



EPiC Series in Language and Linguistics

Volume 2, 2017, Pages 247–255

Professional and Academic Discourse:  
an Interdisciplinary PerspectiveEPiC  
Language  
and Linguistics

# Balancing Tourism Promotion and Professional Discourse: A Corpus-based Analysis of Digital Travel Guidebooks Promoting Venice in English

Daniela Cesiri

“Ca’ Foscari” University of Venice, Venice, Italy  
[daniela.cesiri@unive.it](mailto:daniela.cesiri@unive.it)

## Abstract

The city of Venice is currently promoted through a wide range of channels, either traditional or innovative, using the Internet as their main ‘market place’. Besides Italian, the promotional message is conveyed in English as the main foreign language of communication, addressing not only perspective tourists from English-speaking countries but also tourists of other mother tongues who use English to communicate outside their country.

Considering the multifaceted nature of the language in the field of tourism that balances domain-specific and general terms, the present study investigates the language that is used to ‘sell’ the city of Venice, a complex tourist destination, to international perspective visitors.

By means of corpus linguistics methods of analysis, the study investigates the lexicogrammatical features present in a corpus of digital travel guidebooks, written in English, that are accessible on the Internet. The corpus, purposely collected and annotated, is examined in order to understand how the most peculiar aspects of Venice and its local culture are described to visitors, taking into particular consideration the strategies used by the authors to balance technical terms typical of the language of tourism with promotional terms that might contain specific connotations to attract the readers’ attention.

## 1 Introduction

According to the annual survey conducted by the City of Venice Tourist Board, in 2014, nearly 10 million tourists visited Venice (Miraglia, 2015, p. 12); of these, more than 1.5 million are Italians while nearly 8.5 million are from abroad. The ‘Top 12’ nationalities that visited the city in greater numbers come from European and non-European countries, namely they are – in ranking order – the USA, France, the UK, Germany, China, Japan, Australia, Spain, Brazil, South Korea, Canada and Russia.

These data illustrate the situation that Venetian tourism operators have to face, especially in terms of the language(s) to be used to promote the city, the region but also individual destinations or events. It is clear, then, that the role that English as a *lingua franca* plays in such as composite scenario is fundamental to communicate successfully the promotional message, addressing not only perspective tourists from English-speaking countries as also tourists of other mother tongues who use English to communicate outside their country.

In addition, we have to consider that Venice is also currently promoted through a wide range of channels, either traditional or innovative, using the Internet as their main ‘market place’. Considering the multifaceted nature of the language in the field of tourism (Dann, 1996) that balances domain-specific and general terms, the present study investigates the language that is used to ‘sell’ the city of Venice, a complex tourist destination, to international perspective visitors.

The study will examine digital travel guidebooks (henceforth referred to as DTGs), written in English, which are freely accessible on the Internet. By means of corpus linguistics methods of analysis, a quantitative and qualitative investigation of the lexico-grammatical features used to describe the most peculiar aspects of Venice and its local culture, taking into particular consideration the strategies used by the authors to balance technical terms typical of the language of tourism with generally promotional terms that contain specific connotations to attract the readers’ attention.

The texts will be grouped into a corpus that will be, first, tagged via the *CLAWS Part-of-Speech Tagger* (Kang & Yu, 2011) and, then, searched by means of *Wordsmith Tools 6.0* (Scott, 2012). This procedure will make possible the detection of lexical words, adjectives and verbs that could reveal whether the authors of the travel guides manage successfully to convey their promotional message, at the same time providing effective practical information for the tourists to fully appreciate their visit to Venice.

