

CROATIA AS A VIRTUAL TOURIST DESTINATION: A LINGUISTIC AND SENTIMENT ANALYSIS¹

Miriam Bait
Suzana Jurin
Brigita Bosnar Valković

Abstract

Purpose – Nowadays the virtual access to remote destinations through the use of different platforms has created a new profile of users: the virtual tourists. The question that arises by the emerging of such a new profile of tourist is whether the current structure of Web sites and the adopted communication strategies are adequate to meet the requirements of this type of users. The purpose of our research was to evaluate the effectiveness of the texts that are used to construct a particular image for Croatia as a tourist destination through the analysis of two websites: <https://croatia.hr/en-GB>, the official Croatian Tourism website, in its English version, and <https://www.europeanbestdestinations.com/destinations/search#q=croatia>, a website maintained by an European organisation aimed to promote culture and tourism in Europe which lists Croatia among the suggested best destinations.

Methodology – In our study, two different approaches were combined in a multi-method analysis of the websites. A qualitative analysis within the framework of Discourse Analysis evaluated the linguistic resources used and communicative strategies adopted to promote Croatia. Text linguistic analysis provided data about the language of the selected websites, mostly on its morphology, syntax and lexis which is of great importance for a virtual tourist to situate a virtual text within a cognitive frame.

A quantitative analysis (using AntConc tool) focused upon the frequency and distribution of words in the texts provided further insight and integrated the information value of the results produced by qualitative research. Finally, a sentiment analysis of the contents of the selected websites, and of the general feelings generated on the web by events, news and advertising, using open source tools was performed.

Findings – Through our research several outputs concerning the qualitative linguistics, quantitative and sentiment analysis of the website texts (evaluative terms, positive, negative or neutral sentiments) were obtained. While in industry sentiment analysis is commonly used, in academia it represents an important and innovative aspect.

Contribution – The preliminary results of our research will open up opportunities for future research and practical implications of the results gained that will be implemented in the creation of tourism websites evoking positive sentiments.

Keywords Croatia as a tourist destination, virtual tourist, qualitative linguistic analysis, quantitative corpus analysis, sentiment analysis

¹ This paper is jointly authored, with each author focusing on a specific area of analysis. Suzana Jurin is responsible for section 1, Miriam Bait for sections 2 and 3 and Brigita Bosnar-Valković for composition and preparation of the paper.

INTRODUCTION

The fact that nowadays more and more people are travelling in a virtual way using advanced technology motivated us to analyze Croatia as a virtual tourist destination. The virtual access to remote destinations through the use of different platforms has created a new profile of users - virtual tourists. Virtual tourists are primarily focused on information and an interactive exchange of virtual information and thus have an impact on trends in virtual presentation of destinations. Besides analyzing Croatia as a virtual tourist destination, we also decided to analyze potential reactions of virtual tourists to the provided pieces of information on Croatia as a virtual destination. Based on the aforementioned, the study will be carried out firstly through linguistic analysis focused on texts available on web-sites presenting Croatia as a virtual destination. In order to obtain more complex insights into the research subject, quantitative analysis focused upon the frequency and distribution of words as well as sentiment analysis supported by innovative artificial intelligence methods will also be performed, in order to evaluate the response of virtual tourists to media stimuli. Special attention will be given to positive reactions of virtual tourists because they can encourage a virtual or potential tourist to become a real tourist which means a consumer of offered tourism services and products. Accordingly, this study does not only refer to the present situation, but for the purpose of achieving better economic results in the future, our aim is to contribute to the creation of new trends in the design of the texts representing Croatia as a recognizable and competitive virtual destination.

1. LINGUISTIC TEXT ANALYSIS

Communication is an input to our senses that occurs non-stop. Every input to our senses is a stimulus for reactions, for behaviour, for the way of doing things, for the way of thinking about happenings and surroundings, thinking about other people, communication partners, and for the way of thinking about our communicational reality. Inputs to our senses are stimuli, available for us to interpret and save them as information where from we can derive further information. Everything that we do, see, read and feel is a subjective creation of the sentiment that is present in the very moment of our stimulus - reception. Therefore, we cannot speak about objective experience in any area of our presence and lives, regardless if we consider our private life and making choices about a destination for a great holiday, or we consider our professional life and decisions we make every day in accordance with the corporate ideology. There is no objective experience, all decisions, sentiments and actions are subjective as our brains create the images that we think we perceive. To that extent, objects and sentiments are our own creations and the experiences of them are of subjective, not objective nature (Bateson 1988, 83).

The aforementioned builds a background of the following text analysis which includes texts that introduce Croatia as a virtual destination. The texts of expressive narration will be analysed as well as their effectiveness in creating a particular image of Croatia as a virtual tourist destination. The analysed texts were taken from two websites: <https://croatia.hr/en-GB>, the official Croatian Tourism website, in its English version, and <https://www.europeanbestdestinations.com/destinations/search?q=croatia>, a website

maintained by an European organisation aimed to promote culture and tourism in Europe which lists Croatia among the suggested best destinations. The analysis will be conducted at the linguistic - stylistic level. For the purposes of this analysis, five texts of expressive narration from the above mentioned corpora were investigated.

