



This is a contribution from *Cultus: the Intercultural Journal of Mediation and Communication* 2011: 4

© **Iconesoft Edizioni Gruppo Eurosan**

This electronic file may not be altered in any way.

The author(s) of this article is /are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

 *Cultus*

THE JOURNAL OF INTERCULTURAL
MEDIATION AND COMMUNICATION

CORPUS LINGUISTICS AND CULTURE

2011, Volume 4

Iconesoft Edizioni
Terni - Italy

Registrazione al Tribunale di Terni
n. 11 del 24.09.2007

Direttore Responsabile Agostino Quero
Editore Iconesoft Edizioni
Anno 2011
ISSN 2035-3111

© *Iconesoft Edizioni – Gruppo Eurosan Italia srl*
via Garibaldi 89 – 05100 Terni

La realizzazione di questo volume è stata resa possibile grazie al
contributo dei *Monti dei Paschi di Siena* in collaborazione con
l'Università del Salento



CULTUS

the Journal of Intercultural Mediation and Communication

2011, Volume 4

CORPUS LINGUISTICS AND CULTURE

Editors

David Katan

University of Salento and University of Trieste

Elena Manca

University of Salento

Cinzia Spinzi

University of Bologna

ICONESOFT EDIZIONI
TERNI

CULTUS

the Journal of Intercultural Mediation and Communication

Editorial Board

Michael Agar

Ethknoworks LLC and University of Maryland, College Park, USA

Patrick Boylan

SIETAR

Milton Bennet

Intercultural Development Research Institute, Italy

Patrick Boylan

SIETAR-Italy and past Professor at Roma Tre University, Rome

Ida Castiglioni

University of Milan (Bicocca), Intercultural Development Research Institute

Andrew Chesterman

University of Helsinki, Finland

Delia Chiaro

University of Bologna (SSLMIT), Forlì, Italy

Nigel Ewington

WorldWork Ltd, Cambridge, England

Peter Franklin

HTWG Konstanz University of Applied Sciences, dialogin-The Delta Intercultural Academy

Maria Grazia Guido
University of Salento, Italy

Raffaella Merlini
University of Macerata, Italy

Robert O'Dowd
University of León, Spain.

Anthony Pym
Intercultural Studies Group, Universidad Rovira I Virgili, Tarragona, Spain

Helen Spencer-Oatey
University of Warwick, England

Federica Scarpa
SSLMIT University of Trieste, Italy

Helen Spencer-Oatey
Universoty of Warwick, England

Christopher Taylor
University of Trieste, Italy

Kumiko Torikai
Rikkyo Graduate School of Intercultural Communication, Tokyo, Japan

David Trickey
TCO s.r.l., International Diversity Management, Bologna, Italy

Margherita Ulrych
University of Milan, Cattolica del Sacro Cuore, Italy

Table of Contents

Corpus Linguistics and Intercultural communicative approach: a synergy. <i>Introduction by Cinzia Spinzi</i>	9
A conversation between <i>Mike Stubbs and Sebastian Hoffmann</i>	21
Corpus Linguistics: what it is and what it can do <i>Alan Partington</i>	35
What A <i>TripAdvisor</i> Corpus Can Tell Us About Culture <i>Maria Elisa Fina</i>	59
Perceptions of Europe in British and Italian TV news programmes <i>Silvia de Candia and Marco Venuti</i>	81
Intercultural space in translation education: Parallel corpus as a tool for exploration <i>Xiaoping Jiang and Josta van Rij-Heyligers</i>	99
Notes on contributors	120
Guidelines for contributors	123

What A *TripAdvisor* Corpus Can Tell Us About Culture

Maria Elisa Fina

Abstract

The rise of Travel 2.0 has produced significant changes in the way travellers plan their holidays and TripAdvisor has now become the world's largest travel reviews website. These reviews represent the travellers' voice and can be considered a valuable tourist text type since they reflect the travellers' needs and values.

This study aims to demonstrate that these values can be culturally aligned through a corpus-based analysis of the reviews.

The data include two comparable corpora made up of reviews of accommodations in Puglia (Italy) written by English and Italian travellers. The texts contained in both corpora were subject to linguistic analysis by means of TextSTAT-2.

The analysis carried out in this study is based on the methodology developed by Manca (2011), who proposes a combined quantitative and qualitative approach in corpus analysis, and on the methodology proposed by Tognini-Bonelli (2001) for the identification of functional correspondences across languages.

The analysis of reviews shows that there are significant differences in the way English and Italian travellers perceive the holiday experience. These differences have particular linguistic outcomes, which concern mainly the length of reviews, the description of places and relational issues.

The results obtained by linguistic enquiry are in line with Manca's and are framed within Intercultural Studies. More specifically, they are interpreted in terms of Thinking Orientation (Walker et al. 2003, Katan 2004), Uncertainty Avoidance (Hofstede, 2001) and Affective Associations (Francesconi, 2007).

1. Introduction

Corpus Linguistics has been used in a number of studies as a methodology for investigating language in different linguistic areas and has widely proved to be a valid tool for analysing language use and language variation.

An important contribution to Corpus Linguistics was given by Tognini-Bonelli (2001), whose work is particularly significant in the translation domain, as she develops a three-step methodology for the identification of functionally equivalent units of meaning (*ibid.*).

Over the last few years, Corpus Linguistics as a research domain has been extended to Intercultural Studies. The concept underlying this approach is that the study of language patterns through corpus analysis can tell us not only about the way language is used, but also about those aspects of culture which are “neither taught nor learned, but acquired informally” (Katan, 2004: 46) and conveyed by language.

