

The Formation of the Image of Top-Ranked Hotels through Real Online Customer Reviews: A Corpus-Based Study of Evaluative Adjectives as Image-Formers/Providers

Nuria Edo Marzá¹

¹ Faculty of Humanities and Social Sciences, Department of English Studies, Universitat Jaume I, Castellón, Spain

Correspondence: Nuria Edo Marzá, Universitat Jaume I, Faculty of Humanities and Social Sciences, Department of English Studies, Avda. Sos Baynat s/n, 12071, Castellón, Spain. Tel: 34-964-72-96-12. E-mail: nedo@ang.uji.es

Received: April 17, 2013 Accepted: May 20, 2013 Online Published: July 17, 2013

doi:10.5539/ijel.v3n4p15

URL: <http://dx.doi.org/10.5539/ijel.v3n4p15>

This research has been carried out within the framework and support of a research project awarded by the Ministerio de Economía y Competitividad (former Ministerio de Ciencia e Innovación), with reference FFI2011-2472, and entitled “Análisis léxico y discursivo de corpus paralelos y comparables (español, inglés y francés) de páginas electrónicas de promoción turística”. 2011-2014.

Abstract

This study is grounded on a corpus-based analysis of evaluative adjectives as hotel image-formers/providers conducted from a multifaceted yet complementary perspective which departs from the notion of destination image formation (Gartner, 1993; Baloglu & McCleary, 1999, Gylling, 2004 among others) focused on a specific aspect, that of the hotel of destination, and to a specific genre, the online (hotel) customer review. The corpus of study, which contains more than 250 000 words, is made up of more than 1200 online reviews of top-ranked hotels retrieved from the popular website *Tripadvisor.com*. WordSmith Tools 5.0 was used to analyse and categorise the evaluative adjectives in the corpus from three perspectives: firstly, adjectives were initially classified according to frequency and/or saliency criteria; secondly, in terms of their implicit degree of positivity/negativity; and thirdly, according to the kind of evaluation provided. Subsequently, the main features evaluated in any hotel online customer review (value, rooms, location, cleanliness, and service) and their frequently associated adjectival evaluative acts were analysed to determine their collocational behavior and thus the way such appraisal patterns contribute to “destination image formation”, that is, to form an impression on specific aspects of top-ranked hotels. The paper is aimed at showing that evaluative adjectives in online customer reviews are critical in the interpretation of subjective expected values, and that a correct understanding of the kind of image they transmit in the tourist-to tourist digital interaction through reviews provides an opportunity for destination marketers, in this case hoteliers, to enhance the psychological and symbolic values as well as the benefits of the destination according to improved communication strategies and language usage.

Keywords: evaluative adjective, corpus, online review, collocational behavior, frequency, saliency, evaluation, WordSmith Tools

1. Theoretical Framework and Aims of the Study

1.1 The Notion of “Destination Image Formation”

This study is primarily framed within current research on destination image formation processes (Gartner, 1993; Baloglu & McCleary, 1999, Gylling, 2004, among others). Due to the enormous scope of the “destination image formation” concept and to the specific aims of this study, a concrete aspect of such notion has been chosen for further analysis, that of accommodation, and more specifically that of top-ranked hotel image. Therefore, the theoretical approach to this study has departed from and been constructed on the basis of three main standpoints – (1) the study of “destination image formation (2) through evaluative adjectives (EAs) (3) in a specific digital genre, the online customer review (OCR) –, with the final aim of showing how a correct understanding of the way image is portrayed and transmitted in tourist-to-tourist online interactions through reviews may provide an

opportunity for destination marketers, in this case hoteliers, to learn to enhance the psychological and symbolic values as well as the benefits of the destination according to what potential customers perceive as reliable linguistic use and consider relevant aspects. Contrary to what many hotel marketers seem to believe, the formation of a good/positive destination image does not necessarily have to imply a too blatant over-exaggeration of features and qualities, as some web pages or other promotional genres seem to reflect. Praise is good and an absolute must in promotional genres but clients do also distrust too sweetened descriptions. As online reviews show, the potential client resorts to OCRs precisely because he/she is looking for a more reliable, real and realistic description of hotels, something hotel marketers should take into account when designing their promotional webpages or brochures despite acknowledging the fact that, obviously, the purpose of promotional genres and OCRs is not the same.

As Garzone (2009, p. 31) states “tourism is a first-hand experience, which cannot be tried before it is consumed, it can only be prefigured on the basis of the information available”. Therefore, prospective clients are increasingly interested in looking for information in which no persuasive purpose is inherent, and they find in real customers’ online reviews the best alternative to find the kind of non-biased information they look for before making a decision. That is the reason why OCRs have been the genre chosen for this analysis, because of their immediate, real and non-interested nature which provides an interesting approach to “non-interested” image formation processes. Tourism researchers and academics agree in the fact that a positive destination image results in increased visitation and has an impact upon tourists’ evaluation and selection of destination (Hall, 2003, Morgan & Pritchard, 1998, Echtner & Ritchie, 2003, Baloglu & McCleary, 1999 among others). This statement is logical and obviously true since highlighting the shortcoming or deficiencies of a destination would contravene the basic principles of marketing and go against what tourists are looking for but, at the same time, a closer, more realistic description of the destination (as happens in online reviews) – not so much noticeably determined by the notion of interestedness (Shaw, 2006) always pervading promotional genres – could result into a “franker” image formation process, more valued by the potential customer. Most probably, prospective clients are “too aware” of the fact that no negative aspect will ever be portrayed by the destination marketer and therefore look for more realistic, objective and reliable information from which to form an image/opinion and thus make a decision in a more equal-to-equal relationship of information sharing.