1.1. Text in the virtual communication process

Text is a set of linguistic, cognitive and social actions that gives us a model through which we understand the world (Beaugrande 1997:128). Through this model the text provides information about the units, statements, facts, events or actions. The text gives information about the extra-linguistic reality (which can be understood as a person, an institution and a state of being). In virtual communication about tourist destinations, the text is used as the main tool for providing all relevant information about the destination, its function and its impact on the tourism environment. The text is a result of linguistic activities of people involved in a social (virtual) interaction, and this activity is dependent on the cognitive and sentiment evaluation of participants in communication, whereby it actualizes different knowledges of the recipient. This knowledge is manifested as subjective and emotional, and in texts it is manifested in a specific way so that it constitutes the dimensional structure of texts (Heinemann/Viehweger 1991, 172).

Texts are holistically and emotionally connected and conditioned units, which through their characteristics and components refer to evoked sentiments (Heinemann 2000, 48). Because of this, texts presenting virtual tourist destinations are not closed systems or units, but rather open systems that are dependent on the currently default communicative emotional and cognitive objectives. The analysed texts are open systems that serve the purpose of a virtual communication which implements sentiments and information, as well as cognitive objectives for a virtual tourist.

1.2. Meaning and function of texts of expressive narration - sentiment triggers

Texts of an expressive narration convey feelings and attitudes, and above all they have the function of establishing or maintaining some forms of interpersonal relations. With texts of expressive narration, authors "react" to emerged circumstances, give their opinion or express attitudes and emotions about current facts, situations that have recently emerged or were observed by the author, facts or situations that cannot be overlooked or ignored any more. These new situations and facts can stem from the individual or social domain. In one sentence, texts of expressive narration create new sentiments or evoke old ones.

According to the text linguistic theory, texts with expressive narration can serve to stabilize or destabilize sentiments (Rolf 1993, 227-278). Also, the texts of expressive narration can be divided into texts with the basic function of establishing or maintaining contact, texts with the basic function of expressing a positive or negative attitude towards someone or something, and texts with the principal function of venting feelings (Ivanetić 2003, 73-74). The illocutionary intention of texts with expressive narration is to show an attitude towards someone or something, the direction of action is not pointed out because it is assumed in the statement, and the mental attitude of the author is also different (Ivanetić 2003, 47). Texts presenting a virtual tourism destination belong to the texts

with the basic function of building a positive or negative attitude towards someone (Croatian people) or something (a destination) – basically, these texts can be classified as sentiment triggers².

1.3. Results of the qualitative linguistic analysis of texts presenting Croatia as a tourist destination

Texts that present a tourist destination (as statements for future good wishes) are future-oriented texts (Martin-Cleef 1991, 131), as they express wishes for prosperity and happiness in the upcoming period, in the upcoming holiday, short vacation etc. Thereby the author of the text expresses his/her sympathy as well as desire and hope for amusement, happiness, relaxation and health. It is evident that here we are talking about the positive empathy of the sender (author) directed to the recipient (virtual tourist).

Analysed texts show that they are, as means of communication, are a form of routines in the field of tourism promotion, and therefore such texts confirm predictable content and structure, information emptiness, formalization and stereotyping. But on the other hand, due to their unnecessary metaphorisation, usage of colourful images which through exaggerated beauty of nature try to convince the traveller - tourist to visit the commercialized destination, these texts are typical sentiment triggers.

Texts with expressive narration as a means of tourist destination promotion are on the one hand business oriented and on the other they are emotionally oriented and acting as sentiment triggers. Additionally, they are determined by cultural and business practices that are nurtured in the field of tourism promotion. Communication through expressively narrated texts is future-oriented and showing positive promotion of the destination (sender) to the virtual tourist (recipient).

In the analysed texts of expressive narration for the purpose of a tourist destination promotion, a number of similarities and characteristics was identified at the communicative – pragmatic, thematic and linguistic-stylistic level:

1) All five examined texts are written in a short and concise format which is on the macro-level divided into three sections: introduction, central part and conclusion. The original language of all texts is English. The main theme of the first two texts is sailing along the Croatian coast and renting sailing boats for that purpose, the main theme of the third text is camping in Croatia, the fourth theme is diving in the Adriatic and the fifth main topic is bike riding through the towns at the seaside.

The introductory part of all five texts begins with a strong expressive syntagm or sentence, or even a metaphorical expression - a syntagm, which has the task to involve the reader, recipient, the virtual tourist into the emotional content presented by further reading e.g.:... *Sailing is a symbol of ultimate freedom and independence.....* or... *When*

² For the purpose of this study the term *sentiment trigger* has been introduced by Suzana Jurin. In the context of texts of expressive narration sentiment triggers are texts that evoke either positive, negative or neutral sentiments.

you go sailing along the Croatian coast, you will reconnect with breathtakingly beautiful nature....or ...Dive to discover the sunken wonders.... or... Excitement on wheels... etc.

