni-lj.si/elope

<https://doi.org/10.4312/elope.18.2.51-68>

UDC: [81'373.72:338.48]:81'322

Key *Noun* + *Noun Collocations* in the Language of Tourism: A Corpus-Based Study of English and Serbian

ABSTRACT

In this paper, we investigate the main characteristics underlying *noun + noun collocations* in the English and Serbian language of tourism. Their morpho-syntactic, semantic and communicative features are contrasted and compared in the two languages. Firstly, we compiled two comparable corpora in English and Serbian from the tourism websites of Great Britain and Serbia. Based on their normalized frequencies per 10,000 words, key noun + noun collocations were extracted, using TermoStat Web 3.0 and AntConc. The results showed certain similarities in terms of the prevailing topics in the two corpora, based on the analysis of key noun + noun collocations. However, we found major differences in the two languages in terms of their morpho-syntactic features, communicative focus and the relationship of the collocates. The results of the study have implications for English for Tourism education, tourism discourse studies, language typology and lexicography.

Keywords: noun + noun collocations, contrastive corpus analysis, custom-made corpora, corpus analysis tools, language of tourism

Ključne dvosamostalniške kolokacije v jeziku turizma – korpusna študija angleškega in srbskega jezika

IZVLEČEK

V prispevku raziskujemo glavne značilnosti dvosamostalniških kolokaciji v angleškem in srbskem jeziku turizma. Primerjamo njihove oblikovno-skladenjske, pomenske in komunikacijske značilnosti v obeh jezikih. Najprej smo sestavili dva primerljiva korpusa v angleškem in srbskem jeziku, za katera smo črpali podatke s turističnih spletnih strani Velike Britanije in Srbije. Na podlagi njihovih normaliziranih frekvenc na 10.000 besed so bile z uporabo TermoStat Web 3.0 in AntConc izločene ključne kolokacije, sestavljene iz dveh samostalnikov. Rezultati njihove analize so pokazali določene podobnosti glede prevladujočih tem v obeh korpusih. Vendar smo v obeh jezikih ugotovili velike razlike glede na obliko-skladenjske značilnosti, komunikacijsko osredotočenost in odnos med kolokati. Rezultati študije lahko koristijo pri pouku angleščine za potrebe turizma, študijah turističnega diskurza, jezikovni tipologiji in leksikografiji.

Ključne besede: dvosamostalniške kolokacije, protistavna korpusna analiza, po meri ustvarjeni korpusi, orodja za korpusno analizo, jezik turizma

1 Introduction

In this paper, we present and analyse some of the main features of *noun + noun collocations* in the language of tourism in English and Serbian, based on the contrastive analysis of two comparable custom-made corpora of texts compiled from the promotional tourism websites of Great Britain and Serbia. The initial motives for this study were the preparation of prospective course content for compulsory English for Tourism courses at the Faculty of Sciences, University of Novi Sad, as well as a long-term experience of the author in translating scientific papers in the field of tourism and related fields.¹

Firstly, referring to the genre and discourse analysis of Swales and Bhatia (Swales 1990; Bhatia 1993; 2002), it has been observed that the language of tourism can be considered as a special discourse with its specific language structures and functions, specific written and spoken forms, and what is the most pertinent to this study, with specific vocabulary rich in collocations and idiomatic expressions. Secondly, the language of tourism has a high communicative value for the professionals in this field who understand the nuances of its terminology and aim to communicate successfully with their colleagues. However, it is also a powerful tool for communication with non-professionals, i.e., clients, passengers, guests, and tourists, and its main purpose in this sense is to create a desire to travel and persuade potential clients to depart on a journey (Dann 1996; Thurlow and Jaworski 2010; Jaworski and Pritchard 2005). In addition to being objective and informative, the language of tourism as the language of promotion provides an embellished view of reality, hence its communicative practices can be considered as using the language of power in a very subtle way by providing potential clients with the offer they wished for. Moreover, new ways of communicating with clients through the web have created a sort of hybrid tourism discourse which is opening more doors for language specialization and use, keeping in mind responsible policies and strategies of the tourism industry (Maci 2010).

Finally, to refer back to the main purpose of the paper, in English for Tourism classes it is of the utmost importance to teach and learn new vocabulary and acquire the right terminology of the particular field. It has been observed that *noun + noun collocations* coupled with *adjective + noun collocations* are used quite often when describing a particular tourist destination, accommodation, facilities, food, aspects of culture and entertainment, etc. (Lam 2007; Kang and Yu 2011; Vuković Vojnović 2011). Acquiring adequate collocations in a foreign language can be problematic in particular because of the L1 transfer and different morpho-syntactic characteristics of L1 and L2, which is the case with English and Serbian *noun + noun collocations*. This is why it is important to choose the most adequate and effective collocations in the target language and contrast them with L1 collocations in a way that would assist students and/or tourism professionals to communicate more successfully in the target language (e.g., English) with their prospective clients through electronic media and websites, which are, today, the first step in every journey.

2 Theoretical Background

This study is grounded in the findings from two main theoretical frameworks – corpus-based studies of collocations and contrastive studies of collocations, relying also on tourism

¹ The results and findings presented here are part of a larger study that was conducted for a doctoral thesis.

discourse studies. In this section, we will present a brief overview of the theoretical views and relevant literature.

2.1 Tourism Discourse

Tourism English is recognized as a special register of the English language which serves a special purpose and is different from General English (Lam 2007, 72). Gotti (2008, 22–24) reflects on specialized discourse as a form of general discourse with some additional criteria, including mono-referential and denotative meaning of terminology, which is also true for the tourism domain (e.g., *tour operator*, *package holiday*, *buffet breakfast*). However, the specialist language of tourism is not always neutral in register but has a connotative function as well, because objective information is quite often embellished to emphasize the positive features of a destination, or a service offered to tourists. After comparing a large custom-made Tourism English Corpus with the Freiburg-LOB Corpus of British English (F-LOB)², Kang and Yu (2011) came to the conclusion that Tourism English is descriptive, positive, and aimed at creating enthusiasm and appreciation of aesthetic values among potential tourists. Tourism discourse creates a specific vision of reality, social identities and appreciation of self and the other (Jaworski and Pritchard 2005, 6–7). In tourism discourse, physical features, for example *land*, become visual objects of tourist desire, for example *landscape*, because the actual space and time are reshaped and conceptualized by the language of tourism (Jaworski and Pritchard 2005; Urry 1995). This sort of intervening approach towards reality in the language of tourism is called *the tourist gaze*, a term introduced by Urry (Jaworski and Pritchard 2005, 9).