## 2 DTGs in Today’s Tourism Promotion

According to (Maci, 2013, p. 225) “guidebooks became available as early as the fourth century B.C., covering a vast number of destinations [...]. However, it was under the Roman Empire that international travel became important”. The century-old tradition of using guidebooks during a tourist travel experience is, thus, as old as the tourism phenomenon itself that is commonly believed to have started “as far back as the Babylonian and Egyptian Empires” (Maci, 2013, p. 225). Each one of us, when travelling for leisure or business, has most likely used an authoritative and reliable support that could illustrate the itinerary to take, the things to visit and see, the practical information to reach and move around the destination as well as where to sleep and eat, in other words, to enjoy the experience as much as possible.

New technologies have facilitated tourists access to this kind of information but, at the same time, the role of the traditional, printed travel guidebooks has somewhat been challenged by the new digital devices, so the former were ‘forced’ to reinvent themselves to meet technological advancements. Websites are now available, performing the same function as guidebooks with the advantage of a constant, often real-time, update of the information provided. Applications for tablets and smartphones are now available to consult not only traditional data about the destination but also to search for other tourists’ opinions and reviews of the place to go, also in this case with information constantly updated. Furthermore, websites offer the possibility to download for free, in PDF format, digital travel guidebooks that tourists can consult at the destination even if they do not have access to a Wi-Fi or a roamed connection.

The possibilities offered to tourism operators by these texts are manifold: they can be accessed by the tourist before the travel experience actually takes place, they can help perspective visitors decide which operator can offer the best service, they can also literally follow the tourist during the stay, giving

the impression of being personally taken care of by the author of the files, also because it is possible to access online information to the main website, whose hyperlink is provided regularly in the pages of the DTG.

The potential of, and the role offered by, DTGs make them an interesting genre to be analysed by a linguist in terms of the promotional and technical language used: is it different from the traditional guidebooks, or is it similar so that just the support (paper vs digital format) actually changes? The present study seeks to answer to these questions but also to actually look at the strategies that the authors of the DTGs use to convey their promotional message in introducing visitors to the city of Venice and to guide them during their stay.

### 3 Corpus and Methodology

The main research question that leads the study was developed using corpus linguistics methods to investigate the language of tourism promotion in DTGs describing Venice in English. To this purpose three main phases were followed:

1. **Corpus Compilation:** nine DTGs written in English were collected to form a corpus called VTG, which stands for Venice Travel Guidebooks;
2. **Corpus Annotation:** the VTG corpus was annotated using (Kang & Yu, 2011), the C7 Tagset was used because it contains 166 word tags, including punctuation tags, that could allow an annotation as thorough as possible of the semantic categories present in the VTG corpus;
3. **Corpus Search:** the annotated corpus was searched by means of *Wordsmith Tools Version 6* (Scott, 2012) in order to look for recurring lexico-grammatical features.

After completion of Phase Three, qualitative and quantitative analyses of the outcomes of corpus search were conducted. In particular, lexical nouns, verbs, adjectives, adverbs and pronouns were isolated from the corpus, but the main attention focused on nouns, verbs and adjectives that could help understand the typical expressions used by the authors to describe and promote Venice. In addition, it was seen whether the authors of the DTGs manage successfully to convey their promotional message and, at the same time, whether they were able to provide effective practical information for the tourists to fully appreciate their visit to the city.

The nine DTGs composing the VTG corpus are the following ones:

1. *Venice Guide* (by TomsPortGuides.com),
2. *The Rough Guide to Venice and The Veneto*,
3. *The Venice City Guide* (by Travelplan.it),
4. *Venice Guide* (by Venezia.net),
5. *Venice City Guide* (by Tripomatic.com),
6. *Venice Smart Tourism Guide* [by the Italian Association of people with Down Syndrome, Equal Partners Foundation (Malta), Společnost rodičů a přátel dětí s Downovým syndromem, Os (Czech Republic) and Down Alapítvány (Hungary)],
7. *City Guide* (by WeeaGoo),
8. *Venice* (by Arrival Guides),
9. *Introducing Venice* (by Alison Bing for Lonely Planet).