This part seduces the recipient, slides into the emotions and triggers all wishes, dreams and hopes from the childhood and youth, visualizes a very long time ago and transforms recipients into free and wild explorers, adventurers and romanticists. The trigger effect is created by the usage of strong adjective + noun syntagms e.g.: *...sunken wonders...or...breathtakingly beautiful nature...or...romantic beaches... or...refreshing summer... or...the wild and untouched nature...etc.*

At the syntactical level, the introductory part contains mainly short and simple sentences that express hope for happiness and enjoyment in the possible future vacation, or they provide concise facts about the destination, e.g.: *...Enjoy the mild and pleasant climate.... or... You can moor your vessel in 70 marinas....etc.*

2) The central part of the text contains relatively long explicit dependent clauses in order to describe the destination closer, e.g.: *...Sailing in Croatia is also very popular because you can choose from a wide variety of sailboats for rent...or...The thermal wind during the day is replaced by a light night- time breeze...or ...Cave diving is also becoming more popular especially in the karst regions of Dalmatia... etc.* Very often complex imperative forms occur, in order to assert the decision to consume that type of vacation or destination, e.g.: *... Ride through astonishing landscapes in natural parks... or ...Even challenge your own limits and get excited reviving the five stages of the Tour of Croatia...or... Sleep tight under the stars and wake up your wild side...or ...Free your energy... etc.*

In a number of sentence examples, a text rhyme is observable e.g.: *... Take up your bike, and enjoy the hike ... or ... Try to fly, explore the sky... etc.* Humour that tries to break the formality of an informative text about a certain tourist destination or activity, when the content is distributed through formal communication channels, or when the content is targeting the younger population of virtual tourists, is also evident, e.g.: *... be blind and ride up the sky....or... if you think you are the captain of all times, wait that the queen of all winds start- Bora ...etc.*

The language of the central part of the text is almost of poetic stylistic nature. The aim of usage of such lyrical style is seducing and giving the opportunity of individual, deeply personal interpretation of vivid images which are realized through the usage of colourful metaphors reinforcing them with even more colourful adjectives, e.g.: *...true paradise... or...magical place...or...imposing steep cliffs...or...the simultaneous sounds of crickets...etc.*

At the lexical level, "vacation- holiday" phrases and words, expressive verbs and adjectives that are a statement of social, less than personal - emotional routines, are products of a psychological desire, and they are represented in the texts in following forms: *...Explore the deep going on organized tours...or...Cross the gigantic paths...or...Amaze yourself surrounded by green landscapes...or...Just relax during walking...etc.*, without explicitly using the noun *vacation* or *holiday*. Auxiliary verbs *will* and *shall* (*find*), and verbs *can*, *take*, *see*, *experience* and *wish* are frequently used in

all texts, as well as the verbs *fill* and *fulfill* in the context of fulfilling wishes, e.g.: ... *may you be filled with joy in the ...* or ... *may your innermost desires be fulfilled...* etc.

Adjectives which were registered in each out of the analysed texts from both corpora describe recipient's emotional, physical or mental condition (virtual tourist's) satisfaction, e.g.: *happy, blessed, peaceful, successful, promising ...* etc. Words and phrases loaded with positive connotations, that describe benefits or the desire for happiness and positive future-focused experience are confirmed in each central part of the examined texts, for example: ... *You will feel the energy...* or... *you will start to imagine...or... you will reconnect...or...you will fall deeply in love...* etc.

3) The end part or the conclusion of the analysed texts is reserved for a maxim, moral lesson or an emotional message, very often expressed in an imperative form, e.g.: ...*You can do all that while listening to the wind and nature calling you...or...Whatever you do, you will fall deeply in love...or...Sleep tight under the stars and wake up your wild side...or... Live every experience, live every sensation, come and live Croatia!...or... Amusements never ends...* etc., or simply is an expression of concluding facts about the destination, e.g.: ...*Croatia is a true paradise for lovers of active tourism, nature and adrenaline...* or...*Sailing has been a tradition in Croatia for centuries but it will need little time to conquer you...*etc.

2. QUANTITATIVE ANALYSIS

Corpus linguistics contains a strong quantitative element as its constituent part that allows to reveal salient contextual features, make comparisons among texts and find patterns of grammatical use, or frequently recurring phrases in a corpus of texts, showing differences and similarities. Thus, the aim of our quantitative analysis was to identify the frequency and distribution of words and patterns in the texts, in order to provide further insights and integrate quantitative results with the information value provided by the qualitative analysis.

2.1. Methodology

The aim of the performed quantitative analysis was to identify the frequency and distribution of words and patterns in the texts, by using AntConc Corpus Tool, version 3.4.4.0, a free software tool for text annotation and analysis.

Two corpora were prepared by downloading texts from the two selected websites and transforming them into two corpora in .txt format, as a first step. The first corpus (hereinafter Corpus 1) consisting of texts from the official Croatian Tourism website (<https://croatia.hr/en-GB>) is composed of approximately 30,000 words. The texts forming the second corpus (hereinafter Corpus 2) were retrieved from the European Best Destination website, (<https://www.europeanbestdestinations.com/destinations/search?q=croatia>) and total 13,000 words.