What characterises these studies is a contrastive approach, as they are aimed at studying and comparing the behaviours of different cultures in a specific domain. The theoretical models usually involved in these studies are Hall’s frame of High Context vs. Low Context Cultures (1990) and Hofstede’s Cultural Dimensions (2001).

While in a few studies (Katan 2006, Zappettini 2010) corpus analysis is carried out without using any software, in others (Manca 2009, Spinzi 2010, Wolfe 2008, Cucchi 2010) corpus analysis is software-aided.

Recently, too, there has been a wide interest in cultural differences in the language of tourism. For example, Manca (2009) investigates cultural differences between Britain and Italy in tourism promotion, while Spinzi (2010) applies Manca’s methodology to investigate the discourse of ecotourism in American English, British English and Italian websites. Other intercultural studies in the tourism domain were carried out by D’Egidio (2010), Orlando (*ibid.*) and Katan (*forthcoming*).

The contribution of this study lies in extending previous cross-cultural studies on tourism to travel-related user-generated content delivered by travel communities. It focuses in particular on traveller reviews, whose main aspects are the diversity and variety of travellers’ opinions and values. This study is intended to be a further confirmation of the validity of corpus analysis as a methodology for investigating cultural issues underlying language patterns.

2. Culture: a shifting concept

An important distinction often referred to in intercultural studies is that proposed by Hall (1990) between High Context (HC) communication and Low Context (LC) communication. This distinction is based on whether in communication the information lies more in the text (LC) or in the context (HC), with 'text' being the "transmitted information" and 'context' being "the amount of information the other person can be expected to possess on a given subject" (1983: 61). Hence, Hall distinguishes between HC cultures, which are more context-oriented (e.g. the Mediterranean culture), and LC cultures, which are more text-oriented (e.g. the Anglo-American culture). While HC cultures tend to focus more on feelings and 'being', prefer implicitness and value relationships more than tasks, LC cultures tend to focus more on facts and 'doing', value tasks more than relationships and prefer explicitness (Katan, 2004).

The importance of *contexting* in cross-cultural communication lies in the fact that the features characterising HCCs and LCCs are mirrored in language. Thus, misunderstanding occurs whenever individuals with different backgrounds interact and these features are unrecognized, misinterpreted or simply disregarded. However, Hall's model is meant to represent not a strict HC vs. LC classification but a continuum along which each culture is positioned.

Hofstede (2001) enriched the model by providing five other cultural dimensions, which are: individualism/collectivism, power distance, uncertainty avoidance, masculinity/femininity and long-term/short-term orientation. For each of these dimensions Hofstede attributes a score to over fifty countries.

In this research we will constantly refer to Hall's and Hofstede's models to discuss the findings obtained. Clearly, the risk involved in this type of research is that it can easily lead to oversimplifications of the concept of culture. The models, for example, do not include important factors affecting communication, such as the context of situation in which communication takes place nor the heterogeneity characterising cultural settings. Indeed, Hannerz (1996: 18) rejects the concept of a "cultural mosaic, where each culture would be a territorial entity with clear, sharp, enduring hedges [...]". He argues that cultures are being reshaped by complex and asymmetrical flows of meaningful forms. In this reshaping of cultures an important role is played by globalization,

defined by Hannerz as an increase in “long-distance interconnectedness” (1996: 17), enabled in particular by transport technology and social media. The latter are defined by Hannerz as “machineries of meaning”, (1992: 26) which have an impact “on the distribution of meanings and meaningful forms over people and relationships” (1992: 27) and contribute to “making the boundaries of societies and cultures fuzzy” (1992:30). Thus, culture flows and the outcomes of this process cannot be disregarded. As Hannerz says (1996: 20), “people, meanings and meaningful forms which travel fit badly with what have been conventional units of social and cultural thought”. This means that in intercultural studies it is necessary to take into account that people move, travel, change and that the effects of globalization influence not only people’s cultural identity but also their communicative competence.

As the present research aims to compare the English¹ culture and the Italian culture in relation to preferred communication styles in a specific communicative event (travel reviews), a few remarks are necessary. First of all, it is not possible to conceive an ‘English culture’ or an ‘Italian culture’ as homogeneous entities. Indeed, within a given country features such as North/South differences, different physical environments or ethnic composition imply differences in values, behaviours and orientations within that culture. Hence, we cannot attribute the cultural orientations displayed by our findings to the whole country nor to all subcultures constituting it. We are also aware of the ‘danger’ involved in using Corpus Linguistics as a methodology for cross-cultural studies, i.e. producing cultural generalizations from “a handful of keywords” (Stubbs, 2001:167).

However, software-aided corpus analysis does allow us to spot repeated and recurrent lingua-cultural patterns, and collocates occurring in the same context, from which it is possible to recognize specific “meaningful forms”. Thus, in spite of cultural complexity it is still possible to identify for both the cultures involved patterns of behaviours reflecting underlying values within particular contexts - and here in reference to the tourism domain. Furthermore, according to Hofstede (2001), intra-national differences are normally weaker than cross-cultural differences.² Finally, it should be remembered that, being the largest travel reviews website in the world, *TripAdvisor* represents one of the

¹ Since very few reviews are written by Scottish, Irish or Welsh travellers we will refer to “English corpus”, “English travellers” and “English reviews” rather than “British”.

² See *Cultus* 3 (2010) for an ample discussion on Hofstede.

best examples of globalization through social media. If we accept that when producing reviews travellers leave a part of their identity in their writing then the fuzziness characterising the boundaries of societies and cultures will also be reflected in the travel community reports. Hence, given the scale of the *TripAdvisor* phenomenon, we may expect to be dealing with an amount of cultural complexity large enough to consider this study valid when talking about cultural differences.