2.2 Corpus-Based Studies

The study of authentic language based on a corpus has the following characteristics: “its usage-based corpus methodology, the central role of the notion of collocation, and the technological background of the approach” (Geeraerts 2010, 169). Custom-made corpora in ESP education can be incredibly useful for exploring and analysing authentic specialist materials that would lead to the identification of key specialist terminology and concepts in a specific field or profession (e.g., tourism discourse). Over the past seventy years, there have been a lot of corpus-based studies that dealt with ESP and EAP issues (West 1953; Xue and Nation 1984; Sinclair 1991; Howarth 1998; Coxhead 2000; Gavioli 2005; Nesselhauf 2005; Lam 2007; McClaren 2007; Granger and Meunier 2008; Schmitt 2008; Geeraerts 2010; McEnery and Xiao 2010; Cobb and Boulton 2015; Liu and Han 2015; Hyland and Jiang 2017, etc.). These led to a deeper understanding of lexical features and terminology in specialized languages, using both statistical and qualitative data obtained from large electronic or custom-made corpora, exploring the relationship between the vocabulary and the domain/context, making use of empirical, authentic examples of professional communication, as well as exploring the relationship between L1 and L2 that affects the target language acquisition and production, and creating word lists and collocation lists for Academic English or for specialist fields.

In contrastive corpus studies, researchers can focus on comparing language features in two or more languages using large electronic corpora, such as the Brigham Young University

² See <https://varieng.helsinki.fi/CoRD/corpora/FLOB/>.

Corpus integrated with the British National Corpus³ – BNC or the Corpus of Contemporary American English – COCA, which is a huge dynamic corpus with more than 560 million words⁴. In addition to word frequency studies, corpus-based research can provide qualitative data on collocational information such as a broader context of language use with authentic examples of specific language features related to particular topics (McCarten 2007). When analysing collocations in the corpus, researchers usually produce concordances which show the words in their immediate context, usually according to a Key Word in Context index (Geeraerts 2010, 170). The context shown in concordance lists helps us distinguish frequent co-occurrences of words that are not collocations (e.g., *of the*).

In studies related to language pedagogy, corpora are usually purpose-built from texts in a specialist field. In particular, Gavioli (2005) explored the application of corpora in ESP language teaching, creating special, custom-made corpora that would meet the needs of her language classroom and help students acquire specific vocabulary for their specialist field. Most of the corpora were made up of specialized texts downloaded from the web, keeping in mind certain criteria when choosing the texts for the corpora – avoiding multiple texts written by the same author or from the same journal or source, and ensuring that most of the texts were written by native speakers, with 20% written by non-native speakers (Gavioli 2005, 8). Another way of using corpora for data-driven learning is to create so-called learner corpora, compiled from texts produced by non-native speakers of a language, i.e., students, researchers, and professionals, which are then compared with the native-speaker corpora. In the foreign language learning contexts, if learner corpora are to be adequate and systematic then they need to be collected in such a way that there is little control by the teacher/researcher, such as with a collection of student essays which only have a preset title or through open-ended interviews (Nesselhauf 2005, 40).

As Geeraerts (2010, 181) observes, corpus-based collocational lexical description led to the development of corpus analysis tools which can be easily used even for a large-scale quantitative analysis of the corpus. The tools used for the investigation of custom-made corpora are mostly available online and are constantly being developed and improved (e.g., LancsBox⁵, Sketch Engine⁶, Range⁷, etc.) For the purpose of this study, we have used TermoStat Web 3.0 for the English corpus and AntConc for the Serbian corpus, which will be explained in the section on research methodology.

2.3 Contrastive Approach to Collocations

The contrastive approach to the investigation of collocations and prefabricated language items is the key to providing a more systematic insight into the complexities of acquiring foreign language vocabulary. It has an array of applications in foreign language education, corpus studies, translation studies, and the development of electronic corpora and glossaries. In this subsection, we will present a few examples of relevant contrastive research and also

³ See <https://corpus.byu.edu/bnc>.

⁴ See <https://www.english-corpora.org/coca/>.

⁵ See <http://corpora.lancs.ac.uk/lancsbox/>.

⁶ See <https://www.sketchengine.eu/>.

⁷ See <https://www.range-software.com/>.

establish the definition of collocations and criteria applied in this paper. We will first look into some conclusions derived from empirical research. However, it is interesting to note that there have not yet been any relevant studies comparing N+N collocations in English and Serbian, in particular in an ESP context.

One of the major projects regarding the contrastive analysis of English and Serbian was *The Yugoslav Serbo-Croatian English Contrastive Project* in the 1970s, whose main editor was Rudolf Filipović, and the main institutions participating in the project were the University of Zagreb, the Institute of Linguistics and the Centre for Applied Linguistics from Washington, DC. Some of the findings from this project definitely shaped the understanding of contrastive analysis in this region, and shed light on various linguistic features of the two languages in contrast. The research results published within the project will not be discussed here, however, because they go beyond the research aims and the scope of the paper.

Đorđević (1989) gives one of the most comprehensive contrastive insights into the features of nominals including the classification of noun collocations in English and Serbian with numerous examples in the two languages. Special attention is given to the study of attributive and classifying functions of adjectives and nouns in Adj+N and N+N collocations.

Nation (2007, 328) explains that it is important to investigate collocations for successful vocabulary learning in a foreign language. Firstly, we should decide if the collocations are frequent and/or predictable, and what their common patterns are. If they are frequent and predictable, they will be more easily acquired by the students because they will have multiple chances to come across such collocations in a broader written and spoken context. High-frequency collocations are also often included in dictionaries, so students can look them up easily. In a specialized context, the collocations that are less frequent and less predictable should be collected as glossaries which were found to be useful in teaching ESP at tertiary level (Gajšt 2006; Paradž 2021). Collocation lists can also facilitate the choice of collocations to be taught through direct teaching methods. Otherwise, the foreign language students will have difficulty in acquiring such collocations, because the criteria for them are not as easily transparent and available as is the case with native speakers. In his vocabulary development strategy, Nation (2007) explains that explicit teaching of carefully chosen vocabulary is especially important in the early stages of academic foreign language learning and languages for specific purposes, and that vocabulary lists can boost the incidental vocabulary learning that happens during reading and listening activities.

