

44*ADJECTIVE-NOUN COLLOCATIONS IN TOURIST ADVERTISING BROCHURES ABOUT ISTRIA: A CORPUS-BASED TRANSLATION STUDY

JELENA GUGIĆ

Juraj Dobrila University of Pula, Faculty of Educational Sciences, Pula, Croatia.
E-mail: jelena.gugic@unipu.hr

Abstract The aim of this research is to determine the collocation strength and contrastively analyse adjective-noun collocations in tourist advertising brochures about Istria found on the official web-page of the Croatian Tourist Board assuming that collocations found in the brochures will be stronger, i.e. more typical, and that most collocations in the English language will be direct translation equivalents of the Croatian collocations. The research starts by describing the position and importance of tourism in the overall European and Croatian industry. It then continues with a definition of collocations, their importance for smooth communication, and the description of their different types, highlighting the problems encountered with their proper translation. The empirical part presents the corpus-based methodology applied and offers the analysis results which confirm the tested assumptions. In the end, the study offers suggestions about how to overcome difficulties in the acquisition and use of collocations in the tourism discourse.

Keywords:

congruent
collocations,
non-congruent
collocations,
corpus-based
methodology,
tourism,
typicality

1 Introduction

Tourism as an industry has a leading role in the EU economy (EP, 2021). The Republic of Croatia has been a member of the EU for eight years. It is a country whose economy mostly relies on the tourism industry (15% of the total GDP). The Croatian National Tourist Board (HTZ, 2021), as the leading body taking care of the promotion of Croatia on the tourism market, has launched numerous initiatives to attract potential visitors. In the last few years, a “product line” of videos and brochures starting with the phrase *Croatia full of...* has caught the eye and raised the interest of travellers. The registers of the promotional materials tackle numerous areas of discourse such as nature and eco-tourism, history, food and drink, health, business, and narration, each one being marked by specific vocabulary. Numerous research studies have dealt with the linguistic analysis of the language of tourism (Peeters, 2007; Bednarek & Bublitz, 2007; Vučković-Vojnović & Nićin, 2012; Rata, 2012). In an interesting piece of research conducted by Rata in 2012, 232 different English noun phrases designating different types of tourism were analysed with the aim of providing Romanian equivalents for English adjectival tourism terms, thus showing the richness of this register’s vocabulary.

Given the importance of tourism for Istria and the Republic of Croatia in general, it is not surprising that great efforts and investments have been made into attracting Croatian citizens to join this sector. Considering that top-rate tourism would hardly be achievable without good knowledge of foreign languages, this research investigates the use of a more difficult linguistic segment which shows one’s proficiency in the overall knowledge of a language (Moehkardi, 2002; Nesselhauf, 2003; Košuta, 2012), namely adjective-noun collocations found in the brochures about Istria made available to the public on the web-page of the Croatian National Tourist Board and their translations in the English language brochures published by the same organisation.

2 Theoretical background

2.1 The definition of collocations

Words as lexical units bear a meaning which is inherent to them. Even in cases when words bear more than one meaning, a native speaker, or a proficient speaker of a foreign language, can understand the meaning of a word without too much difficulty. Thus Murphy (2010, p. 22) writes about “native speaker intuition.” However, words often combine, they seek for each other, and co-occur. According to Cowie (2001) word combinations can be divided into semantic and pragmatic combinations. Semantic combinations are further divided into collocations and idioms, whereas pragmatic combinations are divided into proverbs and routine formulae (where he also differs them from speech formulae (Ibid.).

According to the Oxford English Dictionary, collocations were first mentioned at the beginning of the 16th century. In Barnbrook, Mason & Krishnamurthy (2013, p. 6) one can find the definition of the verb form *collocate* from 1513 taken from the OED: 1. a. trans. To place side by side or in some relation to each other; to arrange. b. To set in a place. The online Cambridge English Dictionary defines collocations as “the combination of words formed when two or more words are often used together in a way that sounds correct.” Nattinger and DeCarrico (1992, p.21, as cited in Moehkardi, 2002) defined collocations as “strings of words that seem to have certain ‘mutual expectancy’, or a greater-than-chance likelihood that they will co-occur in any text.”

Collocations are linguistic units which enable smooth and economical communication (Blažević & Košuta, 2016). They are indicators of communication competence since the ability to use collocations typical for a foreign language is often considered as a high level of knowledge of a language. Having that in mind, and considering the education system which includes teaching foreign language vocabulary as isolated lexical units instead of semantic units (Stojić & Murica, 2010), a higher level of oral or written competence among foreign language learners can hardly be expected, especially when it is known that as less fixed, collocations are often subject to “cross-linguistic influence” (Mustapić & Malenica, 2013, p. 209), or “native language interference” (Marton, 1977, p. 53). The survey conducted by Alharbi (2017) among Saudi Arabian university students and teachers shows that

teaching collocations has always been neglected and that teachers themselves do not work on raising collocational awareness. Learners face problems linked to combining two or more words together, and examples such as *feeble tea*, *laugh broadly*, *climb a horse* “show a lack of collocational competence in the learner” (Bahns, 1993, p.56). This problem can be observed at the level of production, especially in the field of foreign language acquisition and translation (Marton, 1977; Nesselhauf, 2003).