It was in the early 1970s when the tourism researcher Hunt (1975) demonstrated that the destination image plays a central role in tourists’ selection process (Cai, 2002; Echtner & Ritchie 2003). As its very same name indicates, tourism researchers and scholars define the destination image as ‘impressions of a place’ or ‘perceptions of an area’ (Echtner & Ritchie 2003: 41), which in this case would be circumscribed to ‘impressions of a hotel’ or ‘perceptions of a hotel’. Analysing how this image is formed through EAs during client (reviewer)-to-(potential) client(s) interaction may be useful and enriching in order to better understand what kind of image/impression/opinion motivates tourists hotel election process.

However, despite the studies aforementioned and the importance of the tourism industry worldwide, there are few empirical studies which have determined and analysed the factors that specifically influence the formation of the destination image (Baloglu & McCleary, 1999), with the exception of, for instance, Gylling’s (2004) study on the destination image of Singapore. This is quite surprising because ‘before image can be used to influence behaviour, it is important to understand what influences image...knowing factors influencing it would help to identify markets and decide which image should be promoted to which segment of the market’ (Baloglu & McCleary, 1999, p. 869). As regards these specific factors, Gartner (1993) signals cognitive, affective and conative components as the ones forming the destination image. However, within a destination marketing context it is mainly the cognitive image – ‘the sum of beliefs and attitudes of an object leading to some internally accepted picture of its attributes’ (Gartner, 1993, p. 193) – the one with a ruling role among destination marketers since it can directly influence the tourist and his/her selection process. This cognitive image is the easiest to control and influence and that is the reason why it should be the focus of attention of hoteliers. Baloglu and McCleary (1999, p. 869) did also develop a model (see Figure 1) illustrating the destination image formation and including more factors than those mentioned by Gartner (1993):

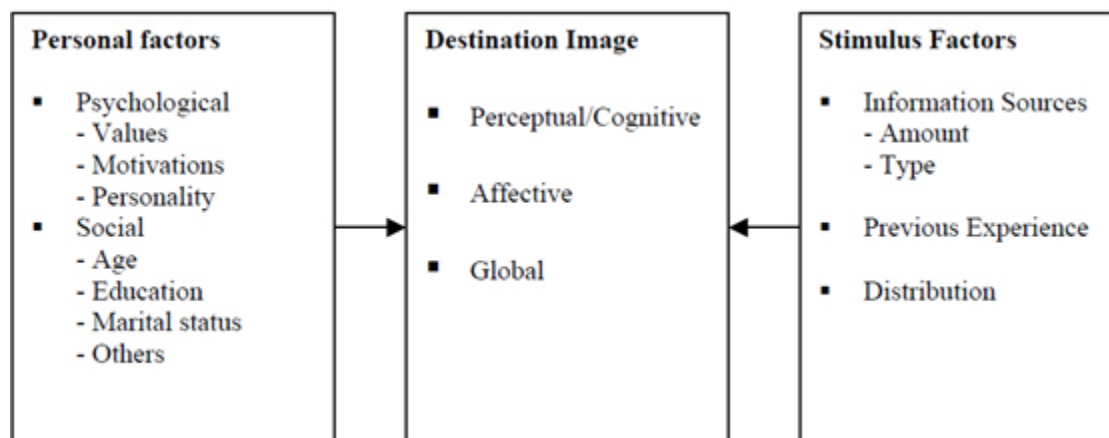


Figure 1. Destination image formation scheme

Description: A general framework of destination image formation according to Baloglu & McCleary (1999, p. 870).

The destination image formation is thus crucial in any destination (hotel) choice since, as Gartner's (1993) theory of destination selection process indicates – in this case extrapolated to hotel selection process-, all destinations are selectable ('total opportunity set') and it is this image what will determine this selection. The potential tourist narrows down the options to a final destination choice through a number of selection processes (Gylling, 2004). The analysis of some of these selection processes and their implications in impression formation processes from the perspective of the adjectives employed in OCRs has been the focus of this study. Only images that are carefully created and positioned in the minds of potential tourists 'survive' the evaluation process (Gylling, 2004) and adjectives very much determine and shape this. However, it cannot be forgotten either that what "good hotel" means to one person in an OCR may not be the same to what it means to another (demand, requirement and expectation parameters may vary from one person to another). Therefore, despite the fact that this small-scale study tries to provide an overview of the top-ranked hotel image formation process through adjectives, it must be kept in mind that particular reviewers and readers constitute individual entities which, despite normally being part of the same or at least similar interpretive communities (Fish, 1980), may interpret or react differently to the same kind of information.

1.2 Netlinguistics and Communication: The Online Customer Review

The development of Information and Communication Technologies (ICTs) in recent years has brought about a proliferation of new means of information exchange and debate that allow people to get in touch and interact with other people from all over the world in an immediate manner. Internet linguistics (Crystal, 2002) or netlinguistics (Posteguillo, 2002, 2003) present and analyse the Internet as a medium of computer-mediated communication (CMC) in which the traditional analytical levels of language, i.e., substance, form, discourse, situation, and ideology (Haynes, 1989), take on a new dimension and a new significance. As Crystal (2002) wisely portrays, the web has become the first port of call for both information enquiry and leisure activity. The impact of communication through the Internet is noticeable nowadays; it breaks conventions and favours genre hybridization and the development of frontier genres. Moreover, the direct participation of travelers questions the exclusive role of experts in information transmission (Calvi, 2010).