Granger (1998) defines her methodology as Contrastive Interlanguage Analysis, first contrasting and comparing language use among native and non-native speakers, and also comparing the target language among non-native speakers with different native languages. After investigating collocations, Granger concluded that non-native speakers most often produce unusual combinations for the target language, which resemble collocations in their native language, and are not the same as collocations made by native speakers of the target language (Granger 1998, 152–53).

Howarth (1998) investigates *verb + noun collocations* (e.g., *reach a conclusion, produce findings*) in academic writing in English among students who are not native speakers of English, and

his main motivation for the study was to provide a descriptive framework for teachers who need to assess student written assignments and find an easier way to give valuable, data-based feedback to students.

Collocations are in the middle of a continuum (Cowie 1981; Howarth 1998) ranging from free combinations to fixed idiomatic expressions, where interchangeability of their constituents is not possible. Collocations also comprise a spectrum of typical word combinations where the secondary collocate can easily be replaced by a synonym, to more restricted collocations made up of a more fixed set of units. Ackermann and Chen (2013, 236) argue that even so-called free collocations or free combinations imply certain syntactic and semantic restrictions regarding the words that co-occur in them. For example, *write an essay* can be regarded as a free combination, but, syntactically, the verb *write* is followed by an object, and, semantically, that object can only be something that is a written product, e.g., an essay or a song. Furthermore, collocations cause additional problems for L2 learners regardless of their proficiency level in L2 because of the element of “grammatical or lexical unpredictability or inflexibility” (Nation 2001, 324 in Ackermann and Chen 2013, 236). Native speakers do not experience such difficulty and can easily recognize collocations (Ackermann and Chen 2013, 236). In the context of foreign language learning, it is important to acknowledge the open-choice principle for making collocations which is used by non-native speakers, as defined by Sinclair (1987, 319). Conversely, native speakers choose combinations which are well-established in the language and even prefabricated, based on the principles that are not transparent or available to non-native speakers, as we have already mentioned. Since the majority of collocational errors in L2 come from L1 interference, contrastive analysis of collocations helps the teacher to eliminate the collocations that are equivalent in L1 and L2, and to choose language-specific collocations to be explored further (Bahns 1993, 56–57).

Although Palmer (1933) was one of the first to deal with collocations in English more deeply in his *Second Interim Report on English Collocations*, it was Firth’s insightful statement⁸ that prepared the ground for a new approach to understanding meaning, moving away from individual words and drawing attention to context (Firth 1957, 11). His approach was upgraded by Halliday (1966) who singled out lexical collocations and drew attention to the lexical level of language, which should be separate from the grammatical level, but he also emphasized their interconnectedness. It was Sinclair (1966; 1987; 1991) who developed Halliday’s model even further focussing on the collocability of words and the influence of context, which led to his development of corpus studies. In more recent times, within the scope of lexical and cognitive semantics, there is also an emphasis on the context, i.e., the domain, which is seen as important for understanding the meaning of a lexeme (Croft and Cruse 2004, 18).

The theory and definition of collocations has revolved around the understanding of lexical, syntactic and semantic principles, producing different approaches across different linguistic disciplines, such as structural linguistics, functional linguistics, generative semantics, lexical semantics, cognitive semantics, etc. Cruse (1986) defines lexical collocations as a series of lexical units that make a standard combination and for which each lexical constituent is

⁸ “You shall know a word by the company it keeps.”

also semantic, so that the meaning is transparent (e.g., *fine weather, torrential rain*). When considering the study of collocations in the Serbian language, Dražić (2014, 73) provides a rather complex definition of collocations, defining them as “a syntagmatic lexical combination of two or more auto-semantic lexemes, organized according to the morphosyntactic rules of a given language, with a greater or lesser degree of permanence and interconnectedness of elements, conditioned by linguistic and non-linguistic, cultural-historical and social data.”

For the purposes of this study, relying on the findings of Van Roey (1990) and Granger (1998), lexical collocations will be defined as combinations of two or more lexemes that form a connection so that one collocate pairs up with a particular lexeme from a pool of synonyms, regardless of their syntactic restrictions.

3 Research Methodology

3.1 Research Aims

The ultimate goal of this research is to single out and analyse *noun + noun collocations* that are specific to the language of tourism in English and Serbian and can be characterized as terminological. More specific objectives of the investigation include determining the morphosyntactic and semantic features of these collocations, classifying collocations around specific topics, and discussing the similarities and differences between the collocations in the two languages. This sort of analysis and its findings provide a platform for a more systematic approach to collocations in terms of their lexical-semantic features and use, and their adoption in the language of the profession, primarily in terms of developing communicative competencies in English among non-native speakers for whom tourism is a broader area of their future or current profession. Hence, one of the goals in the future would be to create a list of key terminological collocations in the field of tourism in English and Serbian.

3.2 Corpus

Firstly, two comparable corpora were collected from electronic texts in English and Serbian that were excerpted from the websites of official national tourism organizations of Great Britain (www.visitbritain.com) and Serbia (www.serbia.travel), and from the websites of two leading travel agencies in the respective countries (www.tui.co.uk and www.kontiki.rs). The English Tourism Corpus (ETC) consists of 98,567 words and the Serbian Tourism Corpus (STC) contains 87,489 words. These particular websites were chosen because they represented the best possible sources at the time of compilation and would cover a wide range of promotional tourism texts. On the one hand, the websites of national tourism organizations present what that country has to offer to domestic and international tourists, i.e., their main business interest lies both in domestic and inbound tourism⁹. On the other hand, the travel agencies mainly deal with outbound tourism,¹⁰ i.e., they take domestic tourists outside their place of residence. This major difference between the two sub-corpora has an effect on the nature of the texts and the choice of amenities, services and attractions to be presented to the potential

⁹ See <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Tourism>.

¹⁰ See <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Tourism>.

clients. The main idea was to compile a collection of texts that would be more neutral and objective in character, and not affected by an individual style of writing, so personal blogs that were part of the websites were not included in the corpora. All the names of hotels, cities, amenities, such as restaurants, museums, hyperlinks, graphs, tables, etc. were excluded from the corpora. The recurring themes in the corpora were the result of a specific season when the corpora were compiled, which was the summer season of 2019.