Collocations are formed by a base (B) and a collocate (C) (e.g. *pretty* (C) *girl* (B); *handsome* (C) *boy* (B); *make* (B) *the bed* (C); *do* (B) *the homework* (C)). The base usually remains unchanged, whereas the collocate determines the meaning of the collocation (Košuta, 2012). Differently from idioms, the meaning of collocations can be derived from the meaning of the single units forming it. However, they are often difficult to remember. Various languages have a specific way of combining components, so by contrasting two languages one can determine the differences between them (Blažević & Košuta, 2016), and thus identify the problematic structures which cause difficulties to learners of a language. Their inability to form collocations properly is seen as a lack of accuracy and fluency (Van Der Meer, 1998; Stojić & Murica, 2010; Alharbi, 2017).

2.2 Types of collocations

There are two types of collocations, namely grammatical and lexical (Benson, Benson & Ilson, 1997; Rata, 2012). The difference between them is that grammatical collocations are combinations of a noun, adjective or verb and a particle (preposition, infinitive, or clause). For instance, *depend on*, *encourage (someone) to do (something)*, *be afraid that (something could happen)*, etc., whereas lexical collocations are composed by two content words (nouns, adjectives, verbs or adverbs in various combinations). According to Stojić and Murica (2010, p.116) the Croatian language has the possibility to form the following types of lexical collocations: verb + noun (*tražiti pravdu*), adjective + noun (*mladi krumpir*), noun + verb (*pas laje*), noun + noun (*prstohvat soli*), adverb + adjective (*smrtno ranjen*) and adverb + verb (*oštro kritizirati*). According to Petrović (2007, as cited in Košuta, 2012) the most common types of collocations in the Croatian language are adjective + noun, verb + noun and adverb + verb. When it comes to the English language, the possible lexical collocation combinations are as follows (Benson, Benson & Ilson, 1997): verb + noun (*write a letter*), adjective + noun (*hard book*), noun + verb (*bees buzz*), adverb + adjective (*deeply*

moved), verb + adverb (*argue heatedly*) and noun-of-noun (*a pack of dogs*). Corpas Pastor (1995, 1996, as cited by Mendoza Rivera, Mitkov & Corpas Pastor, 2018) identified another type of collocations, namely the verb – preposition – noun combination (*take into consideration, jump to a conclusion*). Since in the general language, as well as in the language for special purposes (LSP) “adjectives, modifying, modulating or elaborating the meaning of nouns, play a relevant role in discourse” (Pierini, 2006, p.94), this study was concentrated on finding collocations and their strength or typicality in a context and on comparatively analysing the adjective-noun collocations in the aforementioned brochures.

In terms of collocations translation, it is important to find out what is the type of congruence among collocations of the source and target language. Numerous authors write about full, partial or zero congruence (a division applied in this study as well), or simply about congruent and non-congruent collocations (Marton, 1977; Bahns, 1993; Nesselhauf, 2003; Mustapić & Malenica, 2013; Blažević & Košuta, 2016; Alharbi, 2017; Hashemi & Eskandari, 2017), defining congruent collocations as those which are formed by identical or near-identical lexemes in L1 and L2, whereas non-congruent collocations contain lexemes which are not direct equivalents in the two analysed languages. According to Mustapić and Malenica (2013), for a collocation to be congruent, it should sound sufficiently natural, in both the source and target language. “Among the lexical problems of language learners, collocational errors are the most frequent, they are the most salient markers of non-nativeness” (Kiss & Horvath, 2015, p.167). Being afraid of making mistakes, second language learners tend to use collocations which are congruent in L1 and L2. Word combinations containing a more abstract or unfamiliar word usually show partial or zero congruence. The aforementioned studies also show that the influence of L1 on the translation and use of L2 collocations represents the most serious problem, and that students/translators rely on the ‘hypothesis of transferability’ (Bahns, 1993, p. 61) in their effort to produce correct L2 collocations.