The second step was to pre-process the text corpora, which includes eliminating ‘stopwords’, such as *the, a, an, and*, etc., to enable a lexical analysis on content-driven terms. A wordlist from each corpus was then generated, i.e., a list of words in order of frequency of appearance, and an analysis of the concordance lines i.e., the words which are used before and after the search item, was also carried out in order to investigate words in the context in which they occur in the text, looking for patterns of similarity or contrast. In fact, interesting relevant insights into the structure and usage of a language can be obtained by analyzing words in real texts and finding out what patterns of lexis, grammar and meaning surround them.

2.2. Results of the quantitative linguistic analysis

The analysis shows that the texts in both corpora are mainly descriptive and the verbs that appear at the top of the frequency lists in both corpora have a typically static meaning (*is, are, has, can*). Stative verbs can be seen and compared in the table that follows.

Table 1: Frequency of verbs with static meaning in Corpus 1 and 2

CORPUS 1			CORPUS 2		
Rank	Lemma	Hits	Rank	Lemma	Hits
1	<i>is</i>	599	1	<i>is</i>	269
2	<i>are</i>	231	9	<i>are</i>	62
11	<i>has</i>	120	17	<i>can</i>	40
12	<i>can</i>	111	18	<i>has</i>	38
22	<i>have</i>	71	27	<i>have</i>	30

In Corpus 1, the first example of a lexical verb indicating an action (*offer*) is found at rank 37, immediately followed by *enjoy* at rank 39 in Corpus 1. In Corpus 2, the top ranking action verbs are *visit* at rank 21, and *discover* at rank 35.

Table 2: Frequency of verbs indicating an action in Corpus 1 and 2

CORPUS 1			CORPUS 2		
Rank	Lemma	Hits	Rank	Lemma	Hits
37	<i>offer</i>	54	21	<i>visit</i>	35
39	<i>enjoy</i>	53	35	<i>discover</i>	25
41	<i>located</i>	51	36	<i>enjoy</i>	25
84	<i>find</i>	33	66	<i>take</i>	17
120	<i>protected</i>	25	114	<i>stop</i>	13
132	<i>make</i>	23	119	<i>make</i>	12
135	<i>surrounded</i>	23	132	<i>find</i>	11
171	<i>become</i>	19	133	<i>go</i>	11
197	<i>preserved</i>	18	135	<i>known</i>	11
232	<i>visit</i>	16	137	<i>located</i>	11

Table 2 provides a more extensive view on the verbs most used in the corpora.

Even a superficial reading reveals that the verbal choices in Corpus 1 tend to *describe* what the country can offer to tourists with a special focus on the country's environment (as suggested by *protected*, and *preserved*), while in Corpus 2 the tourist is more encouraged to take action (as suggested by the verbs *visit*, *discover*, *enjoy*, etc). The top frequency of the personal pronoun *you* is consistent with the attempt to engage the potential tourist and confirms a high degree of interactivity, as the following examples show:

you definitely need to visit!
...something *you* definitely need to try.
you listen to the wind and nature calling you.

Table 3: **Concordance 1 (you) in Corpus 1**

CAN	WILL
<i>you</i> can paddle down the river in rafts	<i>you</i> will encounter a large number of facilities
<i>you</i> can be a true adventurer	<i>you</i> will get to see Croatia as you have never seen it before!
<i>you</i> can go whenever you want	<i>you</i> will wipe away all your fatigue
wherever <i>you</i> go, you can see the sea	<i>you</i> will find the perfect way to enjoy sailing
<i>you</i> can take full pleasure in its charm	<i>you</i> will feel close to a fascinating wilderness
<i>you</i> can enjoy the delights of the Kvarner region	<i>you</i> will find a huge variety of ideal trails
<i>you</i> can experience true freedom	Whatever you do, you will fall deeply in love

As we can see from the instances of *you* in Concordance 1, the target consumer "you" is made the explicit actor: the auxiliary *can* exemplifies that the subject is always the tourist, and *can* transmits the idea of possibility, while the occurrences with *will* convey the idea of possibility and certainty.

Table 4: **Concordance 2 (you) in Corpus 2**

Subject position	Object position
You will fall in love with this country	Croatia gives you an adventure of a lifetime
You will dive into the pure waters of Croatia	Istria is one of the places which provides you with constant winds
You will discover a charming little lake	This destination gives you a quiet and peaceful environment
You can find shelter in a safe harbour	Croatia will seduce you with its exceptional nature
<i>you</i> can simply travel on a tourist boat	Croatia gives you Crystal clear water, sandy and pebbly beaches
	Istria ... offers you a paragliding experience

In Corpus 2, if compared to Corpus 1, a very limited use of *can* and *will* was registered. The above table shows that just few examples of *you* with reference to the potential tourist performing an active role, either with auxiliary verbs or lexical verbs, were registered.

On the other hand, the country, Croatia, or a region or an area, are mostly represented as performing an active role, or better, are personified, and *you* features an object position, that is to say it is ‘acted upon’.

Furthermore, a comparison between the two wordlists reveals that Corpus 1 contains a higher number of words related to the country’s natural beauty, as for example *sea, beach, island, park* a fact which is not surprising for a country boasting “Nature at its finest”, and promoting “The route of the 1,000 islands” on its website.