3.3 Research Procedure

The first step was a quantitative analysis of morphosyntactic realizations of collocations in the two corpora, as well as an investigation into the nature of the relationship between the primary and secondary collocates in the analysed collocations. The primary collocate is the main constituent of a collocation that carries its meaning (Prčić 2008, 151), whereas the secondary collocate modifies the meaning of the primary collocate or clarifies it. In the corpus-based studies, the primary and the secondary collocates are referred to as *node* and *collocate* respectively (Sinclair 1966, 415). At this stage, collocations that appeared four or more times in the corpora were extracted to create the corpus and considered as key terminological collocations for the field of tourism promotion on the web. This criterion was taken from the software tool TermoStat Web 3.0,¹¹ used for extracting collocations from the English corpus. Since the two corpora are not the same in size, a normalized frequency of distribution per 10,000 words was calculated for the collocations. The collocations whose normalized frequency (nf) was 1 or greater than 1 were considered to be the most frequent key collocations in the corpora. For ETC, this meant that the absolute frequency distribution (f) for *noun + noun collocations* in this corpus was at least 10 or more, and for STC this meant that the collocations appeared at least nine times or more. This was followed by the qualitative analysis of the key collocations in terms of their lexical-semantic features and their functional and communicative role in the corpora. Hence, specific topic areas emerged in the corpus based on the most frequent key collocations, depicting cultural aspects of the analysed texts that were compiled for the corpora. Finally, similarities and differences between the English and Serbian corpora were juxtaposed.

3.3.1 Corpus Analysis Tools

The corpus analysis tool TermoStat Web 3.0 developed by Patrick Drouin from the University of Montreal was used to extract the collocations in the English corpus. This tool was chosen because it lists *noun + noun* and *adjective + noun collocations* in the so-called technical corpus¹² after its automatic comparison with the non-technical corpus. Furthermore, typical combinations of adjectives and nouns with prepositions can also be sought for. All collocations that appear four or more times in the technical corpus are considered as terminological units and part of the specialized corpus (Drouin 2003). Figure 1 shows how the results are presented in the tool, showing the absolute frequency, log likelihood score, and type of collocation. Both singular and plural forms are calculated together.

¹¹ See <http://termostat.ling.umontreal.ca/>.

¹² Technical corpus is a term used by Drouin (2003) for the specialized corpus that was uploaded for the analysis.

Corpus >> korpus_engleski_1 DraganaVV | A

Résultats

Liste des termes **Recherche** Statistiques Structuration Régrammes

Candidat de regroupement	Fréquence	Score (Log Likelihood) ⬆	Variantes orthographiques	Matrice
sandy beach	12	112.96	sandy beach sandy beaches	Adjectif Nom
spectacular view	8	81.01	spectacular view spectacular views	Adjectif Nom
national park	9	67.52	national park national parks	Adjectif Nom
rock formation	7	66.85	rock formation rock formations	Nom Nom
low tide	7	60.11	low tide	Adjectif Nom
live music	6	52.04	live music	Adjectif Nom
tea room	6	50.35	tea room tea rooms	Nom Nom
seaside town	6	50.35	seaside town seaside towns	Nom Nom
natural beauty	6	46.41	natural beauty	Adjectif Nom
boat trip	5	46.19	boat trip boat trips	Nom Nom
local produce	5	43.99	local produce	Adjectif Nom
minute train	4	43.64	minute train	Nom Nom
beautiful beach	4	43.64	beautiful beach beautiful beaches	Adjectif Nom
bird life	4	43.64	bird life	Nom Nom
seaside resort	6	41.15	seaside resort seaside resorts	Nom Nom
train journey	4	38.65	train journey train journeys	Nom Nom
university city	4	38.65	university city university cities	Nom Nom

FIGURE 1. Partial view of the results of key collocations obtained by TermoStat Web 3.0.

Since TermoStat Web 3.0 is not available for Serbian, another software tool was used for the analysis of the Serbian Tourism Corpus called AntConc (Anthony 2017). There is a huge difference between these tools in terms of their technical characteristics, which, however, do not influence the obtained results. In our analysis we explored all nouns in the Wordlist for the Serbian Tourism Corpus, which also meant going through all paradigmatic forms of these words. The procedure of extracting collocations with AntConc is much slower because the researcher needs first to choose the Wordlist option, then deliberately search for particular key words and finally to look for secondary collocates and list concordances for them. The concordance option is formatted by the principle “key word in context” (KWIC) and it is very helpful for eliminating inadequate examples because it shows the actual context where the collocation appears in the text. Firstly, word combinations that contained foreign words in their original form were eliminated from the analysis, such as *a la carte restoran*, *deluxe soba*, *fitness centar*, *dress code*, etc. Another group comprised word combinations where a foreign word in the attributive function was adapted into Serbian but actually became a so-called prefixoid and formed a compound or a semi-compound noun (Klajn 1978, 188), for example, *aperitiv-bar*, *avio-karta*, *avio-prevoz*, *etno-selo*, *etno-park*, *snek-bar*, etc.

4 Results

In the next few paragraphs, we will look into the final results of the corpus analysis after the inadequate examples were filtered out. These were eliminated because they did not meet the proposed definition of collocations, or the word class was not the right one for this study. For instance, *host royal honeymoon* was extracted as *noun* + [*adjective* + *noun*] collocation by the software tool, but the context clearly showed that the word *host* was a verb in this example. Another combination that was excluded from the analysis was *sun bed*, which is actually written as one word and is clearly a compound.¹³ Although software

¹³ See <https://www.macmillandictionary.com/dictionary/british/sunbed>.

tools were used in extracting collocations, in a few instances, final decisions were made after consulting the *Online Oxford Collocation Dictionary*¹⁴ or *Oxford Collocations Dictionary for Students of English* (2002), as well as monolingual online dictionaries, such as the *Macmillan Dictionary*¹⁵ and *Oxford Learner's Dictionary*¹⁶. The treatment of collocations in the paper is based on phraseological and corpus studies approaches as mentioned before (cf. Cowie 1981; Sinclair 1991; Van Roey 1990; Granger 1998; Howarth 1998; Geeraerts 2010, etc.) and the treatment of collocations in the *Oxford Collocations Dictionary for Students of English* (OALD). OALD applies a pragmatic approach to collocations based on the most typical use of language (2002, viii) and includes collocations ranging from the rather weak ones to the most restricted ones, which were extracted from the 100-million-word British National Corpus. Finally, the software tools used in the paper for extracting collocation were considered as reliable, especially TermStat Web 3.0, which automatically extracted collocations based on corpora comparison, log-likelihood score and mutual information score (Drouin 2003), so the results obtained from this software tool were considered as reliable with some minor exceptions.

4.1 Noun + Noun Collocations in English Tourism Corpus (ETC)

As it is illustrated in Table 1, a total of 102 different key terminological collocations were found in the English Tourism Corpus (ETC). To refer back to the research design, key terminological collocations are those that appeared four or more times in the corpus.