Internet forums and what could be called "interactive reviews" or online reviews are clear examples of computer-mediated information-exchange formats (genres), called digital genres or cybergenres, close to orality in many respects (especially in the case of forums), and with an astonishing potential for making information widely known immediately and with a great visual potential. Both of them are asynchronous in nature (unless participants happen to be connected at the same time so that communication can be synchronic) and, given the rationales of such genres, they are effective media for sharing data and information that users tend to consider true, non-biased, and reliable, since the interventions are authored, at least presumably, by independent "opinioners".

Following Gea Valor (2008), customer reviews (CRs) have their *raison d'être* and are becoming increasingly popular since they may be accepted more readily by an audience of peers. Real customers acting as reviewers

add to the presupposed accuracy and veracity of what is being reviewed, because the influence of what Shaw (2006) calls “interestedness” (the extent to which a recipient may recognise a genre as being used to persuade him/her to do something in benefit of the producer) seems to disappear or at least to be reduced to a minimum. Additionally, the simplicity of access, the scope of influence, the possibilities for updating, the interactive nature, and the constantly and rapidly increasing communicative potential of CMC among many other features make the web the ideal means and environment for customer reviews. This is also reinforced by the views of authors such as Pharo (2008), who stated that genres and their associated purpose, content, participants, time of use, place of usage, media, and form develop over time and this development depends on the community using and creating the genre, and thus on such a community’s needs -OCRs perfectly fitting the needs of tourists’ discourse community (Swales, 1990).

Tripadvisor.com (from now on simply Tripadvisor) is an Internet site containing varied information about hotels, flights, vacation rentals, restaurants, trip ideas, and even has a travel forum, among other utilities. It is generally regarded as the most widely known and visited site of these characteristics on the web and it is advertised as a site containing “unbiased hotel reviews, photos, and travel advice for hotels and vacations”. Tripadvisor is thus a web application housing and providing technical support for information, opinions, and discussion threads about trips (and related issues), in which users are allowed to freely express their views and opinions on such matters, provided that a basic set of rules on respect are obeyed. In this study we have focused on hotel OCRs as offered on the site. Hence, the kind of written material (textual samples) we worked with and compiled for the corpus of study can be illustrated with the sample screenshot shown in Figure 2.



Figure 2. Trip advisor OCR

Description: Sample OCR from Tripadvisor included in the corpus.

It can be observed at first sight that Tripadvisor reviews have a neat layout in which a lot of information is condensed into a relatively small physical space. Visual aids such as graphic ratings are used in order to offer the reader clear, to-the-point information that is easy to assimilate in case he/she does not want to go through the whole review. Thus, at the very beginning of the review we find a summarizing, general appraisal sentence that is always included as a “title” offering a preliminary evaluation and automatically, and rather unconsciously, generating a first impression on the reader. This summarizing sentence will already give the reader a first idea of the overall positive or negative character of the opinion/evaluation provided, which in most cases is more than

obvious and very straightforward (see also examples 1 and 2 from the corpus). It is this kind of information, provided by reviewers (former customers), that first helps the user gain an impression/form a general image of the reviewed hotel.

- (1) *Home away from home for a princess.*
- (2) *The most horrible hotel I have ever stayed in.*

Proceeding with the general layout of the OCRs analysed, a graphic general rating scale of satisfaction is provided to quickly corroborate and “quantify” the preliminary information provided in the title. Then, the exact date of the review is provided, thereby adding to the implicitly assumed rule that a CMC needs to be up-to-date in order to be considered trustworthy. Following this, some details about the reviewer are also provided on the left-hand side of the review so as to reinforce the presumed reliability and real existence of the person writing the review. Thus, the name (normally a nickname), a profile picture (real or chosen from a predetermined set), the city and country of origin (Note 1), the number of reviews previously written on the site, and the number of “helpful votes” previously assigned to the reviewer by other information-seekers are the kind of data that will help the user decide on the usefulness and reliability of the review. Then, the review as such is presented and the options “more” or “less” provide the reader with the option of looking at the full version of the review or just the initial part of it, while also helping to prevent the page from appearing cluttered.

In general, after reading just a few reviews and as expected, the user will realise that the main topics or parameters dealt with in hotel reviews have to do with value (for money), rooms (space, bathroom, distribution, standards of quality, commodities, accessories, views, etc.), hotel location, cleanliness, service (staff kindness, diligence and professionalism), and sleep quality. In fact, all the reviews include a specific rating system for each of these aspects, as shown at the bottom of Figure 2, where users have to value their impressions about the hotel they stayed in on a scale from zero to five marks. Furthermore, Tripadvisor also offers information on the date of the stay and the kind of trip, apart from giving the reader the opportunity to report inconvenient reviews and contributing to the reviewer’s “helpful votes” counter.