3. Corpora and Tools

Specialized corpora are extremely useful because they allow us to obtain lexically homogeneous data. General purpose corpora (e.g. the LOB Corpus) are less useful for investigating cultural issues because they do not focus on specific fields. However, a general corpus can be used as a reference corpus to compare with the more specialized ones to obtain a keyword list, that is a list of words having unusual frequencies in comparison to the reference corpus (see Partington, in this issue).

Corpus data can be processed by using a variety of software. Their basic functions include generating frequency lists, displaying the concordances of a given ‘node word’ and displaying the text passage in which the word occurs. The software used in this study is *TextSTAT-2*.

4. Methodology

The analysis carried out in this study is based on two interrelated methodologies: the three-step methodology developed by Tognini-Bonelli (2001) to identify functionally equivalent units of meaning, and the combined quantitative and qualitative approach in corpus analysis proposed by Manca (2011). Both methodologies are relevant to translation issues, as the former allows us to find translation equivalents through the analysis of meaning and context, whereas the latter takes into account the constraints due to cultural differences that might be involved once a unit of meaning becomes a unit of translation.³

³ While a unit of meaning can be defined contextually at the linguistic level, a unit of translation can only be defined strategically according to factors such as genre, context of situation, context of culture. (Tognini-Bonelli 2001)

The methodology developed by Tognini-Bonelli draws on Sinclair's theory of context and meaning (1996), according to which the meaning of words is strongly influenced by the context in which they occur. The three steps are the following:

1. a node word is investigated in its collocational profile and its main collocates are identified;
2. for each collocate a prima-facie translation equivalent is identified and its collocational profile is investigated;
3. an adequate translation equivalent of the initial node word is identified within the collocational range of the prima-facie translation equivalent.

By applying this methodology to a contrastive study on the concept of 'nature' in websites of Italian *Agriturismo* and British Farmhouse Holidays, Manca (2004) observes a lack of correspondence between the items *natura* and *nature* and identifies the pattern *set in the countryside* as the functional equivalent of the expression *immerso nella natura*.

In her cross-cultural studies Manca (2009; 2011) further exemplifies and develops Tognini-Bonelli's methodology by adding a fourth step which implies framing the results obtained by linguistic enquiry within Intercultural Studies. More specifically, she combines two methodological approaches: Sinclair's theory of meaning and context (1996) for the corpus data and HCC vs. LCC framework (Hall 1990) and Cultural Orientations (Hofstede 2001), as discussed in Katan (2004), for discussion of the results.

Hence, Manca (2009) suggests that the difference in the concept of 'nature' is due to the fact that the Italian culture, which is HC-oriented, tends to represent 'nature' as an abstract entity as to impress the readers, whereas the British culture, being more LC-oriented, tends to provide more objective descriptions.

In the next section we will apply these two methodologies to travel reviews.

5. What *TripAdvisor* Traveller Reviews can tell us about Culture

TripAdvisor is the world's largest travel community where users can post reviews written by other travellers.

The present corpus-based study aims to investigate how Puglia is perceived by *TripAdvisor* travellers through a quantitative and qualitative analysis of the language used in *TripAdvisor* reviews.

The study was carried out on two corpora: the first corpus is made up of reviews of accommodations in Puglia written by English travellers⁴, the second corpus is made up of reviews of accommodations in Puglia written by Italian travellers from all parts of Italy.

In the light of the observations made in section 2, a few remarks are necessary. Both the English and Italian reviews are in theory written in the users' native tongue. However, since personal profiles display the user's location rather than nationality, a few reviews are likely to be written by non-native speakers using English or Italian as their second language. Italian reviews written by non-native speakers were easily recognizable and eliminated. With regard to the English reviews, those written by travellers located in the UK but characterised by frequent mistakes in language and syntax were not taken into consideration. Nevertheless, the English corpus is likely to include a number of reviews written by non-native speakers using English fluently. Either way, this does not affect the validity of the research for two reasons: 1) reviews were selected carefully as to make sure that the vast majority of them were written by native speakers and 2) the potential presence of reviews written by non-native speakers using English fluently is important to this study since they actually reflect the cultural complexity mentioned by Hannerz himself.

The two corpora were assembled in the period ranging from September to December 2010. The Italian corpus has almost 108,000 running words (710 reviews); the English corpus has 101,000 running words (501 reviews). The texts contained in the corpora were subject to linguistic analysis by using *TextSTAT-2*.

The quantitative and qualitative analysis of linguistic data provided interesting results, which were framed within the HCC vs. LCC framework (Katan, 2004) and in particular Thinking Orientation (Walker et al. 2003, Katan 2004), Uncertainty Avoidance (Hofstede 2001, Katan 2004) and Affective Associations (Francesconi, 2007).

⁴ See footnote 1.

5.1 Thinking Orientation: linear vs. systemic

One of the first things researchers do when they are dealing with corpus analysis is to analyse the wordlists of the corpora involved in the study. The analysis of the wordlists is time-consuming but will usually give useful hints for further, more detailed investigation. What is most useful in this first step is to identify the most recurring items and group them into semantic fields for each corpus. In this way it is possible to compare the words occurring in each semantic field along with their frequencies, narrow the research according to what is most worth analysing and start drawing hypotheses about the cultural orientations that could be involved.

From the analysis of the English and Italian wordlists seven semantic fields were identified: ‘accommodation’, ‘services’, ‘surroundings/territory’, ‘food and drink’, ‘buildings’, ‘leisure activities’ and ‘price’.