TABLE 1. Total number of different key terminological noun + noun collocations in ETC.

Collocation Subtype	F	nf
N + N	93	9.43
[N + N] + N	1	0.1
N + [N+N]	1	0.1
N+ preposition 'of' +N	7	0.71
Total	102	10.35

The most frequent subtype is [N + N], i.e., the binary structure of collocations. Ninety three different collocations of this subtype were found in the corpus, with a normalized frequency of 9.34. Structurally, in English [N + N] collocations, the secondary collocate has attributive meaning and is positioned before the noun which is the primary collocate or the node. The secondary collocate has a modifying function similar to adjectives in adjective + noun collocations.

In the group of key [N+N] collocations with binary structure, a total of 23 most frequent key collocations were found, which therefore occurred 10 or more times in the corpus, i.e., whose normalized frequency per 10,000 is one or more. In this group, we will mention the first 10 most frequent collocations extracted by the TermStat Web 3.0: *buffet restaurant* (nf 4.16), *water sports* (nf 3.75), *sea view* (nf 3.04), *afternoon tea* (nf 2.74), *golf courses* (nf 2 , 54)

¹⁴ See <http://www.freecollocation.com/search?word=room>.

¹⁵ See <https://www.macmillandictionary.com/>.

¹⁶ See https://www.oxfordlearnersdictionaries.com/definition/english/next-door_2.

pool bar (nf 2.13), *dress code* (nf 1.93), *sun terrace* (nf 1.93), *boat trips* (nf 1.83) and *street food* (nf 1.62). As mentioned before, although the collocations were automatically extracted by TermoStat Web 3.0, some of the results were also checked in the *Oxford Collocations Dictionary for Students of English* (Dignen, Crowther, and Lea 2002). For example, *afternoon tea* might be considered as a free combination, but it is actually a meal which has a significant impact on the British way of life, and it has its own entry in the above-mentioned dictionary of collocations (Dignen, Crowther, and Lea 2002, 15). Moreover, it can also be found as a separate entry in online dictionaries¹⁷. The collocation *dress code* can be found as a separate entry in dictionaries¹⁸, so it might be analysed as a compound noun, but it is also found in the dictionary of collocations (Dignen, Crowther, and Lea 2002, 240).

In addition to binary collocations, only two extended collocations were found in the group of key collocations, hence appearing four or more times in the corpus. The first type is [N+N]+N, where we have a collocation [N+N] which is a modifier to the noun that follows (e.g., UK *airport lounge access* for Premium Club guests). The second type found in the corpus has a different structure with [N+N] collocation being a node preceded by another noun that modifies it (e.g., *daytime activity program*). The most frequent type of extended collocations is [N+prep. "of" + N] with seven different ones having been found in the corpus. Here are two examples:

- (1) A beautiful *stretch of coastline* offering varied attractions including stunning beaches like...
- (2) You get more than *a mile of sand* at Maleme Beach.

Finally, TermoStat Web 3.0 did not identify the most frequent binomials such as *hot and cold buffet*, *night and day*, *pool and beach bar*, *pool and beach games*, *loungers and parasols*, which appeared in the corpus, so they were not included in this analysis.

4.2 Noun + Noun Collocations in Serbian Tourism Corpus (STC)

Unlike English noun + noun collocations, this collocational type in Serbian is different in morphosyntactic terms. The primary collocate in this type is the noun that is in the first position in the collocation, and the noun that follows is the secondary collocate, in the function of an incongruent attribute.¹⁹ Establishing a hierarchical connection between the constituents in this type of collocations in the Serbian language represents a certain hindrance, which will be discussed later. As can be observed from Table 2, the total number of key noun + noun collocations excerpted from the corpus is 73, which is much less than in the English corpus, which is congruent with the particularities of the Serbian language.

As expected, the most frequent collocations in this group have a binary structure, with a total of 55 different [N+N] collocations excerpted from the corpus. The criterion was the

¹⁷ See <https://www.oxfordlearnersdictionaries.com/definition/english/afternoon-tea>. and <https://dictionary.cambridge.org/dictionary/english/afternoon-tea>.

¹⁸ See <https://www.oxfordlearnersdictionaries.com/definition/english/dress-code?q=dress+code>.

¹⁹ In the context of N+N collocations in Serbian, the incongruent attribute is the post-modifying noun that has a different case than the noun it is modifying.

TABLE 2. Total number of different key terminological noun + noun collocations in STC.

Collocation Subtype	f	nf
N + N	55	6.28
N+ [Adj. + N]	6	0.68
N + [N+ N]	3	0.34
N+preposition+N	9	1.03
Total	73	8.34

same as for English collocations – that they appeared four or more times in the corpus. Nineteen most frequent key [N+N] collocations (with normalized frequency one or more per 10,000) were extracted from the corpus, hence their absolute frequency was nine or more in the corpus. The first 10 most frequent [N+N] collocations in STC are as follows: *cena aranžmana* (nf 8.34), *program putovanja* (nf 7.42), *organizator putovanja* (nf 6.62), *vrste ptica* (nf 2.63), *kupac aranžmana* (nf 2.06), *repcija hotela* (nf 1.71), *vreme leta* (nf 1.48) *rezervacija aranžmana* (nf 1.37), *rezervat prirode* (nf 1.37), *spomenik prirode* (nf 1.26).

There are three different subtypes of extended collocations which were included in the analysis. In the first subtype, there are six collocations in which the primary collocates (e.g., *izvor*, *predeo* and *organizator*) are postmodified with the [Adj.+N] collocation (e.g., *Vranjska Banja ima više izvora mineralnih voda*; *Organizator fakultativnih izleta* je lokalni partner turističke agencije). The next subtype contains three extended collocations with primary collocates *potvrda*, *moгуćnost* and *troškovi* being postmodified with [N+N] collocations (e.g., *Rok za potvrdu rezervacije aranžmana* je...; *troškovi organizacije putovanja*). The third subtype of extended collocations includes nine collocations with the primary collocate being postmodified with a preposition followed by a noun, which is actually the prepositional case construction (e.g., *noćenje sa doručkom*; *fotelja na rasklapanje*; *soba sa pogledom na more*).

5 Discussion

The basic communicative function of binary collocations of the noun + noun type is informative, since they present the facts about a destination, i.e., primarily the offer within the accommodation complex and in its surroundings. In a broader context, they are further modified by adjectives that express attitude, emotion, or opinion, which is beyond the scope of this study.