3 Empirical research

3.1 Research aims

Regarding the importance of advertising for the tourist industry, and bearing in mind that adjectives are words describing nouns (and in the touristic sense these would be destinations, services, offers, gastronomy, etc.), the first aim was to examine collocation strength or typicality (the level to which the appearance of a certain word makes another word more probable to be found nearby) of adjective-noun collocations in tourist advertising brochures about Istria which can be found on the official web-page of the Croatian Tourist Board using the Log Dice score and to compare it to the strength of the same collocation in a general language corpus. It was assumed that the Log Dice score for the analysed collocations would be higher in the tourist brochures because of their typicality in the given context. On the other hand, they would be less typical in a general language context. The second aim of this research was to contrastively analyse the aforementioned collocations. The source language was Croatian, and the aim was to determine the translations of the found collocations into the English language, as well as to find out how collocations were translated and thus define them as fully or partially congruent, or not congruent at all. If the adjective is taken as the collocate and the noun it describes as the base, the research assumed that most collocations in the English language would be direct translation equivalents of the Croatian collocations, i.e. they would be fully congruent.

3.2 Methodology

The corpus-based approach in the study of collocations related to the tourism discourse has been a common method in recent times (Gerbig & Shek, 2007; Pierini, 2009; Kiss & Horv ath, 2015; Bla zevi c & Ko suta, 2016). As found by these authors, the corpus method is extremely applicable in the study of vocabulary and its subsequent use for various purposes such as language learning and grammar or dictionary compilation. For the purpose of this research, a total of 12 brochures found on the official web-page of the Croatian Tourist Board were downloaded, and then their parts regarding the Istria region were singled out. The research was conducted in a few steps. After the texts of the brochures in both Croatian and English were transformed in text form (cleared of pictures and other metadata called

‘noise’ (Fletcher, 2004, as cited by Pierini, 2006), they were uploaded to the concordancing tool *Sketch Engine*. A corpus of 24,982 tokens was created. After the corpus was compiled, the most frequently used nouns were extracted from the corpus and their function as the collocation base was analysed. Next, the adjective-noun combinations were also extracted and they represented the collocations to be analysed. Regarding the first aim of this research – the analysis of collocation strength – all the nouns in the Croatian language brochures which were singled out as the collocation base were analysed by the *Word Sketch* tool in *Sketch Engine*. It offers the possibility to analyse the collocation strength using the Log Dice statistic measure which is usually employed to test collocation typicality. The typicality of the found collocations was then compared to their typicality in the hrWaC corpus.

The Croatian corpus was then compared to the English in order to find the translation of the collocations under study. Although this could be seen as a small corpus, when the abundance of adjectives typically used to describe tourist destinations is considered, the findings in it can reflect the situation in the language of tourism in general. In the end, what was considered was the level of congruence between a Croatian collocation and its translation in English.

3.3 Results

3.3.1 The nouns

The first step in the analysis was to find the most frequently used nouns in the Croatian corpus. It produced a total of 1,399 nouns. The list was then cleared of all proper nouns (locations, wine varieties, wine growers, companies and local Istrian words which could not be translated into English) and unadapted loanwords. The list was narrowed down to 1,135 nouns. Only those occurring in the Croatian corpus ten times or more, namely 48 nouns, were considered for analysis. In this part of the analysis it was considered interesting to compare their frequency with the frequency of these nouns in a general language corpus. For this purpose, the hrWaC corpus (available online on the site <http://nlp.ffzg.hr/resources/corpora/hrwac/> as well as on *Sketch Engine*) was used as the reference corpus. It consists of 1,397,757,548 tokens extracted from texts taken from the web (the .hr domain) which is a good source of texts belonging to different styles and registers. Table 1 shows these nouns

and the frequency of their occurrence in the Croatian language brochures compared to the frequency of their occurrence in the hrWaC corpus.

Table 1: Collocation bases and their frequency in the Croatian language brochures compared to their frequency in the hrWaC corpus

Noun	Croatian language brochures			hrWaC	
	f	%		f	%
Vino	54	0.47		125,499	0.008
Grad	49	0.43		1,131,889	0.08
Obala	38	0.33		109,437	0.007
Ponuda	34	0.3		251,306	0.02
Gradić	34	0.3		22,407	0.001
Stoljeće	29	0.3		217,284	0.01
Tartuf	29	0.25		4,778	0.0003
Poluotok	28	0.24		18,232	0.001
Godina	28	0.24		4,076,110	0.29
More	27	0.24		272,473	0.02
Sorta	24	0.21		21,413	0.001
Unutrašnjost	23	0.2		37,549	0.002
Dio	22	0.19		1,267,701	0.9
Zemlja	21	0.18		875,667	0.06
Park	21	0.18		118,788	0.008
Ulje	20	0.17		152,259	0.01
Mjesto	20	0.17		1,194,473	0.08
Svijet	19	0.17		882,602	0.06
Regija	17	0.15		125,617	0.008
Staza	17	0.15		124,961	0.008
Priroda	17	0.15		179,837	0.01
Vrh	15	0.13		182,352	0.01
Gastronomija	15	0.13		4,936	0.0003
Šuma	15	0.13		92,734	0.006
Jaje	15	0.13		61,825	0.004
Otok	14	0.12		182,917	0.01
Ljepota	14	0.12		86,669	0.006
Povijest	13	0.11		274,843	0.02
Uvala	13	0.11		22,763	0.002
Posjetitelj	13	0.11		99,021	0.007
Amfiteatar	13	0.11		3,018	0.0002
Zaljev	12	0.1		22,773	0.002
Turizam	12	0.1		118,506	0.008
Odmor	12	0.1		105,931	0.007
Rijeka	12	0.1		20,971	0.001
Blizina	12	0.1		96,166	0.007
Područje	12	0.1		621,201	0.04
Gost	11	0.09		242,769	0.02
Dan	11	0.09		2,151,333	0.15
Ruta	11	0.09		17,243	0.001