Some of these DTGs, such as number 3 and 4, are written by Italian authors and have a corresponding text in Italian downloadable from a specific section of the websites. However, after comparison of the English and the Italian versions, it could be assumed that the English version is not

a translation of the Italian one because they contain different information, tailored to meet the demand of a foreign VS an Italian tourist.

In addition, it must also be specified that the DTGs are contained in websites, and reference is made to the possibility of visiting the hyperlink leading to the website itself or to download the corresponding app, but the DTGs are not to be considered a summary of either since an analysis of the three ascertained that they contain information that integrate one another but they illustrate substantially different contents. For instance, while the websites contain commercial pages with links to hotels, companies for local transportation, airports etcetera, the apps focused more on the users' reviews of the destination, while the DTG provided practical information in specific sections and the rest of the text is used to describe itineraries, provide historical and artistic information on individual sites as well as on the city in general.

## 4 Quantitative Analysis

The quantitative analysis used two reference corpora, namely FLOB and BNC. The data for the FLOB corpus were taken from (Kang & Yu, 2011) who conducted a stylistic research on a Tourism English Corpus (TEC) that they collected choosing texts "from official tourism websites in Britain and the U.S." (Kang & Yu, 2011, p. 131). The FLOB (Freiburg-LOB Corpus of British English) is a corpus of general English containing samples from 1991 through 1996, while the BNC (British National Corpus) is a much wider corpus of general English, comprising samples of written and spoken language, that was compiled between the 1990s and 2007. In addition, data from the TEC corpus, as presented by (Kang & Yu, 2011), were also compared to data from the VTG corpus in order to contrast the language used in the latter with both general English and a specific corpus from the same domain.

Table 1 illustrates the comparison of the VTG corpus with the three reference corpora:

	VTG (2015-2016)	FLOB (1991-1996)	BNC 1990s-2007	TEC (2011)
Tokens	299,736	1,237,424	97,860,872	107,981
Types	20,884	45,089	512,588	10,789
Type/Token Ratio	6.97	3.64	0.52	9.99
STTR	44.68	45.52	42.66	44.96
Avg. word length (letters)/sentences	4.64/12,531	4.35/52,674	4.68/4,754,513	4.77/5,166
Average sentence length	23.92	23.49	20.59	20.90

**Table 1:** Quantitative data from VTG and reference corpora

Considering the ratios for the two reference corpora of general English, the STTR for the VTG corpus indicates that it is lexically dense, rich in word use. Mean word length values are similar, and the number of sentences is proportionate to the size of each one of the corpora. The average sentence length is also similar in all the four corpora but, in VTG, sentences are longer than in the FLOB corpus, indicating that sentences in VTG are complex: a feature typical of "a more formal style" (de Haan & van Esch, 2007, p. 198).

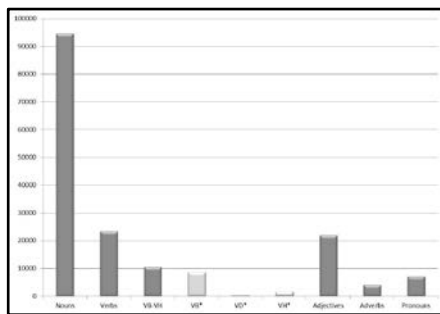
Compared to TEC, it is clear that the two domain-specific corpora have similar figures, demonstrating that the VTG corpus shows stylistic features that are typical of 'Tourism English'; thus, focusing on the values regarding the VTG corpus, we could describe it as lexically dense, with textual complexity typical of the formal register, while the average word length indicates that the words used are content words. Quoting (Diemer & Frobenius, 2013, p. 57), we can say that "usually the combination

of high TTR, high word and high sentence length indicates specialised or possibly even restricted content and elaborate style”. This is also consistent to what (Kang & Yu, 2011, p. 133) affirm about their corpus of English tourism texts, namely that the use of longer words in Tourism English “is to introduce large amounts of practical information such as location, history, culture, etc.”