From a first comparison of the two wordlists two main aspects can be observed. The first one is the use of generic, all-encompassing words in Italian reviews and of more specific words in English reviews. In Table 1 a few examples are displayed:

Semantic field		Italian term and no. of occurrences	English terms and no. of occurrences
buildings/ architecture	generic	struttura (<i>buildings</i>) – 370	building – 58
	specific	hotel/albergo (<i>hotel</i>) – 296 casa (<i>home/house</i>) – 118 (mostly metaphorical use)	hotel – 1200 house – 54 home – 46
surroundings/ territory	generic	natura (<i>nature</i>) – 62 verde (<i>green countryside</i>) – 55	nature – 4 green – 8
	specific	alberi (<i>trees</i>) – 8 giardino(i) (<i>garden/s</i>) – 44 terreno(i) (<i>grounds</i>) – 2 cortile(i) (<i>courtyard/s</i>) – 8	trees – 26 garden(s) – 76 grounds – 42 courtyard(s) – 20

Table 1. General Italian items vs. English specific items

If we look at the frequencies of all the English and Italian words related to the semantic field of ‘buildings/architecture’ (except *casa*, which mainly refers metaphorically to ‘home’ as in ‘feel at home’), we

will notice that there are about twice as many occurrences of the English *hotel*⁵, *building* and *house* compared to the Italian *hotel*, *albergo* and *struttura* (1312 vs. 666).⁶ A first possible explanation could lie in the tendency of the English language towards repetition compared to Italian (Scarpa, 2008: 156). Furthermore, in a number of occurrences the word *hotel* is used as an adjective in English, e.g. *hotel staff* (11) and *hotel grounds* (5). In order to better understand the reason underlying this quantitative difference, the concordances of the word *hotel* were analysed. The analysis showed that, unlike the Italian reviewers, the English tend to use the expression “the hotel has” (23 occurrences) to introduce items or describe the hotel features:

“The hotel has indoor and outdoor pools...” (English Text No. 314)

“The hotel has a garden roof terrace...” (English Text No. 325)

In the Italian corpus, equivalent expressions such as *l'albergo/hotel/struttura ha/è dotato/a di* occur overall only 8 times. In fact, in the Italian corpus, introducing the items and features of a hotel does not appear to be important. A general assumption might be that while in the English reviews the items and features characterising the accommodation tend to be the “rheme” of the sentence (i.e. new information), in the Italian reviews these tend instead to be the “theme” of the sentence, (i.e. given information, Halliday 1985: 38).

However, the validity of this hypothesis can only be verified with larger corpora and by adopting a methodology which also includes non software-aided corpus analysis.

Furthermore, the analysis of the words *hotels* and *alberghi* (in the plural form) suggests that English reviewers tend to make comparisons with other hotels more frequently than Italian reviewers (about 30 occurrences in the English corpus vs. 10 occurrences in the Italian corpus).

“Breakfast was fine (better than in most Italian hotels)...” (English Text No. 74)

⁵ In a number of instances the word *hotel* is part of the name of the accommodation (e.g. “Hotel Pizzomunno”). Since this use is not relevant to our purpose, these occurrences were not taken into account.

⁶ The word *home* was not taken into account as it collocates mostly with *made/cooked* and with the verb *return*.

However, as mentioned above, in order to confirm these tendencies this would need testing with further, more detailed research with larger corpora.

Italian reviewers tend to use the generic word *struttura* very frequently. This word appears almost exclusively in the text rather than in the title of reviews and is used as a synonym for the noun *hotel*. So, for example, the English style tends to simply repeat the specific term, as in the following example:

“The hotel has a secure car park...” (English Text No. 17)

whereas the Italian style tends to look for a hypernym:

“La struttura è accogliente ed elegante, con ampio parcheggio interno.”
[The building is elegant and welcoming, with a big indoor car parking.]
(Italian Text No. 637)

As far as the words *casa* and *house* are concerned, even though the former displays a higher frequency than the latter, they are not used in the same way. Indeed, while *house* refers to buildings, *casa* occurs frequently in the expressions *padroni di casa* (owners) and *sentirsi a casa* (feel like at home) and tends to refer to home environment rather than buildings:

“The house is lovely as are the rooms.” (English Text No. 200)
“Mario e Silvana ci hanno fatto sentire a casa nostra.” [Mario and Silvana made us feel like at home...] (Italian Text No. 388)

A similar pattern occurs also in the description of surroundings. While Italian travellers tend to use generic words such as *natura* (nature) and *verde* (green countryside), English travellers tend to specify what the surroundings consist of. Indeed, the words *trees*, *garden/s*, *grounds* and *courtyard/s* display high frequencies in the English corpus, whereas their Italian equivalents (*alberi*, *giardino/i*, *terreni*, *cortile/i*) display lower frequencies. An exception is the Italian word *giardino/i* which, similarly to its English equivalent *garden/s*, displays a high frequency (44). However, the word *grounds*, literally *terreni* or less likely *cortile/i* (courtyard/s, yard/s), is also a possible alternative to *garden/s*. And, if we sum the number of occurrences of the words *grounds* and *garden/s* we obtain 118 occurrences, nearly three times as many as *giardino/i* (plus *terreni* and *cortile/i*). This

suggests that English reviewers tend to focus on the specificity of (the planted) *gardens* more frequently than their Italian counterparts.

The second aspect resulting from the comparison of the wordlists is that the Italian wordlist displays a series of abstract nouns, whose potential English equivalents either do not appear at all in the English wordlist or display very low frequencies.