Noun	Croatian language brochures			hrWaC	
	f	%		f	%
Odredište	11	0.09		22,270	0.001
Vrijeme	11	0.09		1,726,714	0.12
Put	10	0.08		1,570,603	0.11
Broj	10	0.08		838,099	0.06
Selo	10	0.08		166,839	0.01
Hotel	10	0.08		191,544	0.01
Brežuljak	10	0.08		7,035	0.0005
Kuhinja	10	0.08		3,665	0.0002

Source: Sketch Engine and <http://nlp.ffzg.hr/resources/corpora/hrwac/>

It has to be emphasized that in the analysis of the hrWaC corpus the words *rijeka* and *kuhinja* caused difficulties linked to their polysemy. The word *rijeka* has two meanings – ‘river’ and ‘name of a Croatian city’ – but in finding only the first meaning the problem was easily solved by extracting only those nouns starting with the lower case. The situation with the polysemy of the word *kuhinja* – meaning ‘a room in the house,’ ‘the furniture in such a room,’ and ‘cuisine’ – was harder to solve because there were 76,481 occurrences of the word. To get the most reliable result possible, the collocations with the node *kuhinja* (*građanska, pučka, tradicionalna, domaća, visoka*) were extracted from the Croatian brochures and exactly these collocations were searched for in the hrWaC corpus which gave, as can be seen in the table, the result of 3,665 occurrences or 0.0002% of the corpus (*građanska kuhinja* occurs 20 times, *pučka kuhinja* 2,593 times, *tradicionalna kuhinja* 356 times, *domaća kuhinja* 696 times, whereas the collocation *visoka kuhinja* was not found).

The most frequently occurring noun in the Croatian language brochures was *vino* (*wine*), and it occurred 54 times. However, the collocation base which produced the largest number of collocations, but occurred 34 times in the Croatian corpus, was the noun *ponuda* (*offer*) and it formed a total of 23 collocations. Regarding the fact that proper nouns found in the analysed brochures, and expected to be the most frequent, were omitted from analysis (such as Istria, Croatia, and various other toponyms), words such as *vino, ulje, tartuf, more, poluotok, ponuda*, etc. were expected to be in the frequency list since they are all semantically linked to the Mediterranean and its geographical characteristics.

3.3.2 Collocations typicality

The next step employed the use of the *Word Sketch* tool in *Sketch Engine* by which the most common adjective-noun collocations were extracted for the Croatian language. After checking the obtained collocations in their context (to find whether *Sketch Engine* made some mistakes in their extraction due to syntactic reasons), and omitting those for which the collocate was a proper noun (e.g. *istarski poluotok*, *francuska kuhinja*), the total number of collocations in the Croatian language was 257. Although some collocations were found in the text as not only binary structures, but multi-word units consisting of more than two elements (for instance, *ekstra djevičansko maslinovo ulje – extra virgin olive oil*), which is usually done to economise the use of vocabulary, for the purpose of this study, they were separated into binary units.

Further, the analysis of collocations comprised an insight into the typicality of their co-occurrence. Namely, the aim was to find the most typical collocator to each collocation base in the Croatian language brochures and compare their strength to the same collocation's strength in a general language corpus, namely the hrWaC (Table 5). The strength of English brochure collocations was not analysed because in this paper they were solely studied as translations of the Croatian collocations. For that purpose, the Log Dice score was used. It is a statistic measure of typicality based on the co-occurrence of the base and the collocate. Log Dice is standardised and fixed at its maximum value of 14 which makes it easy to interpret (the closer the value to 14, the more typical the collocation) and its results can be compared across corpora, irrespective of their size.