A good point about online review sites such as Tripadvisor is that they are interactive and feedback may be provided constantly. Contrary to what happens with more “traditional” reviews (normally about films, books or particular products), this *extant variant* cybergenre (Watters & Shepherd 1997) offers an interactive approach that takes advantage of the best of traditional genres but also incorporates all the possibilities offered by the new media. For instance, a especially interesting feature is the fact that Tripadvisor offers hotel managers the possibility of responding to customers’ reviews.

As regards the style and language used in online reviews, some types of *novel* (whether *emergent* or *spontaneous*) cybergenres like, prototypically, chats constitute instances of oral(ised) written texts (Yus 2001). In this respect, reviews *per se* cannot normally be considered pure instances of oral(ised) written texts but a considerable percentage of them do contain significant traits of orality, thus tending toward informality. In short and in general, it could be stated that we have detected two large profiles of reviewers: those more concerned with the language, style and spelling of their reviews, that is, those more concerned with formality (see example 3), and those less concerned with these aspects who prefer a more conversational, expressive form of communication that comes closer to orality and thus closer to oral(ised) written texts and informality (see examples 4 and 5).

- (3) *Many different nationalities but all with the same exemplary level of service blended with a genuine interest in their guests.*
- (4) *The one BUT is the lack of wifi. The hotel advertises and charges (why??) for in-room wifi.*
- (5) *Great first night in London! Gorgeous hotel, rooms were fantastic, bed soooo comfortable!*

In this paper, however, we are not trying to elucidate which tendency (formal or informal) is more consistently deemed trustworthy by review readers. The focus of this study is the observation and analysis of evaluative adjectives (EAs) as image or opinion-formers/providers in OCRs. However, the appearance of adjectives such as *fab* (11 instances in the corpus) and *comfy* (38 instances) in the corpus, as well as other purposefully misspelled formations of the *brilliantttttt** type, does nothing but confirm this tendency toward informality also in many EAs. According to this and as a brief overview of the main stylistic, formal and sociolinguistic features of cybergenres, it can be stated that they are giving way to a new kind of literacy. The views in this respect are varied and despite the catastrophist view of those thinking that “technospeak” or CMC in general are the beginning of the end of presently accepted standards, it must also be acknowledged that the Internet is fostering new and creative ways of communication. As Crystal (2005) pointed out, the key is to be aware that (within certain acceptable and reasonable limits) “if new informal varieties of language are emerging, then a fresh

relationship between nonstandard English and Standard English is likely". According to this same author, the properties that differentiate CMC from speech (from a formal perspective) include the lack of simultaneous feedback (critical to successful conversation), the absence of a nonsegmental phonology (or tone of voice, etc.), which can be partially imitated with other means such as capitalisation (see example 6 from the corpus), and its ability to carry on multiple interactions simultaneously (for example in chatrooms).

(6) *Honestly, If anyone is thinking of going there, please, AVOID IT!!!*

Also according to Crystal (2005), the properties which differentiate CMC from writing include its dynamic dimension (through effects such as animation and page refreshing), its ability to frame messages (as in email cutting and pasting), and its hypertextuality. From a sociolinguistic perspective, Crystal (2005) contended that CMC should be regarded as a positive fact, since "the internet is allowing us to once more explore the power of the written language in a creative way", whereas from a stylistic perspective, he considers the emergence of digital genres as a new stage in the evolution of the written language.

Finally, it must be kept in mind that the focus of this study is not on checking the accomplishment of linguistic quality standards or the in-depth analysis of online reviews as examples of oralised-written texts. Instead, it concentrates on the analysis of the way reviewers portray and give their opinions about hotels through EAs and on the way this opinionated portrayal may help prospective clients form an image about such hotel and thus help them make a decision as regards their choice of hotel.

1.3 The Role of Evaluative Adjectives as Opinion-Formers/Providers

Adjectives are words capable of condensing evaluations, classifying entities or describing qualities into a single lexeme. Despite the differences in the syntactic form adopted, all languages provide some means of modifying or elaborating the meanings of nouns (Fellbaum, Gross & Miller 1993). Adjectives are the most prototypical way to express the qualification of a noun despite the fact that words originally pertaining to other syntactic categories can also function as adjectives, as is the case of the present and past participle of terms and also of nouns premodifying other nouns.

On the other hand, "evaluation", as conceded by Hunston and Thompson (2000), is a slippery and complex notion which has been assigned several different labels depending on the author: Martin (2000) and Martin and White (2005) adopted the term *appraisal*, whereas Biber et al (1999) and Conrad and Biber (2000) used the term *stance*, and Barton (1993) preferred *evidentiality*. As has already been explained, evaluation cannot only be achieved by means of EAs but it is difficult to find an "easier", "more condensed" way of expressing appraisal than with adjectives. In the line of the aforementioned variation in the term "evaluation", adjectives with an evaluative role have also been assigned different nomenclatures, including the already-mentioned "evaluative" adjectives (Tucker, 1997; Hunston & Francis 2000; Swales & Burke 2003; Samson 2006), "attitudinal epithets" (Halliday 1985) or "subjective evaluative adjectives" (Kerbrat-Orecchioni 1980), among others.