These are listed in Table 2:

Italian abstract nouns	Frequency
Cortesia (<i>Kindness/Politeness</i>)	78
Disponibilità (<i>Helpfulness</i>)	78
Gentilezza (<i>Kindness</i>)	67
Professionalità (<i>Professionalism</i>)	53
Cordialità (<i>Cordiality</i>)	44
Pulizia (<i>Cleanliness</i>)	110
Pace (<i>Peace</i>)	36
Tranquillità (<i>Tranquillity</i>)	58
Gusto (<i>Taste</i>)	43
Bellezza (<i>Beauty</i>)	54

Table 2. Italian abstract nouns

Following Manca's study (2007), the abstract nouns occurring in the Italian corpus were analysed in detail in order to determine their main collocates, which are displayed in Table 3:

Italian abstract nouns	Collocates
cortesia (<i>kindness/politeness</i>) cordialità (<i>cordiality</i>) disponibilità (<i>helpfulness</i>) gentilezza (<i>kindness</i>) professionalità (<i>professionalism</i>)	staff personale (<i>personnel/staff</i>) proprietari (<i>owners</i>)
pulizia (<i>cleanliness</i>)	camere (<i>rooms</i>) stanze (<i>rooms</i>)
tranquillità (<i>tranquillity</i>), pace (<i>peace</i>)	pace, tranquillità oasi (<i>oasis</i>)
bellezza(e) (<i>beauty/beauties</i>)	natura, posto, luogo, struttura (<i>nature, place, building</i>)
gusto (<i>taste</i>)	arredato (<i>furnished</i>)

Table 3. Collocates of the Italian abstract nouns

From Table 3 it is possible to identify three main semantic fields:

1) people's qualities, with the nouns *cortesìa*, *cordialità*, *disponibilità*, *gentilezza* and *professionalità* tending to co-occur in the description of the hotel staff:

“...e può sufficientemente descrivere la professionalità, disponibilità e cortesìa di tutto lo staff...” [...and this is enough to describe the staff's professionalism, helpfulness and courtesy...] (Italian Text No. 161)

2) indoor environment, with the noun *pulizia* (cleanliness) used to describe rooms and accommodations in general and the noun *gusto* (taste) referring mainly to furniture:

“...pulizia delle camere e di tutti i luoghi comuni eccellente.” [...the cleanliness of the room and of all common spaces was excellent.] (Italian Text No. 610)

“La camera a noi riservata era pulita e arredata con gusto...” [The room we booked was clean and tastefully furnished...] (Italian Text No. 651)

3) outdoor environment, with the nouns *pace* and *tranquillità* occurring together to describe places in general and the noun *bellezza(e)* referring to the surroundings:

“...è un luogo incantato per chi cerca pace e tranquillità.” [...it's a charmed place for those seeking peace and tranquillity.] (Italian Text No. 274)

“...la bellezza della campagna e degli ulivi secolari.” [...the beauty of the countryside and centuries-old olive trees.] (Italian Text No. 24)

Then, the prima-facie translation of each noun and its frequency were investigated in the English corpus:

Node word	Prima-facie Translation	Prima-facie Translation frequency
Cortesìa	Kindness/Politeness	0
Disponibilità	Helpfulness	0

Gentilezza	Kindness	0
Professionalità	Professionalism	0
Cordialità	Cordiality	0
Pulizia	Cleanliness	6
Pace	Peace	17
Tranquillità	Tranquillity	0
Gusto	Taste	6
Bellezza(e)	Beauty/Beauties	4

Table 4. Prima- facie translations

As can be seen in Table 4, in most cases the prima-facie translations of the Italian abstract nouns were not found in the English corpus. The only exceptions are the noun *peace* (17), which displays a few occurrences with the nouns *quiet* and *solitude*, the noun *beauty*, which occurs only in the singular form (4) and refers to beauty treatments, and the noun *taste*, which occurs only 6 times. This could mean that either the concepts conveyed by the Italian abstract nouns are not expressed in the English corpus or they are expressed in a different way.

The English equivalents of the collocates of the Italian abstract nouns were then investigated.

Table 5 shows the functional equivalents of the Italian abstract nouns. These were retrieved by analysing the collocational profiles of the English equivalents of the words occurring with the Italian nouns:

Italian Corpus abstract nouns	Italian Corpus collocates	English Corpus equivalents of collocates	English Corpus collocates + frequencies
cortesia, cordialità, disponibilità, gentilezza, professionalità	staff personale proprietari	staff owners	friendly – 183 helpful – 140 kind – 11 professional – 8
pulizia	camere stanze	rooms	clean – 193
tranquillità	pace	peace	quiet (adj.) – 75 (indoor and outdoor environment) quiet (noun) – 4
bellezza(e)	luogo, natura	words related to surroundings	beautiful – 153
gusto	arredato	furnished/furniture	tasteful/tastefully –14

Table 5. Functional equivalents of the Italian abstract nouns

Thus, we can say that the concepts expressed in the Italian corpus by abstract nouns generally tend to be expressed in the English corpus by their corresponding adjectives, most of which display high frequencies. Exceptions are the adjective *professional*, whose low frequency leads us to think that for English travellers it could be more important that the staff be friendly and helpful rather than professional, and the words *tasteful* and *tastefully*, which overall display only 14 occurrences but are used in the same way as *gusto*.

The Italian equivalents of the adjectives found in the English corpus, e.g. *disponibile* (helpful), *professionale* (professional), *gentile* (friendly), *pulito* (clean), *tranquillo* (quiet) etc. were then investigated in the Italian corpus. They all display high frequencies and we may deduce that while adjectives referring to the semantic fields mentioned above tend to occur in both English and Italian reviews, abstraction is a feature characterising Italian reviews only.

Overall, these findings are in line with Manca's (2007), who suggests that the frequent use of adjective nominalisation in Italian is due not only to the Italian language system but also to the "need for abstraction which characterises the Italian language" (*ibid.*).