In the English Tourism Corpus (ETC), binary collocations of the [N+N] subtype show a rather stable connection between collocates and are often mono-referential, i.e., they express a term that has a unique meaning. The replacement of the secondary collocate would lead to the loss of that specialized meaning, as in *bell tower*, *coral reef*, *craft beer*, *golf courses*, *home game*, *mountain range*, *nature reserve*, *patron saint*, *resort centre*, *sun lounger*, *steam room*, *souvenir shop*, etc. We can also find more flexible connections, with a greater possibility of replacing the secondary collocate and maintaining the meaning slightly modified (e.g., *pool bar*, *cocktail bar*, *snack bar*, *rooftop bar*, *lounge bar*). Another example of an even more flexible connection is when the secondary collocate modifies a noun of a more general meaning (e.g., *pool area*, *play area*, *lounge area*, *lobby area*). Even though in the general language, there would

be a much wider range of such word combinations, in the specialized language they form a set of collocations that refer to key areas in hotels, and thus can be considered as key collocations for this particular context.

In the Serbian Tourism Corpus (STC), unlike in English, the most predominant connection between the constituents in [N+N] collocations is that the first noun is the primary collocate which is postmodified by the secondary collocate functioning as an incongruent attribute, most often in genitive case without a preposition, or in another case with a preposition. The primary collocate in these collocations is often a noun with an incomplete, general or broad meaning. For instance, a collocation whose primary collocate is a noun with an incomplete meaning can refer to *part-whole* relationship (e.g., *parče hleba*, *grupa navijača*) (Dražić 2014, 144). In such cases, the communicative focus is most often on the secondary collocate, so these collocations should be observed in the interaction of meanings of their constituents. This interpretation is supported by the literature and more thoroughly explained by Dražić (2014, 74), who states that “the idea of the ranking of its constituents, i.e., the existence and definition of the so-called main (key) word, largely neglects the unique collocation meaning” (translated into English by the author), which becomes very obvious when analysing noun + noun collocations in Serbian.

In some cases, the meaning and the communicative focus are clearly conveyed by the primary collocate (e.g., *kupac aranžmana*, *predstavnik agencije*, *rezervacija aranžmana*, *recepција hotela*, *organizator putovanja*) because of the situational and communicative context. These types of collocations in Serbian can often have a unique denotative meaning in the specialized context, and smaller collocability (e.g., *podnožje planine*, *padina planine*, *dolina reke*).

Furthermore, in STC, the primary collocate in [N+N] collocations can often be a noun that has an abstract or incomplete meaning, such as *vrsta*, *vreme*, *izbor*, *cena*, *pozicija* so the secondary collocate is needed for the concretization of meaning. Such nouns have a very broad and general meaning which is completed by the postmodifying noun adding a more concrete meaning (cf. Dražić 2014, 99). In addition, the communicative focus is on the secondary collocate and such collocations need to be observed as units and not to be analysed by the meaning of their individual constituents. This interaction of constituent meaning, and its concretization, is best observed in the examples from STC: *vrsta smeštaja*, *vrsta biljaka*, *vreme obroka*, *vreme večere*, *izbor (više) jela*, *(veliki) izbor hotela*, *pozicija sobe*, *cena aranžmana*, *cena izleta*. An incomplete meaning is also reflected as fragmentary, so the primary collocate is postmodified with particularizing genitive (e.g., *ostaci utvrđenja*, *ostaci crkve*).

In STC, it has been observed that there is an overlapping of collocations with similar meanings, such as *lokacija sobe* or *pozicija sobe*, even though there is a Serbian equivalent *položaj sobe*, which is not represented in the corpus. Such examples are the result of globalization of the language of tourism, which is under the strong influence of the English language. There are other examples of *noun + noun collocations* with similar meanings throughout the STC, such as *izbor hotela* and *hotel po izboru*, *pogled more* and *pogled na more*.

Considering N+N collocations in the two corpora, significant differences have been observed, starting from the total number of extracted binary collocations, i.e., 93 in ETC compared

to 55 in STC. As we can see from the data, the total number of N+N collocations in the Serbian corpus is much smaller, which shows that such collocations are not as frequent as in English. The two corpora also differ in terms of the morpho-syntactic, semantic and lexical characteristics of the collocations which have been discussed in the previous paragraphs. In English, it is common for nouns to be used in pre-modifying, attributive functions as secondary collocates, which is not the case in Serbian, where the first collocate is always the primary collocate or the node in N+N collocations. Moreover, some of the most frequent *noun + noun collocations* from ETC have *adjective + noun collocations* as translation equivalents in STC (e.g., *summer months* – *letnji meseci*, *city life* – *gradski život*). Another difficulty is when the translation option is a word or a phrase without a transparent meaning, e.g., *buffet* translates into Serbian as *švedski sto*, so the collocation *buffet breakfast* would be translated into Serbian as *doručak na bazi švedskog stola*. Such instances could be challenging for the English language learners whose native language is Serbian and would have to be addressed in teaching.

5.1 Predominant Topics in the Corpora

If we observe the collocations in both corpora, we can easily conclude that texts describe a tourist offer for the summer season which can be corroborated with some examples from the ETC: *large outdoor swimming pool*, *sun terrace with sun loungers and umbrellas*, *children's pool*, *water sports*, *boat trip*. The corpus also provides an insight into the cultural aspects of living in the described destination, as evident in examples 3–6 below:

- (3) *Daily afternoon tea* and biscuits from 4–5.30pm.
- (4) Enjoy one of the most beautiful *train journeys* in Britain, Glasgow.
- (5) Being a *fishing town*, Cala Bona really delivers on the *seafood front*.
- (6) Tuck into innovative fare on the city's trendy *street food scene* or treat yourself to something really special.

What can be observed from the examples are some of the essential elements of the British way of life – enjoying afternoon tea as a special meal during the day, travelling by train, fishing and enjoying seafood.

In the ETC, there are four predominant topics in the analysed texts, such as *additional amenities and facilities* (e.g., *beach bar*, *pool area*, *boat trip*, *golf course*), *local cultural elements* (e.g., *city centre*, *evening entertainment*, *street food*), then collocations describing *food and beverage arrangements in the hotel* (e.g., *buffet breakfast*, *afternoon tea*, *cocktail bar*, *snack bar*), and the last group refers to some *specific hotel accommodation features* (e.g., *lounge area*).