In this study, we have used the term *evaluation* applied to adjectives in its broadest semantic sense, as understood by Hunston and Thompson (2000, p. 5): "[...] evaluation is the broad cover term for the expression of the speaker or writer's attitude or stance toward, viewpoint on, or feelings about the entities or propositions that he or she is talking about". In fact, the notion of EA adopted here would combine semantic, pragmatic, and functional criteria. Therefore, from now on throughout this paper, EAs will refer to those adjectival instances combining the Hallidayan concept of "attitudinal epithets" (expressing the speaker's subjective attitude toward the referent) and Kerbrat-Orecchioni's notion of subjective adjectives (including modal adjectives).

As this paper tries to show, EAs play a paramount role in the genre analysed, i.e., hotel OCRs, since the rationale of reviews is precisely to express opinion on the basis of a real true experience involving the item reviewed (in this case a hotel) in order to help others make a decision about how well such an item fits their objectives or subjective expected values. Evaluation is thus a key aspect in any opinion-related genre. Adjectives in general can account for merely objective but also for more subjective features and as such they can reveal much about the speaker's/writer's attitude toward the textual content, which is exactly what the review reader is looking for in order to form an impression/image of his/her own. This is why when convincing, reasoning, narrating, and telling, adjectives are always present and thus have a strong interpersonal dimension, especially in the case of EAs.

The choice of one evaluative adjective instead of another in the text when commenting key aspects about hotels may make the meaning of the text change completely. Hence, we firmly believe that EAs play a crucial role in tourists' destination image formation and decision-making processes. When it comes to forming a preliminary image of the suitability and convenience of booking a room in one hotel instead of another, reading about a

wonderful hotel obviously does not normally have the same effect as finding it described as a *decent hotel*. The evaluated entity (Hewings, 2004) is the same – the hotel – but the adjectival evaluative acts (Hewings, 2004) *wonderful* and *decent* are not expressing the same degree of excellence or satisfaction. In any case, evaluation on the part of equals seems to be considered more trustworthy than biased promotional discourse, even though “language is never neutral and texts are never innocent” and “all selections are ideological” (Stubbs, 1998, pp. 371-372).

Finally, a recurrently appreciated feature such as the super-positive or super-negative (although in a smaller proportion in our corpus) portrayal of certain aspects has also been analysed as a highly significant feature of EAs in the genre under analysis. As McCarthy and Carter (2004) stated, hyperbole (also referred to as exaggeration or overstatement in order to achieve some effect) has been studied in rhetoric and in literary contexts, but only relatively recently in banal, everyday contexts like the one under analysis here. Hyperbole is quite a common strategy among reviewers due to its strong emotional potential for the reader. Hyperbole, which always carries with it increased expressivity, may be produced in a wide variety of forms (also through simple, monolexematic EAs) and constitutes an instance of “extreme case formulation” (Norrick, 2004; Sidnell, 2004). These EAs that show increased intensity in their meaning carry what Martin and White (2005) have called *force*, which includes functions such as intensification and quantification, up-scaling, and maximisation. In this sense, as Edo Marzá (2012) concedes, the euphoria technique is frequently used in the tourism context, which makes the extended and hyperbolic use of adjectives a common trait also among hotel OCRs.

2. The Study

This study is grounded on a corpus of OCRs of top-ranked hotels from four different geographical locations, all of them having in common the fact that they were published online on the Tripadvisor website during June 2011. Thus, in this study an *ad-hoc* corpus of top-ranked hotel reviews amounting to 251 271 tokens was compiled. This corpus includes reviews of the top-10 most popular, the top-10 budget and the top-10 luxury hotels (according to the Tripadvisor official site) from New York, Paris, London, and Rome. More specifically, 10 reviews were gathered for each of the top-10 hotels in each of the three categories mentioned (popular, budget, and luxury) and in each of the four locations under analysis. We are aware of the fact that the composition of the corpus (OCRs of top-ranked hotels) implies a deliberate bias of the corpus toward positivity. This bias was created purposefully to fit the aim of the study, which is precisely the analysis of how the image of top-ranked hotels is formed via EAs (and the corpus was compiled accordingly). The focus on top-ranked hotels is due to the fact that, at least presumably, these reviews will show a higher lexical richness (obviously implying a tendency toward praise) than other more neutrally appraised hotels and this could be translated into a significant presence of EAs. Moreover, we also considered that both positive and negative extremes in evaluations would be more easily noticeable in top-ranked hotel reviews (despite the clear predominance of the former) and extremes in evaluation always facilitate and speed up, consciously or unconsciously, the image-formation process. Moreover, we believe that the analysis of negative evaluations in top-ranked hotels will be especially significant in the image formation of top-ranked hotels because of their unexpectedness for the reader, which will probably result in an increased specific weight of these negative evaluations in the impression-formation process.

Despite this purposeful “top-rank” bias, mixing different hotel categories when compiling the corpus responds to the desire and need to present a sample of language and reviewers’ profiles that is as real and varied as possible. Altogether, 1200 different hotel reviews were finally gathered in the corpus. For the purposes of this study, all the contributions compiled were checked to make sure that they were written in English by reviewers from English-speaking countries and that the quality of the language made the reviews understandable and at least average-acceptable from a linguistic and stylistic perspective.

The corpus was analysed and results retrieved with the software concordance program WordSmith Tools 5.0 (WST), which in turn consists of three applications: WordList, KeyWords, and Concord. In this study the three applications were used to perform the statistical analysis and obtain corpus frequency counts (WordList), to retrieve the list of keywords or words with significant saliency in the corpus (KeyWords), and to analyse the in-context behavior of EAs and hospitality-related key/allusive lexical items, mainly through concordance lines and collocates analysis (Concord).