Interestingly enough, Corpus 2 lays more emphasis on Croatia’s monumental attractions as museums, cities, towns, together with names of cities such as Zagreb, Zadar, Krk, Opatija. Therefore, marked prominence is assigned to the artistic heritage, thus providing a different image of Croatia that might not correspond to the tourists’ expectations (and stereotypes).

Given the fact that the main purpose of promotional texts is to be descriptive, but more importantly, to be persuasive, evaluative lexis is present throughout both corpora, as table 5 shows.

Table 5: Evaluative lexis in Corpus 1 and Corpus 2

CORPUS 1			CORPUS 2		
Rank	Lemma	Hits	Rank	Lemma	Hits
24	<i>beautiful</i>	67	7	<i>beautiful</i>	63
44	<i>famous</i>	48	25	<i>best</i>	31
70	<i>best</i>	36	40	<i>unique</i>	23
94	<i>ideal</i>	30	47	<i>perfect</i>	21
97	<i>interesting</i>	29	48	<i>ideal</i>	20

However, what the results also appear to show is that in Corpus 1, relatively less evaluation appears to occur in favour of adjective connoting the country, such as *natural, Adriatic, Croatian, national, Mediterranean*, thus confirming a stronger focus on the location’s physical and cultural characteristics and, implicitly, on the authenticity (Dann 1996, 98) of the place.

An analysis of the concordance lines generated for *beautiful*, the most recurring adjective in both corpora, revealed interesting usage of this pre-modifier.

Table 6: **Concordance lines for adjective *beautiful* in Corpus 1 and Corpus 2**

Corpus 1	Corpus 2
the most <i>beautiful</i> locations in the world	<i>Beautiful</i> towns
surrounded by <i>beautiful</i> nature	admire the <i>beautiful</i> view
numerous <i>beautiful</i> rivers	the most <i>beautiful</i> experience in the Adriatic Sea
with its <i>beautiful</i> waterfalls	the most <i>beautiful</i> places in Croatia
the most <i>beautiful</i> littoral rivers	the most <i>beautiful</i> cities of Croatia
the <i>beautiful</i> clean sea	<i>beautiful</i> romantic sunsets
the most <i>beautiful</i> beaches in the world	It is particularly <i>beautiful</i> in spring and summer
the largest and most <i>beautiful</i> bay	the most <i>beautiful</i> Christmas market in Europe
the most <i>beautiful</i> and the tidiest beach on the Adriatic	<i>beautiful</i> permanent exhibitions
breathhtakingly <i>beautiful</i> nature	<i>beautiful</i> street lined with restaurants and cafes
the most <i>beautiful</i> natural sights of Europe	<i>beautiful</i> atrium of the Convent of St. Claire
the most <i>beautiful</i> orchid in Europe	a <i>beautiful</i> and charming little town
its centuries-old highland forests, <i>beautiful</i> meadows, pastures, clear rivers and lakes	beautiful cities on the Croatian coast
the most <i>beautiful</i> European garden	a <i>beautiful</i> and well equipped hotel

Beautiful is the preferred adjective in Corpus 1 whenever natural elements are being described, while in Corpus 2 the same adjective appears to be more culturally connoted as it does refer to places or general terms such as *view* or *sunsets*.

3. SENTIMENT ANALYSIS

To make our research more comprehensive we performed a sentiment analysis (Pang and Lee 2008, 36; Godbole, Srinivasaiah and Skiena 2007, 25) of the contents of the selected websites, in order to obtain information about the general mood of the text, i.e., to assess if it is positive, neutral or negative. In fact, the aim of sentiment analysis, also called opinion mining, is the analysis of people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. Obviously, opinions are central to almost all human activities because they are key influencers of our behaviors.

In the past, when an organization or a business needed to retrieve public or consumer opinions about their products and services, they conducted surveys, opinion polls, and focused groups in order to gather public opinions. Nowadays, there is an abundance of such information publicly available on the Web, but the average human reader will have difficulty extracting and summarizing the opinions in them. Thus, automated sentiment analysis systems are needed, and that is where tools derived from Artificial Intelligence (A.I.) comes in (Nasukawa and Yi 2003, 179; Lohr 2012, 62). Thanks to cloud-based tools and procedures based on Natural Language Processing A.I. algorithms, it is possible to extract sentiment from content and from the web just as any individual would do, but much faster and allowing to obtain more consistent results.

3.1. Methodology

To perform our analysis we used two open source tools (www.danielsoper.com/sentimentanalysis/; <http://werfamous.com/sentimentanalyzer/>), aimed at detecting feelings emerging from the words used in the texts. We investigated the texts at two levels: document level and sentence level.

Document-level sentiment classification assumes that each document expresses opinions on a single entity (in this case, the promotion of Croatia).

Analysis of the sentence level determines whether each sentence expresses a positive, negative, or neutral opinion. Neutral usually means no opinion. This level of analysis is closely related to subjectivity classification (Wiebe, Bruce and O'Hara 1999, 148), which distinguishes sentences (called objective sentences) that express factual information from sentences (called subjective sentences) that express subjective views and opinions.