#### 4.1 Content Words in the VTG Corpus

Graph 1 shows the distribution of content words present in the VTG corpus. The content words searched were nouns, adjectives, adverbs, pronouns and verbs. The latter are indicated in the items ‘Verbs’ as one group including lexical verbs, modals and auxiliaries, while the columns in a different colour indicate the auxiliaries ‘to be’ and ‘to have’ grouped together in the ‘VB-VH’ column in Graph 1 and, then, differentiated in the columns ‘VB’ and ‘VH’. The column labelled with ‘VD’ illustrates occurrences of ‘to do’.

Data about the occurrence of content words in the VTG corpus are in line with (Kang & Yu, 2011) analysis of their domain-specific corpus, so the same assumptions can be applied to the present case, i.e., nouns are more frequent than verbs. This can be interpreted as evidence of “a well-built, concise, clear but not complicated structure” (Kang & Yu, 2011, p. 133), aiming at attracting the readers’ attention. This is also confirmed by the higher percentage of adjectives over lexical verbs because they serve to describe the destination, but also individual parts or single aspects of it, and make it more attractive to potential visitors.



**Graph 1:** Occurrence of content words in the VTG corpus

### 5 Qualitative Analysis: Keywords

A keyword analysis was also conducted taking the BNC as reference corpus because more similar to VTG than FLOB in terms of average word length\*. The keyword analysis of the most frequent and most relevant words occurring in the VTG corpus produced 2,594 entries, from which the first twenty nouns, verbs and adjectives were selected (Table 2).

Categories	Keywords
First 20 nouns	Venice, century, city, church, <i>San Marco</i> , <i>campo</i> , <i>palazzo</i> , <i>piazza</i> , canal, <i>calle</i> , building, time, <i>via</i> , bridge, side, <i>doge</i> , water, <i>Rialto</i> , rooms, place was, were, has, have, had, been, see, built, sat,

\* TEC is not available to the public so it could not be used as a reference corpus.

First 20 verbs	open, would, could, stop, walk, painted, cross, should, takes, need, want
First 20 adjectives	Venetian, most, well, small, best, gothic, major, free, central, famous, modern, expensive, Roman, beautiful, original, southern, moderate, eastern, ancient, popular, huge

**Table 2:** The first 20 keywords in the VTG corpus divided *per* category

As we can see in Table 2, the most frequent nouns refer to topographical elements of the city of Venice and to the city itself. The prevailing language used is English but terms in Italian are used as well. They refer to very specific elements and this might indicate that, even though an English equivalent term exists, the authors preferred to report them in the local language to fill the cultural gap between the city and the foreign visitor, who will then be aware of the terms s/he will find upon arrival at the destination.

This is typical of the language of tourism promotion, which uses a verbal technique called ‘linguaging’, defined as “the impressive use of foreign words, but also a manipulation of the vernacular, a special choice of vocabulary, and not just for its own sake” (Dann, 1996, p. 184). According to the sociolinguistic approach to the language of tourism, the use of these foreign words is particularly appropriate when the reader is believed to have little knowledge of the concepts they convey, thus the author of the promotional text induces feelings of inferiority in the reader and transforms him/herself into a trustful authority.

This technique is also called ‘language crossing’; its function is presented as that of “creating a linguascape of the travel destination” (Jaworski, Thurlow, Lawson, & Ylänne-McEwen, 2003, p. 17) that anticipates the meeting with the local culture of the destination. The foreign words used in these cases are usually taken from the field of eno-gastronomy, or refer to very specific natural, architectural or cultural elements of the host community; generally, they also refer to less-known aspects of the destination and its culture.

As for the most frequent verbs, we clearly see that they are functional to the touring experience and refer to actions that the visitors are very likely to perform once arrived at the destination. The most recurrent adjectives, on the other hand, define specific qualities of the destination; the use of the superlative is also in line with the typical features of the language of tourism, characterised by a bombastic style that tends to emphasise at the extreme the qualities of the promoted destination, creating in the tourists specific expectations and stimulating their desire to actually experience them (Dann, 1996).