If we look at the predominant topics in the STC, it is easy to see that there is a slight shift from the English corpus. In addition to some similar topics, a large part of STC is made of collocations pertaining to the *general organization of the trip* (e.g., *(opšti) uslovi putovanja*, *program putovanja*, *potvrda rezervacije aranžmana*), then *travel arrangements* (e.g., *vreme leta*, *sastanak putnika*, *dolazak u hotel*), and *payment arrangements* (e.g., *kupac aranžmana*, *cena aranžmana*, *uplata akontacije*). Furthermore, another predominant group is concerned with

the *description of natural-geographical characteristics* (e.g., *vrste ptica, podnožje planine, rezervat prirode, spomenik prirode*).

In total, there were seven different topics covered in the Serbian corpus, whereas only four in the English corpus. The topics that were the same covered accommodation descriptions and hotel amenities which, in ETC, included a sub-topic for food and beverages, as well as segments that described amenities outside the hotel or in the vicinity of the tourist resort. On the one hand, STC was more informative in character and also more focused on some practical aspects of travelling, which included descriptions of the legal details related to the organization and realization of the travel package and payment methods. This was found in the sub-corpora extracted from the travel agency website that deals with outbound tourism. On the other hand, the part of STC that dealt with inbound tourism was dominated by those topics related to the natural and geographical characteristics of the region, as well as cultural and historic elements. As such, tourism professionals in Serbia used the tourism organization website to promote the country in the most comprehensive way they could, since Serbia is still an emerging destination on the tourism market. However, Great Britain is already a well-established travel destination, so the promotional texts were more focused on the amenities and activities.

6 Conclusion

Linguistic contrastive analysis of lexical collocations in the language of a specific profession, especially in the context of custom-made corpus analysis, provides new insights and deepens the knowledge of morphosyntactic and lexical-semantic characteristics of the compared languages, creating an opportunity for observing some universal features as well as the differences between them. In this way, the lexicon of one language is systematized and a more detailed insight into the semantic relations between the lexical units of the same language is provided. This makes it easier to compare the same systems and subsystems of different languages.

Considering the results and discussion presented above, it is evident that there are some significant differences in *noun + noun collocations* in English and Serbian in terms of their morphosyntactic characteristics as well as some communicative features, which have been discussed in the previous chapter. The methodology of corpus analysis and collocational analysis presented here can contribute to more efficient language teaching and learning, because in this way both key collocations and key vocabulary can be selected and included in data-driven direct teaching methods. This is important in adult or academic foreign language courses that are often shorter and have fewer teaching hours per semester. In this way students' attention will be drawn to the similarities and differences in the two languages, and the emphasis would be on the ways how to assimilate adequate and appropriate collocations in L2.

In the context of ESP and EAP, custom-made corpora provide an insight into the domain specific texts, so the teachers and learners can gain a deeper understanding of lexical features and terminology related to a specific field of study, such as tourism studies. Since ESP classes at the tertiary level are often limited by the number of teaching hours, teachers can focus on specific language issues by creating custom-made corpora and using corpus analysis tools,

or in this case choose key terminological collocations related to the destination promotion to be taught through direct teaching methods. Students' attention will be drawn to specific language issues and the specific context-related jargon. Furthermore, the relationship between cultural and professional communication can be discussed in class by exploring the underlying topics in the promotional texts. They could discuss the impact of the specific language use on tourism promotion and focus on diverse cultural values promoted by tourism texts. Finally, as a step further, teachers can instruct students how to use custom-made corpora and corpus analysis tools to explore specialist texts, which can be accompanied by student interviews to see the effects of such assignments (Vuković Vojnović 2020). Although it should be noted that this approach has limitations regarding the number of teaching hours per course and the actual language proficiency of the students.

Further developments of this study should be focused on the creation of glossaries (cf. Gajšt 2011; Paradiž 2021) and lists of key terminological collocations in the field of tourism in English and Serbian, and other relevant languages. This would contribute to the systematization of vocabulary knowledge in this area, more high-quality texts in the field of tourism, and it would improve the teaching of foreign languages at the academic level. In the context of the language of tourism, students or domestic experts in the field of tourism would get acquainted with the nature of the language of tourism in Great Britain, and gain insight into the effects these texts have on potential tourism clients.

References

- Ackermann, Kristen, and Yu-Hua Chen. 2013. "Developing the Academic Collocation List (ACL) – A Corpus-Driven and Expert-Judged Approach." *Journal of English for Academic Purposes* 12: 235–47. <https://doi.org/10.1016/j.jeap.2013.08.002>.
- Bahns, Jens. 1993. "Lexical Collocations – A Contrastive View." *ELT Journal* 47 (1): 56–63. <https://doi.org/10.1093/elt/47.1.56>.
- Bhatia, Vijay K. 1993. *Analyzing Genre: Language Use in Professional Settings*. London: Longman.
- . 2002. "Applied Genre Analysis: A Multi-perspective Model." *Iberica* 4: 3–19.
- Cobb, Tom, and Alex Boulton. 2015. "Classroom Applications of Corpus Analysis." In *Cambridge Handbook of English Corpus Linguistics*, edited by Douglas Biber, and Randi Reppen, 478–94. Cambridge: Cambridge University Press.
- Coxhead, Averil. 2000. "A New Academic Word List." *TESOL Quarterly* 34 (2): 213–38. <https://doi.org/10.2307/3587951>.
- Croft, William, and David Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Cruse, David Alan. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Dann, Graham M. S. 1996. *The Language of Tourism: A Sociolinguistic Perspective*. Wallingford: CAB International.
- Dignen, Sheila, Jonathan Crowther, and Diana Lea, eds. 2002. *Oxford Collocations Dictionary for Students of English*. Oxford: Oxford University Press.
- Đorđević, Radmila. 1989. *Engleski i srpskohrvatski jezik. Kontrastivna gramatika imeničke grupe*. Beograd: Naučna knjiga.
- Dražić, Jasmina. 2014. *Leksičke i gramatičke kolokacije u srpskom jeziku*. Novi Sad: Filozofski fakultet.
- Drouin, Patrick. 2003. "Term Extraction Using Non-Technical Corpora as a Point of Leverage." *Terminology* 9 (1): 99–117. <https://doi.org/10.1075/term.9.1.06dro>.
- Firth, John Rupert. 1957. "A Synopsis of Linguistic Theory, 1930–1955." In *Studies in Linguistic Analysis, Special Volume of Philological Society*, edited by John Rupert Firth et al., 1–32. Oxford: Basil Blackwell.