3.2. Results of the sentiment analysis

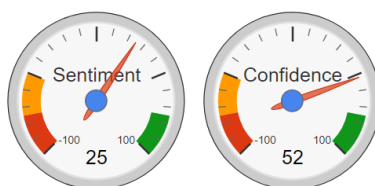
First of all, we started our investigation by using www.danielsoper.com on the content of the official Croatian tourism website. The system computes a sentiment score which reflects the overall sentiment, tone, or emotional feeling of the input text. Sentiment scores range from -100 to +100, where -100 indicates a very negative or serious tone and +100 indicates a very positive or enthusiastic tone.

The general evaluation stated that the texts have an overall very positive/enthusiastic sentiment or tone.

However, an in-depth analysis performed with www.werfamous.com sentiment analyzer revealed that sentences have an average score of 25% on the positive side. This means that the text elicits positive sentiments, although much effort is needed to make the message fully positive (green part of the dashboard, Figure 1).

Figure 1: Sentiment analysis of corpus 1 performed with www.werfamous.com

Your score is : 25% , with a confidence of 52%



As a result, one could claim that the message is not engaging and persuasive, but only mostly informative. To make a message more effective in eliciting positive, enthusiastic sentiments, writers should choose more positive words. The confidence of the analysis is about 52% which is a very good score, considering the informative nature of the text. The following Figure number 2 shows that at the sentence level the overall sentiment is low positive or neutral. This is an index of lack of enthusiastic terms in the text and confirms that the text is mostly informative and does not engage the readers (an example is given in Figure 2 below). The closer the score is to 1, the more the sentiment is positive, the closer is to 0, the more is negative.

Figure 2: Sample of sentiment analysis at the sentence level in Corpus 1

Sentiment by sentences

	Sentence	Sentiment
1	Especially prominent in the development of the MICE industry are the islands of Brač and Hvar and the coastal towns of Makarska, Brela among others.	1
2	It is ideal for incentive and team-building programs!	1
3	The best five hiking spots in Croatia are Učka Mountain, Risnjak National Park, Northern Velebit, Plaklenica and Biokovo Mountain.	1
4	The best place to get lost, the best place to find yourself!	1
5	Baranja is especially interesting.	1
6	Vela Beach features excellent facilities and there are plenty of events there.	1
7	Actually 97.96% of 884 sampled were qualified excellent for bathing.	1
8	Its allure is perhaps best illustrated by the fact that in the 16th century it was also called Devil's Peak.	1
9	The symbols of the island are Pag sheep's cheese, known as one of the best in the world, lamb meat, lace and salt.	1
10	However, access from the sea is ideal for reaching Stiniva beach.	0,9
11	The lunar landscapes of the Kornati islands are ideal destinations for "Robinson Crusoe" tourism.	0,9
12	It is ideal for families with children.	0,9

Another useful result is provided by the list of the significant words classified in order of elicited sentiment. The tools detected 300 significant words: for each, sentiment was obtained (the closer it is to 1, the more positive it is, and vice versa) and the confidence of the analysis as well. Where the confidence might appear ambiguous, in actual fact, it means that the word itself does not provide information about the positive or not positive connotation of the term. Figure 3 shows some sample words from the top, the middle and the bottom of the list.

Figure 3: List of significant sample words in Corpus 1

Word	Occurrence	Sentiment	Confidence
1 most		154	0,5
2 many		82	0,5
3 natural		74	0,1
4 beautiful		67	0,85
5 enjoy		53	0,4
6 famous		48	0,5
7 more		47	0,5
8 other		44	-0,12
9 popular		37	0,6
... omissis			
179 unusual		2	0,2
180 friendly		2	0,38
181 lame		2	-0,22
182 inspirational		2	0,5
183 acquainted		2	0,5
184 adventurous		2	0,5
185 raw		2	-0,23
186 captivating		2	0,5
187 excited		2	0,38
188 better		2	0,5
189 less		2	-0,17
190 limited		2	-0,07
191 dried		2	-0,2
192 military		2	-0,1
193 imaginative		2	0,6
194 artistic		2	0,33
195 lucky		2	0,33
196 recognizable		2	0,25
197 frequent		2	0,1
198 fast		2	0,2
... omissis			
289 rose		1	0,6
290 grey		1	-0,05
291 happiness		1	0,7
292 delicious		1	1
293 inspring		1	0,5
294 dark		1	-0,15
295 colossal		1	0,3
296 golden		1	0,3
297 pure		1	0,21
298 beloved		1	0,7
299 lovely		1	0,5
300 wise		1	0,7

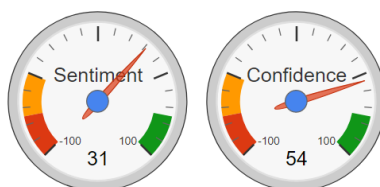
3.3. Results of the sentiment analysis in corpus 2

In Corpus 2 www.danielsoper.com sentiment analyzer detected a sentiment score of -1.3 which proves that the overall sentiment or tone of this text is slightly negative.

The analysis performed with the sentiment analyzer on www.werfamous.com reveals that the sentences have, in general, an average score of 31% on the positive side. This means that the text elicits positive sentiments, although, as in the previous case, the message is mostly informative and text writers should make the message fully positive (green part of the dashboard, figure 4) resorting to a larger number of fully positive words.