## 5.1 Venice vs. Venetian: Collocates Analysis

The first noun and the first adjective that were produced with the keyword analysis were ‘Venice’ and ‘Venetian’, respectively. This is not surprising considering that the DTGs deals with the city itself, its inhabitants and the landscape that surrounds it. One of the research questions that led this study was how the authors of the DTGs managed to describe Venice to perspective visitors, also considering the complex, multifaceted nature of the destination. In order to answer to this question a qualitative analysis was conducted to see what are the words that collocate with greater frequency with both ‘Venice’ and ‘Venetian’. Thus, the first 50 collocates occurring in first position on the right and on the left of the two items were selected. Results are shown in Table 3.

First 50 collocates (L1)	Item (nr. of occurrences)	First 50 collocates (R1)
	Venice (2256)	

<p>in, of, to, the, from, for, destination, century, and, central, that, around, into, on, by, <i>Marco</i>, with, between, as, or, visit, <i>Venice</i>, Renaissance, like, which, is, about, over, outside, guide, rolling, when, after, see, <i>Dorsoduro</i>, connects, <i>Schiavoni</i>, than, your, against, pass, <i>Polo</i>, through, at</p>	<p>'s, the, and, is, phone, was, in, to, it, city, from, Italy, a, by, as, has, are, Internet, but, card, you, for, at, had, where, on, there, with, this, itself, were, international, that, became, or, when, which, Venice, its, however, <i>Marco</i>, Carnival, see, address, since, than, more, we, if, can, who, area</p>	
<p>The, of, a, in, contexts, to, and, by, century, traditional, typical, many, with, some, two, great, famous, distinctively, typically, other, all, for, young, on, now, few, first, old, native, under, true, early, popular, any, beautiful, genuine, real, when, various, uniquely, wonderful, specifically, style, proto, city, <i>cicheti</i>, eminent, discover, celebrated</p>	<p>Venetian (824)</p>	<p>Republic, painting, dialect, cuisine, Gothic, and, art, history, architecture, state, lagoon, style, palaces, life, to, merchants, in, for, glass, specialities, food, where, who, housing, aristocracy, artist, empire, government, Renaissance, Byzantine, social, buildings, paintings, standards, folklore, had, houses, building, military, navy, institutions, painters, nobleman, is, palace, high, society, it, have, painter</p>

**Table 3:** Right and left collocates for keywords 'Venice' and 'Venetian'

The items that collocate with Venice do not use any particular evaluative language that could communicate to the tourists particular sensations towards the city. We find function words such as the genitive, articles, pronouns, prepositions, other nouns referring to topographical elements of Venice. The collocates analysis of the item 'Venice', then, confirms the keyword analysis of the nouns in the VTG corpus: they are used to convey practical information to the tourist with the occasional use of terms in the local language.

The items that collocate with 'Venetian', on the contrary, express the complexity of Venice and its local culture and people. The terms that we see refer to the history of the city, to its past of political supremacy but they also draw the tourist's attention on aspects that are labelled as being typical and authentic of the local culture. In this case, the use of local terms is limited to just one item ('*cicheti*', generally described as 'Venetian *tapas*'), a very specific element of the local eno-gastronomic tradition, bringing the tourist's attention on something that is unique to Venice and, thus, unmissable during the visit.

The collocate analysis of the adjective 'Venetian' and the functions performed by the items in the presentation of Venice to foreign visitors are not really surprising: if nominal elements generally indicate the object, not some particular quality, this role is normally played by adjectives. What is really interesting to notice is the predominant use of the equivalent term in English instead of the Venetian or Italian one. The use of the 'linguaging' technique, then, is here quite limited with the result that, linguistically, the distance with the host community is reduced with the use of English as a *lingua franca*; culturally, however, this distance is intensified because the tourist, upon his/her arrival at the destination, will find a different terminology for the same topographical elements, increasing the sense of alienation with respect of the city and its inhabitants.