- Gajšt, Nataša. 2011. "A Glossary-Based Approach to ESP Terminology Acquisition." *Inter Alia* 2: 21–41. <http://www.sdutsj.edus.si/InterAlia/2011/Gajst.pdf>.
- Gavioli, Laura. 2005. *Exploring Corpora for ESP Learning*. Amsterdam: John Benjamins.
- Geeraerts, Dirk. 2010. *Theories of Lexical Semantics*. Oxford: Oxford Linguistics.
- Gotti, Maurizio. 2008. *Investigating Specialised Discourse*. Bern: Peter Lang.
- Granger, Sylviane. 1998. "Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae." In *Phraseology: Theory, Analysis and Applications*, edited by Anthony P. Cowie, 145–60. Oxford: Clarendon Press Oxford.
- Granger, Sylviane, and Fanny Meunier, eds. 2008. *Phraseology. An Interdisciplinary Perspective*. Amsterdam: John Benjamins.
- Halliday, M.A.K. 1966. "Lexis as a Linguistic Level." In *In Memory of J.R. Firth*, edited by Charles Ernest Bazell, John Cunnison Catford, M. A. K. Halliday, and Robert Henry Robbins, 148–62. London: Longman.
- Howarth, Peter Andrew. 1998. "The Phraseology of Learners' Academic Writing." In *Phraseology. Theory, Analysis, and Applications*, edited by Anthony P. Cowie, 161–86. Oxford: Clarendon Press Oxford.
- Hyland, Ken, and Feng K. Jiang. 2017. "Is Academic Writing Becoming More Informal?" *English for Specific Purposes* 45: 40–51. <https://doi.org/10.1016/j.esp.2016.09.001>.
- Jaworski, Adam, and Annette Pritchard. 2005. *Discourse, Communication and Tourism*. Clevedon: Channel View Publications.
- Kang, Ning, and Qiaofeng Yu. 2011. "Corpus-Based Stylistic Analysis of Tourism English." *Journal of Language Teaching and Research* 2 (1): 129–36. <https://doi.org/10.4304/jltr.2.1.129-136>.
- Klajn, Ivan. 1978. "O prefiksoidima u srpskohrvatskom jeziku." *Naš jezik* XXIII (15): 187–98. Beograd: Institut za srpskohrvatski jezik.
- Lam, Peter Y. W. 2007. "A Corpus-Driven Lexico-Grammatical Analysis of English Tourism Industry Texts and the Study of its Pedagogic Implications in English for Specific Purposes." In *Corpora in the Foreign Language Classroom*, edited by Encarnacion Hidalgo, Luis Quereda, and Juan Santana, 71–90. Leiden: Brill. https://doi.org/10.1163/9789401203906_006.
- Liu, Jia, and Lina Han. 2015. "A Corpus-Based Environmental Academic Word List Building and its Validity Test." *English for Specific Purposes* 39: 1–11. <https://doi.org/10.1016/j.esp.2015.03.001>.
- Maci, Stefania Maria. 2010. *The Language of Tourism*. Bergamo: CELSB.
- McCarten, Jeanne. 2007. *Teaching Vocabulary: Lessons from the Corpus, Lessons for the Classroom*. Cambridge: Cambridge University Press.
- McEney, Tony, and Richard Xiao. 2010. "What Corpora Can Offer in Language Teaching and Learning?" In *Handbook of Research in Second Language Teaching and Learning* 2, edited by Eli Hinkel, 364–80. London: Routledge.
- Nation, I. S. Paul. 2007. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nesselhauf, Nadja. 2005. *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.
- Palmer, E. Harold. 1933. *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
- Paradiž, Martina. 2021. "CoLecter KIN: Mobile Application for Collaborative Bilingual Glossary Compilation in the ESP Classroom." *ELOPE* 18 (1): 91–106. <https://doi.org/10.4312/elope.18.1.91-106>.
- Prčić, Tvrtko. 2008. *Semantika i pragmatika reči*. Novi Sad: Zmaj.
- Schmitt, Norbert. 2008. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Sinclair, McHardy John. 1966. "Beginning the Study of Lexis." In *In Memory of J. R. Firth*, edited by Charles E. Bazell, John Cunnison Catford, M. A. K. Halliday and Robert Henry Robbins, 410–30. London: Longman.
- . 1987. "Collocation: A Progress Report." In *Language Topics: Essays in Honour of Michael Halliday II*, edited by Ross Steele and Terry Threadgold, 319–32. Amsterdam: John Benjamins.
- . 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Swales, John M. 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

- Thurlow, Crispin, and Adam Jaworski. 2010. *Tourism Discourse – Language and Global Mobility*. London: Palgrave Macmillan.
- Urry, John. 1995. *Consuming Places*. London: Routledge.
- Van Roey, Jacques. 1990. *French-English Contrastive Lexicology. An Introduction*. Louvain-la-Neuve: Peeters.
- Vuković Vojnović, Dragana. 2011. “‘Low Budget’ and ‘High Season’-Delve into the ‘Vast Ocean’ of Collocations Related to the Language of Tourism.” *Primenjena lingvistika danas – između teorije i prakse – Zbornik radova sa Trećeg kongresa*, 397–406. Beograd: Društvo za primenjenu lingvistiku Srbije, Filozofski fakultet u Novom Sadu, Filološki fakultet u Beogradu.
- . 2020. “Enhancing ESP Vocabulary Practice with Online Corpus Analysis Tools.” In *Contemporary Research in Foreign Language Teaching and Learning*, edited by Dana di Pardo Leon Henri and Bhawana Jain, 163–83. Newcastle upon Tyne: Cambridge Scholars Publishing.
- West, Michael. 1953. *A General Service List of English Words*. London: Longman.
- Xue, Guoyi, and I.S.P. Nation. 1984. “A University Word List.” *Language Learning and Communication* 3 (2): 215–29.

Corpus – Electronic Sources

- Kontiki Travel Agent: <https://www.kontiki.rs>.
- Thomson-Tui Travel Agency: <https://www.tui.co.uk>.
- Tourism Organisation of Great Britain: <https://www.visitbritain.com>.
- Tourism Organisation of Serbia: <https://www.serbia.travel>.

Electronic Tools for Corpora Analysis

- Anthony, L. 2017. *AntConc* (Version 3.5.0). Tokyo: Waseda University. <https://www.laurenceanthony.net/software>.
- TermoStat Web 3.0. http://termostat.ling.umontreal.ca/index.php?lang=en_CA.