Figure 4: Sentiment analysis of Corpus 2 performed with www.werfamous.com

Your score is : 31% , with a confidence of 54%



The confidence of the analysis is about 54% that is again very good, considering the informative nature of the text.

As in the case of Corpus 1, the sentence analysis indicates that, in general, the sentiment is low positive or neutral confirming the informative nature of the texts which is shown in Figure 5.

Figure 5: Sample of sentiment analysis at the sentence level in Corpus 2

Sentence	Sentiment
1 The temperature of the sea doesn't fall below 15 degrees, perfect for this type of activity.	1
2 In summer it is also the opportunity to ride on Zagreb heights for a magnificent view of the city.	1
3 From Opatija in the north to Karlobag in the south there are excellent marinas on the coast and surrounding islands.	1
4 In Rovanjaska you'll see Modrič cave which is almost completely untouched and is an impressive flowstone cave.	1
5 Especially interesting are black-winged stilt and Kentish plover.	1
6 Welcome to Saint-Tropez in Croatia!	1
7 Rovinj is the perfect place to start your nautical trip.	1
8 Welcome!	1
9 A city of history with a delicious climate.	1
10 Croatia is the perfect destination for a boating summer vacation.	1
11 Among the best known localities in the Park are Zavižan, the Velebit Botanical Garden, the Premužić Trail, Štirovača, Alan, Lubenovac.	1
12 There is a breathtaking view on Zadar from the boat.	1
13 Welcome to Saint-Tropez in Croatia!	1
14 A city of history with a delicious climate.	1
15 Welcome to nature!	1

Results emerging from the list of the significant words in order of elicited sentiment are similar to the results obtained from Corpus 1 (cf. Figure 6).

Figure 6: List of significant sample words in Corpus 2

Word	Occurence	Sentiment	Confidence
1 beautiful	63	0,85	1
2 most	51	-0,5	0,5
3 best	31	1	0,3
4 many	29	0,5	0,5
5 enjoy	25	0,4	0,5
6 unique	23	0,38	1
7 perfect	21	1	1
8 ideal	20	0,9	1
9 cultural	17	0,1	0,1
... omissis			
123 real	2	0,2	0,3
124 difficult	2	-0,5	1
125 late	2	-0,3	0,6
126 fresh	2	0,3	0,5
127 complex	2	-0,3	0,4
128 gentle	2	0,2	0,8
129 far	2	0,1	1
130 down	2	-0,16	0,3
131 suitable	2	0,55	0,8
132 serious	2	-0,33	0,7
133 fast	2	0,2	0,6
134 vibrant	2	0,17	0,3
135 trendy	2	0,6	0,9
136 incomparable	2	0,4	0,6
137 calm	2	0,3	0,8
138 unusual	2	0,2	1
... omissis			
194 delicate	1	-0,3	0,9
195 shallow	1	-0,33	0,5
196 captivating	1	0,5	1
197 pink	1	-0,1	0,3
198 unknown	1	-0,1	0,6
199 sweet	1	0,35	0,7
200 lively	1	0,67	0,9
201 busy	1	0,1	0,3
202 whole	1	0,2	0,4
203 cool	1	0,35	0,7
204 powerful	1	0,3	1
205 fitting	1	0,5	0,5
206 early	1	0,1	0,3
207 common	1	-0,3	0,5

CONCLUSIONS AND FURTHER RESEARCH

The evaluation of the effectiveness of the texts that are used to construct a particular image for Croatia as a tourist destination through the analysis of two websites: <https://croatia.hr/en-GB>, the official Croatian Tourism website, in its English version, and <https://www.europeanbestdestinations.com/destinations/search?q=croatia>, showed within the linguistic text analysis performed on five texts of expressive narration that they are dynamic units created to satisfy vivid communicational necessities in the society, and that they change in accordance with different circumstances and needs of the contemporary changeable and fluctuating lifestyle.

The analysed texts present Croatia as a virtual tourist destination with their basic function to show a positive attitude towards a tourist destination through evoking positive sentiments in order to assert to the tourist the decision about purchasing that destination. Due to this fact all analysed texts can be classified as sentiment triggers, and at the text linguistic level they show a certain almost standardized and schematized pattern. The analysis showed that they, as a means of communication, are a form of routines in the field of tourism promotion, and therefore such texts confirm predictable content and structure, information emptiness, formalization and stereotyping. But on the other hand, because of an excessive metaphorisation of colourful images they try to convince the traveller, i.e. tourist to visit the commercialized destination, so that the analysed texts proved to be characteristic sentiment triggers.

The quantitative analysis showed that the texts in both corpora are mainly descriptive and that the verbs from both corpora that appear at the top of the frequency lists have a typically static meaning (*is, are, has, can*). The verbal choices in Corpus 1 tend to *describe* what the country can offer to tourists with a special focus on the country's environment (as suggested by *protected, and preserved*), while in corpus 2 the tourist is more encouraged to take action (as suggested by the verbs *visit, discover, enjoy* etc.). Corpus 1 contains a higher number of words related to the country's natural beauty, as for example *sea, beach, island, park* a fact which is not surprising for a country boasting "Nature at its finest", and promoting "The route of the 1,000 islands" on its website.