Thus, the following specific aspects of the OCRs of top-ranked hotels as related with the image formation process through EAs were analysed with WST:

1. Frequency and saliency data as regards the most relevant and characteristic EAs in the corpus. On this basis we also tried to determine the character or quality of EAs in top-ranked OCRs according to a positivity *continuum* and to a functional, semantic and pragmatic classification of their values in the

corpus. This first classification from a triple vertex perspective constitutes a focus on the frequency, saliency, quality, and value of evaluations. The applications employed for this purpose were WordList and KeyWords.

2. The collocational behavior of EAs when portraying the main evaluated entities normally associated with the concept “hotel” and thus with the hospitality industry. This second aspect analysed constitutes a focus on the contextual behaviors, meaning, and form of the adjectives evaluating key hospitality-related aspects. The application used for this task was Concordance, and more specifically the tools Concord and Collocates.

With the kind of analysis undertaken and the results obtained, in the next section we intend to provide an enlightening view of the way top-ranked hotels are portrayed by reviewers in OCRs by means of EAs and thus on the way readers (potential customers) may interpret and consequently form an image of the hotel reviewed.

3. Results

3.1 Frequency, Saliency, Quality, and Value: A Proposed Classification of Evaluative Adjectives in the Corpus

As indicated in the WST “statistics” tab (see Figure 3), the frequency list obtained for this study shows a total of 9314 corpus types (Note 2) before preselection and a total of 8915 after the preselection of grammatical words. Of all these types, 3146 are instances of *hapax legomina* or words occurring just once in the corpus. From this frequency list obtained from the corpus, we focused on the first (top-frequency) 2000 word types, from which we later manually selected the top-frequency EAs used by Tripadvisor hotel reviewers talking about top-ranked hotels. Additionally, in this initial part of the study we also selected for further analysis those EAs regarded as key or salient by the Keywords list. Hence, the manual selection of EAs among these 2000 top-frequency types and keywords list in the corpus resulted in 249 EA types to be analysed.

Regarding tokens, the corpus as a whole has 251 271 tokens. From a perspective based purely on frequency, the 2000 top-frequency types accounted for 83 910 of these tokens; of these 2000 top-frequency types, 214 were instances of EAs amounting to 9709 tokens. Altogether (including both frequent and key/salient types), the 249 types belonging to the category of EA and analysed in this study amounted to 11 378 tokens. Of all the adjectives considered for the study, none of them presented a frequency lower than seven according to the Wordlist frequency list or below four according to the Keywords list.

	N	Overall
text file		Overall
file size		1,402,059
tokens (running words) in text		251,271
tokens used for word list		248,151
sum of entries		
types (distinct words)		9,314
type/token ratio (TTR)		3.75
standardised TTR		34.73
standardised TTR std.dev.		64.09
standardised TTR basis		1,000
mean word length (in characters)		4.34
word length std.dev.		2.30
sentences		15,065
mean (in words)		16.47
std.dev.		13.55
paragraphs		1
mean (in words)		248,151.00
std.dev.		
headings		
mean (in words)		

frequency | alphabetical | statistics | filenames | notes

77 | Type-in

Figure 3. Wordlist results

Description: Statistical data from the corpus provided by the WST tool Wordlist.

As initially expected, corpus counts showed that the number of EAs in the corpus analysed is significantly higher than that of more objective/descriptive adjectives, since the genre analysed is deeply subjective and opinionated in nature. The statistical data provided also reflects the fact that in this study we are working with a specialised corpus, since a specific genre characteristic from a specific domain and accomplishing specific aims within a given discourse community is being dealt with. Even though the degree of specialisation is not as high as that of other more prototypically specialised subject topics might be (the TTR of the corpus of study is 3.75 and its standardised TTR is 34.73), specialisation and specificity in content with respect to a reference corpus are easily noticeable. Regarding corpus composition, Wordlist frequency results seem to indicate coherence in the corpus compiled for the study, which appears to be well balanced and representative. For instance, the top-frequency lexical items in our corpus (showing a frequency above 1000) are *hotel* (3743 instances), *staff* (1345), *stay* (1141), and *location* (1128), all of them logical and to be expected given the domain and the genre dealt with in the study. Table 1 shows the 20 top-frequency units in the corpus, which also seem to be expectable, representative units in view of the topic.

Hence, the point of departure of this study was the analysis of grammatical-word-free frequency lists. It is enlightening to notice that of the 20 top-frequency lexical units (types) in the corpus (see Table 1) – and excluding grammatical words, already removed from corpus counts – seven of these types are adjectives. We have to wait until position eight to find the first highly frequent EA in the corpus, which is *clean* (625 instances) and then we subsequently find *nice* (584), *friendly* (491), *perfect* (451), *helpful* (430), *excellent* (423), and *comfortable* (409); all of these top-position adjectives can be considered evaluative, whereas the other 13 types in the corpus are nouns. From these first, preliminary corpus data it seems logical to believe that evaluative adjectives play a determining role in the genre under study (approximately 35 percent of the 20 top-frequency words in the corpus are EAs).