Table 2: Strength of the Croatian language brochures collocations compared to their strength in a general language corpus (hrWaC)

Croatian collocation	Brochure Log Dice	hrWaC Log Dice
Desertno vino	12.0	8.0
Antički grad	11.2	6.0
Istočna obala	12.4	10.2
Bogata ponuda	11.2	9.4
Srednjovjekovni gradić	12.0	8.3
Prošlo stoljeće	13.8	11.1
Bijel tartuf	12.9	6.7
Cijeli poluotok	12.0	3.7
Cijela godina	12.1	8.7
Kristalno more	12.4	7.4

Croatian collocation	Brochure Log Dice	hrWaC Log Dice
Autohtona sorta	11.5	10.7
Zelena unutrašnjost	12.5	/
Vršni/donji/razveden/bitani/zapušten/unutarnji/jugoistočni dio	11.0	Vršni /; Donji 9.3; razveden /; bitan 7.2; zapušten /; unutarnji 5.7; jugoistočni 5.9
Crvena/plodna zemlja	11.9	Plodna 7.0;
Nacionalni park	13.6	10.6
Maslinovo ulje	13.5	12.8
Romantično mjesto	11.4	/
Vanjski svijet	13.0	8.6
Uspješna regija	11.0	/
Biciklistička staza	12.5	11.1
Netaknuta priroda	13.3	10.6
Sam vrh	12.8	9.6
Tradicionalna gastronomija	11.4	4.2
Glasovita šuma	11.8	/
Domaće jaje	12.6	4.2
Obliznji otok	12.1	7.9
Prirodna ljepota	13.0	10.0
Bogata povijest	12.5	8.9
Skrovita uvala	12.2	8.1
Odvažan posjetitelj	13.4	/
Očuvan amfiteatar	11.5	6.9
Zmijolik zaljev	12.3	/
Obalni turizam	12.4	4.6
Aktivni odmor	13.1	9.0
Podzemna rijeka	14.0	6.7
Neposredna blizina	14.0	13.0
Nastanjeno/vinorodno područje	11.2	Nastanjeno /; vinorodno 5.4
Zahtjevan gost	13.0	6.8
Bistar/vedar dan	12.7	Bistar /; vedar 5.1
Nautička/biciklistička ruta	11.8	Nautička 4.4 ; biciklistička 9.4
Idealno odredište	12.3	6.3
Loše vrijeme	12.0	7.8
Plovidbeni/dišni put	12.2	Plovidbeni / ; dišni 9.2
Nemali broj	12.0	7.5
Slikovito selo	12.0	5.8
Luksuzan hotel	12.4	10.6
Visok brežuljak	11.2	2.0
Građanska kuhinja	12.0	3.5

Source: Sketch Engine

As expected, the Log Dice scores for the collocations given in Table 5 were lower in the hrWaC corpus than in the Croatian language brochures. This speaks in favour of the typicality of these collocations being higher in the touristic context than in a general language one. The linguistic setting is of crucial importance when it comes to the expectancy (i.e. frequency) of a word, or word combination, occurrence. This was confirmed by Gablasova, Brezina and McEnery (2017, p.167): "...variation in collocational strength due to linguistic setting should also be expected..." The comparison of the Log Dice scores in the two corpora shows that most of the collocations which are typical for the Croatian brochures can be considered typical in the hrWaC corpus as well (although, less typical). However, there is a number of them showing a great discrepancy (*cijeli poluotok, tradicionalna gastronomija, domaće jaje, obalni turizam, nautička ruta, visok brežuljak, građanska kuhinja*), and some which could not be found in a general language corpus like the HrWaC (*zeleno unutrašnjost, vršni/razveden/zapušten dio, romantično mjesto, uspješna regija, glasovita šuma, odvažan posjetitelj, zmijsolik zaljev, nastanjeno područje, bistar dan, plovidbeni put*). Such collocations can be considered typical and truly typical collocations for the language of tourism found in the Croatian language brochures.

3.3.3 Translation analysis

Concerning the translations of the nouns into English, it could be immediately noticed that one Croatian noun could be translated with more than one English noun which shows the difficulty that learners of English could face in accumulating the vocabulary necessary for more proficient communication. For instance, the Croatian word *grad* was translated as *town* and *city*; *obala* was translated as *coastline*, *coast*, *shore*, even *cliff*; *vrh* was translated as *peak* and *summit*; *šuma* became *woodland* and *forest*; *uvala* was translated as *cove* and *bay*; *unutrašnjost* became *hinterland* and *interior*; *mjesto* was a *place*, *town* and *location*; *staza* was translated as *route*, *trail*, *path* and *footpath*; *odmor* became a *holiday* and *vacation*; *područje* was translated as *area* and *region*.

All the collocations were marked in the original text and by conducting a comparative analysis, the translations of the previously determined collocations were also found and marked. Omitting those which had not been translated in any way (or whose translation was avoided), there were 260 possible collocations in English. However, when the collocations which were paraphrased or not translated at all are considered, the number of possible solutions in the English brochures rises to 276.