Interestingly enough, Corpus 2 laid more emphasis on Croatia's monumental attractions as museums, cities, towns, together with names of cities such as Zagreb, Zadar, Krk, Opatija. Therefore, marked prominence is assigned to the artistic heritage, thus providing a different image of Croatia that might not correspond to the tourists' expectations (and stereotypes).

An analysis of the concordance lines generated for *beautiful*, the most recurring adjective in both corpora, revealed interesting usage of this pre-modifier. *Beautiful* is the preferred adjective in Corpus 1 whenever natural elements are being described, while in Corpus 2 the same adjective appears to be more culturally connoted as it does refer to places or general terms such as *view* or *sunsets*.

The general evaluation based on sentiment analysis of Corpus 1 stated that the texts have an overall very positive/enthusiastic sentiment or tone. However, an in-depth analysis performed with www.werfamous.com sentiment analyzer reveals that sentences have an average score of 25% on the positive side. This means that the text elicits positive sentiments, although much effort is needed to make the message fully positive. To make a message more effective in eliciting positive, enthusiastic sentiments, writers should choose more positive words. In Corpus 2 www.danielsoper.com sentiment analyzer detected a sentiment score of -1.3 which proves that the overall sentiment or tone of this text is slightly negative. The analysis performed with the sentiment analyzer on www.werfamous.com revealed that the sentences have, in general, an average score of 31% on the positive side. This means that the text elicits positive sentiments, although the message is mostly informative and text writers should make the message completely positive resorting to a larger number of completely positive words.

The texts promoting Croatia as a tourist destination are usually accompanied and reinforced by illustrations which are very carefully chosen. Illustrations clearly offer passengers expected, if not stereotyped visual information whose main function is indirectly to encourage potential tourists to make their choice and visit Croatia. Our research might be further completed including visual analysis (Kress and van Leeuwen 2006, 96) of the websites, in order to interpret representational, interactive and compositional meanings of visual images. Overall, the obtained results based on the text linguistic, quantitative and sentiment analysis present a valuable contribution and basis for further research aimed at creating texts presenting a virtual destination that generate positive/enthusiastic sentiment which is an important prerequisite in making a final decision on visiting a certain destination.

REFERENCES

- Beaugrande, R. de (1997), *New Foundations for a Science of Text and Discourse: Cognition, Communication and the Freedom of Access to Knowledge and Society*, Norwood, NJ: Ablex.
- Brinker, K. (2007), "Textfunktionale Analyse", Walter de Gruyter, *Text und Gesprächslinguistik*, Berlin, New York, 175-186.
- Dann, G. (1996), "Tourists' Images of a Destination: An Alternative Analysis", *Journal of Travel and Tourism Marketing*, 5 (1-2), 41-55.
- Godbole, N., Srinivasiah, M., & Skiena, S. (2007), *Large-Scale Sentiment Analysis for News and Blogs*, ICWSM, 7, 21.
- Heinemann, W. (2000), "Textsorte-Textmuster-Texttyp", Brinker, K. et.al. (Hg.), *Text und Gesprächslinguistik*, 1. Halbband, Walter de Gruyter, Berlin, 507-523.
- Ivanetić, N. (2003), *Uporabni tekstovi*, Zavod za lingvistiku FF, Zagreb.
- Kress, G. and van Leeuwen, T. (2006, 1996), *Reading Images: The Grammar of Visual Design*, London and New York, Routledge.
- Lohr, S. (2012), "The age of big data", *New York Times*, 11.
- Marten-Cleef, S. (1991), *Gefühle ausdrücken. Die expressiven Sprechakte*, Kümmerle, Göppingen.
- Rolf, E. (1993), *Die Funktionen der Gebrauchstextsorten*, De Gruyter, Berlin.
- Nasukawa, T., & Yi, J. (2003), "Sentiment analysis: Capturing favorability using natural language processing", *Proceedings of the 2nd international conference on Knowledge capture* (pp. 70-77). ACM.
- Pang, B., & Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and trends in information retrieval*, 2(1-2), 1-135.
- Wiebe, J., Bruce, R., and O'Hara, T. (1999), "Development and use of a gold standard data set for subjectivity classifications", *In Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*, 246-253.

Miriam Bait, Assistant Professor
Università degli Studi di Milano
Facoltà di Scienze Umane
Dipartimento di Filosofia
Via Festa del Perdono 7, Milano 20122, Italy
Tel: +39 02 50312742
E-mail: miriam.bait@unimi.it

Suzana Jurin, PhD, Assistant Professor
University of Rijeka
Faculty of Humanities and Social Sciences
Department of German language and literature
Sveučilišna avenija 4, 51000 Rijeka, Croatia
Tel. + 385 51 265 563
E-mail: sjurin@ffri.hr

Brigita Bosnar Valković, PhD, Assistant Professor
University of Rijeka
Faculty of Tourism and Hospitality Management
Department of Foreign Languages in Business
Primorska 42, 51410 Opatija, Croatia
Tel. + 385 51 294 687
E-mail: bbosval@fthm.hr