Table 1. Frequency list from the corpus of study

N	Word	Freq.
1	HOTEL	3743
2	STAFF	1345
3	STAY	1141
4	LOCATION	1128
5	SERVICE	922
6	STAYED	800
7	BREAKFAST	700
8	CLEAN	625
9	NICE	584
10	NIGHT	579
11	TIME	560
12	FRIENDLY	491
13	ROME	455
14	PERFECT	451
15	PARIS	437
16	HELPFUL	430
17	EXCELLENT	423
18	LONDON	414
19	COMFORTABLE	409
20	BED	394

Description: As shown in Table 1, 20 top-frequency lexical units (types) in the corpus obtained with Wordlist.

It is undeniable that frequency figures are very useful in corpus-based research and that they are normally interpreted in terms of the relevance/importance of the word in the field. Nevertheless, as conceded by Taljard and De Schryver (2002), reading through top-frequency words is obviously an unrefined procedure not necessarily implying significance. This is why, in this paper, the analysis of the image of top-ranked hotels provided by reviewers through EAs will initially be based on the classification of such EAs through a double perspective, based on the frequency but also on the saliency of these EAs. Together with the purely frequency-based analysis, with the saliency analysis we determine which significant terms, as regards both their frequency and their specialised, infrequent character if compared with a general reference corpus, are characteristic of OCRs and thus significantly help to prototypically form the image of top-ranked hotels. Hence, in this section we present a first set of results in the form of a classificatory table (Table 2) in which EAs have been categorised according to three different but complementary perspectives:

- Their aforementioned frequent and/or salient character
- Their quality and degree of positivity/negativity (according to a continuum)
- The classification of their value according to functional, semantic and pragmatic criteria

However, prior to the initial classification, that is, during the manual selection of EAs in both the frequency and the KeyWord lists, a series of problematic issues arose. As is the case in almost any study in which the author's intuition and subjectivity play an important part, there were cases in which the "evaluative" feature depended entirely on the context of use, and there were also dubious instances as regards the accuracy of the label "evaluative" in some adjective types. For instance, in our corpus, the adjective *accessible* was found with an evaluative and non-evaluative dimension (see examples 7 and 8) both in the top-frequency and keywords list:

(7) EV: *It is a small hotel so the people there are very personable and accessible.*

(8) NON-EV: *A control panel for the entire room is accessible at various spots.*

In these cases in which both evaluative and merely descriptive shades of meaning are given in a single adjective, the decision as to whether or not to consider the adjective as an EA depended on its general, predominant or more widespread use in the corpus. Consequently, EAs such as *accessible* were not included because they were considered mainly non-evaluative in our corpus, whereas some other dubious instances such as *private* were finally considered evaluative according to corpus numeric and co-textual evidence.

In the same way, since this study was limited from the very beginning to EAs in their strictest sense, nouns with a semantic adjectival status were not considered for the count. The same happened with the category of participles. Despite Quirk and Greenbaum's (1979) and Bosque's (1990) literature in this respect (according to which the difference between adjectives and participles is often not clear-cut and depends on the verbal force retained by the latter), we decided not to include them in our final count since, after a deep in-context analysis, we observed a very weak evaluative force in most of these past participles.

As previously stated, we initially classified the EAs in the corpus in three different categories (lists): a list of both top-frequency and top-saliency EAs in the corpus (list 1), a list of "merely" frequent EAs (list 2), and a list of "merely" salient EAs (list 3). Thus, in this paper, list (1) (that of both frequent and salient EAs) was created by matching the EAs appearing in both the WordList and the KeyWords lists. Scott (1998, p. 236) defined the term "keyword" as "a word which occurs with unusual frequency in a given text". As Taljard and De Schryver (2002) state, unusual frequency is normally related to outstandingness and implies that a word has an unusually high (but also an unusually low) frequency in a text/corpus in comparison to its occurrence in a *reference corpus* of some kind since a keyword list therefore gives a measure of *saliency*, whereas a simple word list only provides *frequency* (Baker 2006). Hence, KeyWords can thus be said to complement the kind of analysis offered by WordList frequency lists.

As Berber Sardinha (2000) goes on to state, the software also identifies words whose frequencies are statistically lower in the study corpus, which are called "negative keywords", in contrast to positive keywords, which have higher frequencies in the study corpus. However, negative keywords are not the issue to be discussed in this paper and, whenever a keyword is mentioned, a "positive" value is implied. We also disregarded those words rendered as keywords by the program simply because their incorrect spelling makes them salient with respect to the reference corpus. This is the case of the "false keywords" *comfortable**, *helpfull**, *pleasent**, *wonderfull**, *seperate**, *beautifull**, *excellant**, *freindly**, *wonderfull**, *phantastic** and *awfull**. We attribute the recurrent appearance of these words in the corpus mainly to a simple lack of attention on the part of the reviewers. In the case of this study, the reference corpus was the British National Corpus (a reference corpus of 100 million words of written and spoken general British English).

As Table 2 shows, after comparing the top-frequency EAs from the corpus of study with the positively key/salient EAs in the corpus (175 EAs in total were rendered as keywords (Note 3), some of which did not appear among Wordlist's top 2000 lexical units), a significant number of coincidences were noted. For example, 115 EAs appeared in both frequency and saliency lists (units belonging to list (1)), revealing the most representative EAs in the corpus. This high number of coincidences is, however, to be expected since keyness is also assigned on the basis of frequency in comparison with the general, reference corpus.