When collocational congruence is considered, most collocations and their translations show full congruence. More precisely, there were 195 collocations which were lexically and semantically congruent (e.g. *tradicionalna kuhinja* – *traditional cuisine*, *odabrani broj* – *chosen number*, *nautička ruta* – *nautical route*; *vedar dan* – *bright day*, *zahtjevan gost* – *demanding guest*). Since the number of obtained fully congruent collocations is large, Table 2 shows the Croatian collocations for the most productive base (*ponuda*) and their translations in English.

Table 3: List of Croatian collocations for the most productive base (*ponuda*) and their translations in English.

Croatian collocation	Translation in English
Bogata ponuda	Copious attractions; abundance of services; extensive facilities; rich (cultural) events; once without translation
<i>Turistička ponuda</i>	<i>Tourist offer; tourism offer</i>
Visokokvalitetna ponuda	Luxurious offer
Odlična ponuda	Excellent selection
<i>Gastronomska ponuda</i>	<i>Gastronomic offer</i>
Kulturna ponuda	Culture; cultural events
Agroturistička ponuda	Agro offer
<i>Respektabilna ponuda</i>	<i>Respectable offer</i>
Rekreacijska ponuda	Recreational facilities
<i>Sveobuhvatna ponuda</i>	<i>Comprehensive offer</i>
<i>(Pomno) osmišljena ponuda</i>	<i>Carefully-designed offer</i>
Zabavna ponuda	Entertainment programme
Razvijena ponuda	Wide range of services
<i>Raskošna ponuda</i>	<i>Luxurious offer</i>
<i>Primjerena ponuda</i>	<i>Exemplary offer</i>
<i>Kongresna ponuda</i>	<i>Congress offer</i>
Raznolika ponuda	Diverse attractions
Enološka ponuda	Delightful wines
Sportska ponuda	Sports facilities
Raznovrsna ponuda	Varied attractions
<i>Izvrсна ponuda</i>	<i>Outstanding offer</i>
<i>Morska ponuda</i>	<i>Marine offer</i>
Vrhunska ponuda	Exemplary offer

Source: own

It can be immediately noticed that of the 23 collocations formed with the most productive Croatian base *ponuda*, only 10 were fully congruent. However, the initial analysis showed that for all the Croatian collocations (N=257), there were 195 (75.87%) fully congruent English translations. The conclusion can be derived that the English translation of the Croatian brochures shows a high level of both accuracy and fluency (or collocational competence), and this is confirmed by the high percentage of fully congruent English collocations. The collocations presented in Table 1 which were not pointed out (bolded and italicised) belong to either the partially congruent or non-congruent group of collocations. Why the translators of the said brochures decided not to find fully congruent collocations in English can only be guessed.

Partial congruence was established in 49 cases. It should be further pointed out that the partiality was achieved in two ways: either the base remained the same (equivalent translation) and the collocate changed (*crno vino* was translated as *red wine*, not *black wine* (although it is correctly translated, it can be treated as partially congruent), *vrhunska ponuda* – *exemplary offer* instead of *top offer*, *zanimljiva povijest* – *specific history* instead of *interesting history*, *zavučeni zaljev* – *indented bay* instead of *tucked-away bay*), or the base changed retaining the same collocate (for instance, *poznat grad* – *well-known area* instead of *well-known city*, *strma obala* – *steep cliff* instead of *steep coast/shore*, *kulturna ponuda* – *cultural events* instead of *cultural offer*, *zelen dio* – *green countryside* instead of *green part*, *poznata zemlja* – *well-known area* instead of *well-known land*).

When it comes to the notion of zero congruence, it was manifested in the brochures in four ways. First, collocations were translated with other collocations which had completely different lexemes used for both the collocate and the base, but in the given context, they retained the original meaning (e.g. *bogata ponuda* was never translated as *rich offer*, but as *copious attractions*, *abundant services*, *extensive facilities*, *wellness opportunities*, *neotkrivena obala* was translated as *hidden gem*, not as *undiscovered coast*, *enološka ponuda* – *delightful wines* instead of *enological offer*). Sometimes only the base could be found while the collocate was missing (*uređena staza* – *footpath* instead of *landscaped footpath*, *razvijen grad* – *city* instead of *developed city*, *turističko mjesto* – *town* instead of *tourist town*), sometimes, on the other hand, only the collocate (*vršni dio* – *top* instead of *top part*). Another interesting feature of the collocation translations which could be regularly found in the analysed brochures is that the Croatian collocations were often not translated with collocations but paraphrased (*maslinarska*

regija, maslinarsko područje, cijeli svijet, davni dan, novo vrijeme). In the end, there were collocations which were not translated at all (*nekadašnji grad, pejzažni park, osebujan poluotok, plovidbeni put, registriran tartuf*). A similar finding was reached by Marton in 1977. He gave his advanced-level students a Polish text and their task was to translate it into English. The purpose of the test was to “elicit about twenty English conventional syntagms” and the results showed that students constructed grammatically correct expressions, but failed to sound natural. Those students, like the translators in this study, used different production strategies, most commonly avoidance, overgeneralisation and circumlocution.