On the basis of all the data analysed so far, the hypothesis we suggest is that Italian travellers and English travellers tend to provide travel information in different ways. More specifically, we may assume that, due to the use of all-encompassing words and abstraction, Italian reviews tend to be more generic in descriptions. In contrast, due to the use of more specific words and a more factual style, English reviews tend to be more detailed and itemised. As to exemplify this hypothesis, we will analyse the following examples:

- a) "Il posto è incantevole: immerso nella natura..." [The place is wonderful: set in green countryside] (Italian Text No. 269)
- b) "...set in the countryside, with lots of olive trees and tobacco plants..." (English Text No. 164)
- c) "... un'oasi di pace, cortesia, disponibilità..." [...an oasis full of peace, kindness and helpfulness...] (Italian Text No. 23)
- d) "It's very quiet and has a pleasant garden and a nice little pool." (English Text No. 498)

As can be noticed, there is a great difference between examples a) and b) in the way green spaces are described. In example a) the all-encompassing word *natura* (nature) does not provide any information

about what the surroundings are made of. Furthermore, the expression *immerso nella natura* suggests a tendency towards a concept of ‘nature’ as an abstract entity, as widely discussed in Manca (2009). In contrast, example b) shows a more detailed and factual description of the surroundings (*countryside, olive trees, tobacco plants*). Similar differences can be noticed also between examples c), and d), with the former characterised by abstraction and indefiniteness (*oasi di pace*) and the latter by a more detailed and denotative description (*quiet, pleasant garden, nice little pool*). Obviously, more detailed descriptions could be found also in Italian reviews, but descriptions similar to examples a) and c) tend to occur far more frequently.

In order to further exemplify our hypothesis, we will provide an example related to the use of the adjectives *bello/bellissimo* (512) and *beautiful* (135) in the description of rooms.

In the Italian corpus, the adjectives *bello* and *bellissimo* (beautiful) are often used to describe hotel rooms:

“Le camere sono poi davvero bellissime.” [The rooms are really beautiful] (Italian Text No. 551)

In the English corpus, instead, the adjective *beautiful* refers mainly to the surroundings:

“Set in beautiful grounds of olives, oak and lavender.” (English Text No. 191)

Since the English corpus shows that the adjective *beautiful* is not associated with rooms, it would be interesting to find out what the equivalent of *bello* in room description is. The corpus shows that the word *room* often occurs with the adjective *lovely*, as in the following example:

“A lovely room with a balcony sea view in a peaceful hotel...” (English Text No. 50)

Therefore, in the description of rooms the functional equivalent of *bello* is *lovely*.

The adverb *beautifully* was analysed, too. The corpus shows that this adverb is used with adjectives such as *appointed, restored, furnished, decorated*.

Here are a few examples:

“The suites are spacious and beautifully appointed.” (English Text No. 64)

“The jacaranda suite was beautifully furnished with antiques.” (English Text No. 90)

Since the adverb *beautifully* occurs with words associated with rooms, it could be considered an equivalent of *bello/bellissimo*. However, there is a significant difference in the way rooms are described in the two corpora. Italian reviewers tend to describe rooms as simply *belle/bellissime* without saying what in particular was ‘beautiful’ about the room (e.g. furniture, decoration, etc.). In contrast, English reviewers tend to explain what aspect of the room was ‘beautiful’:

beautifully decorated = the decoration was beautiful

beautifully furnished = the furniture was beautiful

These results can be discussed in terms of Thinking Orientation (Walker et al. 2003, Katan 2004). Thinking style can be either systemic or linear. According to Walker et al. (ibid.), systemic thinking “places value on the context and not just on data; it assumes that without understanding the context, we have no real understanding of what the data mean” (2003: 50). As Katan suggests, while systemic thinking is “holistic, and tends to look at the full picture” (2004: 243), linear thinking tends to “look for detail, precision and minute cause and effect” (ibid.). In Italian reviews, the use of general, all-encompassing words and the use of nominalization suggest that Italian travellers tend to focus on the full picture and grasp only its general features. In contrast, English reviews tend to be more detailed and itemised.

These findings reflect Hall’s classification of Italy and England within the HCC vs. LCC framework. Systemic thinking is a feature characterising higher context cultures, like the Italian culture, whereas linear thinking is characteristic of lower context cultures, like the English culture (Katan, 2004).

These data show that our hypothesis can be considered valid, and they also confirm the validity of semantic field analysis as a methodology for the investigation of language patterns and cultural reasons underlying them.

The features described in this section could be linked to differences in the length of reviews. A previous analysis of the reviews (Fina, 2011) showed that English travellers actually tend to write longer reviews than

Italian ones, with a number of words ranging from 100 to 400 words considerably outnumbering Italian reviews with the same number of words (136 English reviews vs. only 72 Italian reviews)⁷. This tendency might appear to be related to KISSy (Keep It Short and Simple) and KILCy (Keep It Long and Complete) styles (Katan, 2004). More specifically, we might assume that English reviews are longer because English travellers tend towards a KILCy style, whereas Italian reviews are shorter because Italian travellers tend to use a KISSy style. However, the driving factor here is not the KISSy/KILCy difference but the linear vs. systemic orientation. Indeed, the KISSy and KILCy styles identified by Katan refer to sentence level, whereas the linear and systemic styles identified respectively in the English and Italian reviews operate at discourse level. Hence, the terms “Text KISSy” and “Text KILCy” would be more appropriate here, with the previous referring to the Italian travellers’ tendency towards abstraction and generic descriptions, which probably make reviews shorter, and the latter referring to the English travellers’ tendency towards more detailed, concrete and factual descriptions, which include more topics and make reviews longer.