## 6 Conclusive Remarks

The present study has shown that the VTG corpus showed specific stylistic features that lead to specific interpretations: the VTG corpus is lexically dense which means that its content is specialised; sentence complexity is typical of a formal and elaborate register. This aspect is reflected also in the average word length which indicates that the terminology used is complex, while the consistent use of content words, nouns and adjectives in particular, reflects the complex narratives in tourism discourse as well as the interdisciplinarity of the tourism discourse community, which employs terms from domains such as the arts, architecture, history, geography, to name only a few.

These data show opposite results to those presented in (Maci, 2013, p. 243), in which “traditional guidebooks show a language that is “informal and colloquial”. One possible explanation is that the authors differentiate the role of the DTG present in a website from that of the website itself, thus using a more specialised language in the DTGs and a more colloquial register in the website, while in traditional printed guidebooks the guide is the only text that the tourist has, so it must be instructional but not boring or too formal.

Keyword analysis showed that the most recurring nouns refer to specific ‘landscape’ elements of the city, while adjectives complete the description emphasising intrinsic qualities, dimensions or other similar characteristics of what is being described. Verbs refer to general actions typical of the touring experience and are signals of the authors’ attitude to control the tourists’ actions during their stay in Venice.

The contrastive analysis of two representative words, such as ‘Venice’ and ‘Venetian’, revealed that the noun ‘Venice’ is used to convey practical information, collocating with content words that do not transmit any particular connotation to the noun they accompany. This might also be interpreted as a preference for periphrastic expressions rather than the creation of expectations through just one term. On the other hand, the adjective ‘Venetian’ collocates with words describing aspects and qualities of the city in the present day as well as in its past history, thus bearing a wide range of connotations – all positive – that make the tourist desire to visit the city, presented in all its complexity.

So, the initial research question, whether the authors of the DTGs showed a preference for technical or for promotional terminology, has obtained a quite composite answer: though they use an emphatic, clearly promotional kind of language, the authors keep their style emotionally balanced, thus giving the impression that, even though they are freely available and connected to a more commercial website, the DTGs are all the same reliable and professionally-made supports to the tourist experience.

## References

- Dann, G. (1996). *The Language of Tourism. A Sociolinguistic Perspective*. Wallingford UK: CAB International.
- de Haan, P., & van Esch, K. (2007). Assessing the Development of Foreign Language Writing Skills: Syntactic and Lexical Features. In E. Fitzpatrick, *Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse* (pp. 185–202). Amsterdam/New York: Brill Rodopi.
- Diemer, S., & Frobenius, M. (2013). When Making Pie, All Ingredients Must Be Chilled. Including You. Lexical, Syntactic and interactive Features in Online Discourse – A Synchronic Study of Food Blogs. In C. e. Gerhardt, *Culinary Linguistics. The chef's special* (pp. 53-81). Amsterdam/Philadelphia: Benjamins.
- Jaworski, A., Thurlow, C., Lawson, S., & Ylänne-McEwen, V. (2003). The Uses and Representations of Local Languages in Tourist Destinations: A View from British TV Holiday Programmes. *Language Awareness* 12 (1), 5-29.



Kang, N., & Yu, Q. (2011). Corpus-based Stylistic Analysis of Tourism English. *Journal of Language Teaching and Research* 2 (1), 129-136.

Maci, S. M. (2013). *Tourism Discourse. Professional, Promotional and Digital Voices*. Genova: ECIG.

Miraglia, A. (. (2015). *Annuario del Turismo 2014*. Venezia: Comune di Venezia: Assessorato al Turismo.Scott, M. (2012). *Wordsmith Tools Version 6.0*. Stroud: Lexical Analysis Software.