What follows is the presentation of the translation context in cases when the translation of collocations was either paraphrased or avoided (Table 3 and Table 4).

Table 4: The context in cases when Croatian collocations were paraphrased in English

Collocation in the Croatian language with suggested congruent collocation	Solution in the English language - paraphrasing
Maslinarska regija – olive-growing region	<i>It is known as an area that produces outstanding olive oil and wine.</i>
Maslinarsko područje – olive-growing area	<i>The towns of Buje and Brtonigla are the centres of this famous area known for vines and olive cultivation.</i>
Cijeli svijet – the whole world	<i>...from around the world.</i>
Davni dan – long ago	<i>...this is why it has long been a popular stop for travellers.</i>
Novo vrijeme – recent times	<i>...recently becoming populated...</i>

Source: own

Table 5: The context in cases when the translation of Croatian collocations was avoided

Collocation in the Croatian language with suggested congruent collocation	Solution in the English language – not translated
Nekadašnji grad – former city	<i>Find out why Dvigrad simply disappeared from the face of the earth in the 17th century.</i>
Pejzažni park – landscape park	<i>Brijuni are renowned for its indented nature and the unspoiled Mediterranean vegetation...</i>
Osebujan poluotok – peculiar peninsula	<i>The island interior features miniature fairytale locations and green hilly scenery...</i>
Plovidbeni put - fairway	\emptyset
Registriran tartuf – registered truffle	<i>The largest white truffle ever found...</i>

Source: own

It is interesting to notice that *osebnjan poluotok* was translated as *island*, while in the case of *plovidbeni put*, the translator omitted the whole sentence. Both mistakes were probably unintentional.

4 Conclusion and further implications

“As a recurring phraseological phenomenon, collocations represent one of the most important aspects of LSP” (Mustapić & Malenica, 2013, p. 207). In terms of typicality, the collocations analysed in this research proved to be more typical in the tourist brochures than in the general language context confirming the thesis that the linguistic context plays a major role in the choice of vocabulary.

In their acquisition it is important to consider the level of congruence between a collocation in the source and target language. As this study shows, the largest number of Croatian to English collocations are fully congruent, and this is typical for all those frequently used. This confirms the understanding that congruent collocations are easier to learn, while special attention should be paid to teaching the non-congruent ones. The brochures analysed in this study were translated by, hopefully, professional translators who had mastered collocations as part of their higher (linguistic) education and this is why more than 75% of the collocations were correctly translated. However, whenever they were not sure about a collocation, and being aware of the possibility to sound unnatural, translators (and this can be reflected in second language learners) used different techniques to translate or even avoid the translation of collocations thus making them fall into the group of non-congruent ones.

Considering the importance of overall communication in all the fields of the tourism industry aiming at satisfied customers, it is obvious that attention should be paid to teaching collocations for better fluency and accuracy. It is necessary to expose second language learners to texts where collocations are used substantially so that the input they receive makes them acquire the foreign language in a more natural way. However, as learning them requires considerable mental effort, they should be consciously learned, memorised and reused in various contexts. Another important thing is to teach collocations as units, not as individual words, thus aiming at the prevention of mistakes occurring due to the influence L1 has on L2. Cross-linguistic comparisons can be a good starting point in the acquisition and proper use of

collocations in a foreign language, and translation analyses similar to the present one can be a good basis for the identification of linguistic elements disrupting the natural communication between speakers of two different mother tongues using one of them as a common language.

References:

- Alharbi, R. M. (2017). *Acquisition of Lexical Collocations: A Corpus-Assisted Contrastive Analysis and Translation Approach*. Newcastle: Newcastle University.
- Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal*, 47(1). Oxford University Press, 56–63.
- Barnbrook, G., Mason, O. & Krishnamurthy, R. (2013). *Collocation. Applications and Implications*. Palgrave Macmillan.
- Bednarek, M. & Bublitz, W. (2007). Enjoy! The phraseological culture of having fun. In: P. Scandera (Ed.), *Phraseology and Culture in English* (pp. 109–136). Berlin: Mouton de Gruyter.
- Benson, M., Benson, E. & Ilson, R. (1997). *The BBI Combinatory Dictionary of English. A Guide to Word Combinations (revised edition)*. Amsterdam: Benjamins.
- Blažević, N. & Košuta, N. (2016). Kolokacije u opisu turističke destinacije. *3rd International Conference Foreign Languages and Tourism, conference proceedings*. Opatija: Faculty of Tourism and Hospitality Management
- Cambridge dictionary (2021). Retrieved from <https://dictionary.cambridge.org/>
- Corpas Pastor, G. (1995). *Un Estudio Paralelo de los Sistemas Fraseológicos del Inglés y del Español*. Málaga: SPICUM.
- Corpas Pastor, G. (1996). *Manual de Fraseología Española*. Madrid: Gredos.
- Cowie, A. P. (2001). Speech formulae in English: Problems of analysis and dictionary treatment. In: G. van der Meer, & A. ter Meulen (Eds.), *Making sense: From lexeme to discourse* (pp. 1–12). Groningen: Center for language and cognition.
- Croatian National Tourist Bord / Hrvatska turistička zajednica (2021). Retrieved from <https://www.htz.hr/>
- European Parliament (2021). Retrieved from <https://www.europarl.europa.eu/>
- Fletcher, William H. (2004). Making the web more useful as a source for linguistic corpora. Retrieved from https://www.researchgate.net/publication/233649548_Making_the_Web_More_Useful_as_a_Source_for_Linguistic_Corpora
- Gablasova, D., Brezina, V. & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing and interpreting the evidence. *Language Learning*. Wiley Periodicals, 155–179.
- Gerbig, A. & Shek, A. (2007). The phraseology of tourism: A central lexical field and its cultural construction. In: P. Scandera (Ed.), *Phraseology and Culture in English* (pp. 303–322). Berlin: Mouton de Gruyter.
- Hashemi, M. R. & Eskandari, R. (2017). The learning of congruent and incongruent collocations utilizing dynamic assessment. *The Language Teacher*, 41(6), 9–14.
- Kiss, I. & Horváth, J. (2015). Sheltered beaches: A tourism collocational approach to CLIL vocabulary teaching. In: M. Lehman, R. Lugossy & J. Horváth (Eds.), *UPRT 2015: Empirical Studies in English Applied Linguistics* (pp. 166–176). Pécs: Lingua Franca Csoport.
- Košuta, N. (2012). O pitanju stjecanja kolokacijske kompetencije u stranome jeziku. *Strani jezici*, 41(3), 263–276.
- Marton, W. (1977). Foreign vocabulary learning as problem No. 1 of language teaching at the advanced level. *Interlanguage Studies Bulletin*, 2(1), Sage Publications Ltd., 33–57.

- Mendoza Rivera, O., Mitkov, R. & Corpas Pastor, G. (2012). A flexible framework for collocation retrieval and translation from parallel and comparable corpora. In: R. Mitkov, J. Monti, G. Corpas Pastor & V. Seretan (Eds.), *Multword Units in Machine Translation and Translation Technology* (pp. 165–180). Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Moehkardi, R. R. D. (2002). Grammatical and lexical English collocations: some possible problems to Indonesian learners of English. *Humaniora*, 14(1), 53–62.
- Murphy, L. M. (2010). *Lexical Meaning*. Cambridge: Cambridge University Press
- Mustapić, E. & Malenica F. (2013). Collocations as a part of an English LSP Course – The importance of congruence for achieving English language proficiency. *The International Language Conference on the Importance of Learning Professional Foreign Languages for Communication between Cultures 2018*, 207–214.
- Nattinger, J.R. & DeCarrico, J.S. (1992). *Lexical Phrases and Language Teaching*. London: Longman
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242.
- Peeters, B. (2007). Australian perceptions of the weekend. In: P. Scandera (Ed.), *Phraseology and Culture in English* (pp. 79–108). Berlin: Mouton de Gruyter.
- Petrović, B. (2007). Razvijanje kolokacijske kompetencije u hrvatskome kao stranom i drugom jeziku. *Strani jezici* 36(1), 31–38.
- Pierini, P. (2009). Adjectives in tourism English on the web: A corpus-based study. *CÍRCULO de Lingüística Aplicada a la Comunicación (clac)*, 40/2009, 93–116.
- Rata, G. (2012). Adjectival tourisms. In: G. Rata, I. Petroman & C. Petroman (Eds.), *The English of Tourism* (pp. 19–34). Newcastle: Cambridge Scholar Publishing.
- Stojić, A. & Murica, S. (2010): Kolokacije – teorijska razmatranja i primjena u praksi na primjerima iz hrvatskoga i njemačkoga jezika, *Fluminensia*, 22(2), 111–126.
- Van Der Meer, G. (1998). Collocations as one particular type of word combinations. Their definition and character. *Euralex '98 Proceedings*, 313–322.
- Vučković-Vojnović, D. & Nićin, M. (2012). English as a global language in the tourism industry: A case study. In: G. Rata, I. Petroman & C. Petroman (Eds.), *The English of Tourism* (pp. 3–18). Newcastle: Cambridge Scholar Publishing.