5.2 Affective Associations: Uncertainty Avoidance

In tourism advertising the destination is often represented as the ‘dream home’ or as a place where tourists will ‘feel like at home’. This aspect can be analysed from the travellers’ point of view in order to see whether they do feel like at home when they are in Puglia. To do so, the words *casa* and *home* were investigated. These words occur 109 and 46 times respectively. What is interesting for us is to check how many times the words *casa* and *home* occur in the expressions “sentirsi (come) a casa” and “feel like at home” or similar. In the Italian corpus the expression *sentirsi a casa* (in all its variants) occurs 39 times. Here is an example:

“Ci siamo sentiti come se fossimo a casa nostra.” [We felt as if we were at home.] (Italian Text No. 38)

⁷ Since *TextSTAT-2* does not help with the calculation of the number of words in single texts, a sample analysis was necessary. 200 Italian reviews and 200 English reviews were selected and the number of words for each review was determined using the tools provided by Microsoft Office Word.

In the English corpus the expression *feel like at home* (in all its variants) occurs only 13 times.

A similar pattern can be noticed also in the use of the words *famiglia* and *family*, which can be considered to belong to the semantic field of 'home'. The analysis of the concordances of *famiglia* (93) shows that this word often occurs in sentences such as "sentirsi in famiglia" (to feel like a family member), as in the following example:

"Siamo stati accolti come persone di famiglia." [We were welcomed as if we were family members.] (Italian Text No. 386)

In the English corpus the word *family* occurs 56 times, but is used mainly in reference to travel type (*family holiday*) and to people managing the accommodation (*family-run*).

These findings can be discussed in terms of 'environmental bubble'. Cohen (2004) investigates the tourists' tendency to expose themselves to the strangeness of the host environment or to encapsulate themselves "within the familiarity of the 'environmental bubble'" (2004: 183). In particular, he suggests that while in organised mass tourism tourists are "little exposed to danger" (ibid.), tourists travelling alone or with families frequently leave their environmental bubbles and expose themselves to strangeness. Thus, travellers tend to recreate their environmental bubble within the tourist establishments. Being an essential part of the services offered by the accommodation, the hotel staff can play a significant role in recreating the travellers' environmental bubble, up to a point that Italian travellers tend to let the hotel staff temporarily enter their private sphere.

The data obtained can be further interpreted in terms of what Francesconi (2007) defines *affective associations*, that is "a reflection of the personal feelings and attitudes of the speaker/writer towards the target of the utterance" (2007: 45). Further references to feelings and relational issues could be elements such as salutations (*un saluto a...*) and thanking (*grazie*), whose equivalents do not occur in the English corpus.

The need for a 'home away from home', the tendency to recreate one's own environmental bubble, and affective associations could be indicators of Uncertainty Avoidance, which Hofstede defines as "the extent to which the members of a culture feel threatened by ambiguous or unknown situations" (2001: 167). Indeed, not only do these elements represent the linguistic expression of feelings and emotions which are usually associated with home environment, but they also indicate a

tendency to reject what is unknown and unfamiliar. Hofstede considers the UK as a low uncertainty avoidance culture. Indeed, its score (35) can be considered very low if compared to that of Italy (75). On the basis of this classification we may assume that, unlike English travellers, Italian travellers tend towards a low tolerance for unknown situations and when they are on holiday they tend to search for a ‘home away from home’, that is they look for familiar clues that will help them feel comfortable in a new environment. In contrast, having a higher tolerance for the unknown, English travellers tend to look for a ‘home away from home’ to a lesser extent than Italian travellers. Instead, what they seem to express is a nostalgic longing for a past time (Francesconi 2007, Dann 1996) where what is ‘old’ is valuable, as the use of the adjective *old* with a positive semantic prosody (*beautiful, lovely*) suggests:

“Lecce itself is fabulous and there is lots to do in the area - beautiful old churches, castles, coastal walks.” (English Text No. 92)

All the findings obtained in this research can be considered the linguistic expression of what Urry (2002) defines the “tourist gaze”, that is the way travellers perceive, interpret and evaluate the holiday experience according to their model of reality.

6. Conclusions

The main assumption in this work is that in a travel community the travellers’ values and beliefs affect the language they use in reviews. In order to understand in what way cultural orientations affect the travellers’ perception we applied a methodology based on a combined quantitative and qualitative approach (Manca 2011).

In Italian reviews aspects such as abstraction, generic descriptions, the need for a ‘home away from home’ and affective associations are in line with Hall’s classification of the Italian culture as HC-oriented. In contrast, in English reviews aspects such as more detailed descriptions and few references to relational issues seem to reflect the classification of the English culture as more LC-oriented.

However, as in all other studies, these findings must be interpreted in terms of general tendencies and their validity cannot be extended to other domains, since what is true in a corpus focused on a given domain might not be true in corpora focused on different domains, on different

areas of the same domain or, most importantly, in other types of communication (i.e. transactional vs. interactional, Katan 2004).

Furthermore, we are aware of factors, due to cultural complexity and globalization, which might determine significant differences also within the same ‘culture’, though we believe they will not be as significant as those across cultures.

This issue is even more significant in a context such as *TripAdvisor*. Travel reviews differ from promotional texts in that they are not written by travel experts but by non-specialized travellers with different backgrounds and limited common interests (apart from travel and writing opinions for *TripAdvisor*). Nevertheless, as the results show, in spite of the diversity and variety characterising the reviews, through corpus-based analysis it was possible to retrieve commonly shared values from the language.

In terms of areas for further research, it would be fruitful to compare travel reviews with promotional texts such as brochures and guidebooks in order to see whether and how they differ in terms of language patterns and cultural issues. This type of analysis would be possible by using the *keyword* function of *WordSmith Tools*, which would allow us to use a corpus made up of promotional texts as a reference corpus and obtain a list of words displaying unusual frequencies compared to this reference corpus. However, first it would be necessary to enlarge the two *TripAdvisor* corpora.

References

- Cohen, E. 2004. *Contemporary Tourism: Diversity and Change*. Amsterdam: Elsevir.
- Cucchi, C. 2010. “Hofstede’s cultural dimensions: Italian national identity in ELF usage”. *Cultus. The Journal of Intercultural Mediation and Communication*, 2010, 3. Terni: Iconesoft, pp. 137-158.
- Dann, G. M. S. 1996. *The Language of Tourism. A sociolinguistic perspective*. Wallingford: CAB International.
- D’Egidio, A. 2010. “Eyeing Puglia: comparing the tourist gaze in English, Italian and German travel articles”. *Rivista Internazionale di Tecnica della Traduzione*, 2010, EUT, n.11/2009, pp. 201-211.
- Fina, M. E. 2011. *TripAdvisor and Trust: an investigation of the phenomenon*. Thesis, University of the Salento (Puglia, Italy), unpublished.

- Francesconi, S. 2007. *English for Tourism Promotion. Italy in British Tourism Texts*. Milano: Ulrico Hoepli Editore.
- Hall, E. T. 1983. *The Dance of Life*. New York: Doubleday.
- Hall, E. T. 1990. *The Silent Language*. New York: Doubleday.
- Halliday, M. A. K. 1985. *An Introduction to functional grammar*. London: Edward Arnold.
- Hannerz, U. 1992. *Cultural complexity: studies in the social organization of meaning*. New York: Columbia University Press.
- Hannerz, U. 1996. *Transnational connections. Culture, people, places*. London: Routledge.
- Hofstede, G. 2001. *Culture's Consequences. Comparing Values, Behaviours, Institutions and Organizations Across Nations*. Thousand Oaks: Sage Publications.
- Katan, D. 2004. *Translating Cultures. An Introduction for Translators, Interpreters and Mediators*. 2nd Ed., Manchester: St. Jerome.
- Katan, D. 2006. "It's a Question of Life or Death: Cultural Differences in Advertising Private Pensions". In N. Vasta (Ed.), *Forms of Promotion: Texts, Contexts and Cultures*, Bologna: Patron Editore, pp. 55-80.
- Katan, D. (forthcoming) "Translating the tourist gaze: from heritage and 'culture' to actual encounter". *Pasos Journal of Tourism and Cultural Heritage*.
- Manca, E. 2004. "The Language of Tourism in English and Italian: Investigating the Concept of Nature between Culture and Usage". In *ESP Across Cultures: Volume 1, Single Issue*, pp. 53-65.
- Manca, E. 2007. "Beauty and Tranquillity in the Language of Tourism: Linguistic and Cultural Reasons". In De Stasio C. & Palusci O. (Eds.), *The Languages of tourism. Turismo e mediazione*, Milano: Edizioni Unicopli, pp. 113-128.
- Manca, E. 2009. "Aspetti culturali a confronto: la promozione turistica in Italia e Gran Bretagna". In G. L. De Rosa e A. De Laurentiis (Eds.), *Lingue policentriche a confronto: quando la periferia diventa centro*, Milano: Polimetrica Publisher, pp.153-164.
- Manca, E. 2011. "Corpus Linguistics and Cultural Studies: a combined approach in the translation process". In M. Bondi (Ed.), *RILA, Rassegna Italiana di Linguistica Applicata*.
- Orlando, C. 2010. "Blogging about London: comparing the Italian and Anglo-American tourist gaze". *Rivista Internazionale di Tecnica della Traduzione*, 2010, EUT, n.11/2009, pp. 189-199.

-
- Partington, A. 2011. "Corpus Linguistics: What it is and what it can do". *Cultus. The Journal of Intercultural Mediation and Communication*, 2011, 4. Terni: Iconesoft
- Scarpa, F. 2008. *La traduzione specializzata. Un approccio didattico professionale*. 2nd Ed., Milano: Ulrico Hoepli Editore.
- Sinclair, J. 1996. "The Search for Units of Meaning". *Textus* IX (1): 71-106.
- Spinzi, C. 2010. "How this holiday makes a difference': the language of environment and the environment of nature in a cross-cultural study of ecotourism." In *Quaderni del CeSLiC. Occasional papers*, Centro di Studi Linguistico-Culturali (CeSLiC), Bologna.
- Stubbs, M. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell Publishing.
- Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam: Benjamins.
- Urry, J. 2002. *The Tourist Gaze*. 2nd Ed., London: Thousand Oaks, New Delhi: SAGE Publications.
- Walker, D., Walker, T. & Schmitz, J. 2003. *Doing Business Internationally: The Guide to Cross-Cultural Success*. 2nd Ed., New York: McGraw Hill.
- Wolfe, M. L. 2008. "Different cultures – Different discourses? Rhetorical patterns of business letters by English and Russian speakers". In U. Connor, E. Magelhout & W. Rozycki (Eds.). *Contrastive Rhetoric. Reaching to intercultural rhetoric*, Amsterdam: John Benjamins, pp. 87-121.
- Zappettini, F. 2010. "Negotiating LGBT identities in Italy: an intercultural perspective". *Cultus. The Journal of Intercultural Mediation and Communication*, 2010, 3. Terni: Iconesoft, pp. 115-135.