Departing from the classification of EAs according to their level of frequency and/or saliency, we proceeded with the analysis of such EAs from the perspective of their degree of positivity/negativity. Reviewers often evaluate products or aspects with multiple attributes that undoubtedly contribute to the informative nature of OCRs and thus to image formation. However, these attributes which will act as opinion formers/providers for the potential customer, trigger a rather unconscious initial cognitive process in him/her consisting in simply regarding the information as positive or negative. Subjective expectations also play an important role in this initial image-formation stage. In this specific case, in which top-ranked hotels are the focus of the study, these expectations (as well as EAs themselves) will presumably tend toward positivity, whereas negative information will have, in our opinion, an even greater effect on the reader because of its more unexpected nature. This is the onset of image or impression formation: the consideration of an evaluation as being positive or negative. EAs may be positive, neutral or negative in valence and may trigger, in this case, the action of booking or disregarding a hotel on the part of a prospective client; they can thus be considered "image formers/providers" and consequently "action-triggers". The placing of EAs on the positivity-negativity continuum established in Table 2 was the second stage in the proposed classification. The classificatory process also attempted to reproduce the natural process by which EAs are understood, assimilated and interpreted by readers when forming an image or impression of the hotel reviewed. This is why, in the final stage, the image formed in the readers' mind and thus the action of liking or disliking, of booking or not, will be carried out depending on the more reflective specific evaluation provided for each of the aspects under analysis and on the accomplishment of expectations on the part of the reader for each of these aspects. Accordingly, Table 2 shows the results obtained from the corpus of study categorised in terms of a triple perspective: the placing of EAs in list (1): frequent and salient EAs, list (2): only frequent EAs, or list (3): only key/salient EAs; their placing on a quality continuum (ranging from the ultra-positive, positive, and more neutral or average, to the negative, or ultra-negative); and their placing according to the kind of evaluation provided as regards five basic values or categories (Note 4) for EAs established on functional, semantic and pragmatic criteria: aesthetic appeal, general appraisal, deviance (Note 5), emotional/sensory appeal (not aesthetic), and size/strength-related adjectives.

Table 2. Suggested categorisation of evaluative adjectives

	++	+	+/-	-	--	
	Majority extremely positive evaluation	Majority positive evaluation	Neutral subject to interpretation evaluation	Majority negative evaluation	Majority extremely negative evaluation	
Freq. and salient EAs	impeccable (45) immaculate (23) magnificent (21) sumptuous (8) phenomenal (9) spectacular (20)	beautiful (237) elegant (64) pretty (95) stylish (26) trendy (19) classy (14) modern (116)				Aesthetic appeal
LIST (1)	perfect (451) excellent (423) amazing (251) fantastic (221) wonderful (301) superb (93) awesome (40)	convenient (111) fine (91) flawless (16) faultless (15) clean (625)	worth (125) touristic (9) ready (77) particular (29) basic (118)	bad (100) (not) worth (6) (not) (as) wonderful (5) (not) ideal (1)	terrible (35)	General appraisal

	gorgeous (47) memorable (32) nicest (19) brilliant (46) exceptional (39) terrific (17) ideal (43) unbeatable (8)	spotless (25) efficient (57) tidy (16) professional (61)				
	outstanding (61)	cheap (150) pricey (24) special (154) cheapest (12) affordable (16) exemplary (10)	nearest (23) quick (53) loud (30)	expensive (124) dirty (52) busy (69)		Deviance
	impressive (26) delicious (67) exquisite (20) delightful (19) incredible (50)	happy (107) impressed (37) enjoyable (24) unforgettable (14) romantic (33) reasonable (56) gracious (17) fun (49) cheerful (116) quiet (260) friendly (491) helpful (430) nice (584) comfortable (409) lovely (322) attentive (61) pleasant (90) favorite (31) welcoming (61) luxurious (45) charming (71) comfy (38) polite (65) cozy (29) warm (108) relaxing (46) intimate (21) tasteful (8) safe (90) fresh (82) cosy (24) cool (47) knowledgeable (25) nicer (17) handy (23)	quieter (9) walkable (10)	noisy (68) unfriendly (19) rude (31) uncomfortable (20) intrusive (10) negative (37) disappointing (24) picky (8) stuffy (16) (not) impressive (1)		Emotional/ sensory appeal (not aesthetic)

		spacious (149)	huge (87) short (119) little (247)	smallish (8)		Size-strength related
Only freq. EAs LIST (2)		smart (13) chic (23)				Aesthetic appeal
	fabulous (109) finest (16) divine (11)	adequate (36) central (156) main (108) unique (41) bright (25) original (21) suitable (9) genuine (22) useful (17)	real (67) major (56) normal (13) simple (41) common (34) usual (27) correct (16) acceptable (15) typical (26) regular (23) true (55) personal (46) numerous (12) ordinary (10)	awkward (7) weird (8) odd (12) strange (17) rare (16) wrong (Note 6) (24) low (36) missing (28)	awful (19) horrible (13) massive (10)	General appraisal
		historic (14) immediate (15) traditional (14) cheaper (18) famous (18) easy (135) positive (37)	fast (19) frequent (13) popular (12) unusual (8) brief (9) classic (18) ancient (11)	impossible (9) worse (19) difficult (37) poor (52)		Deviance
		attractive (11) generous (11) reliable (11) careful (10) sophisticated (8) thoughtful (8) peaceful (16) secure (12) functional (12) magical (12) soft (32) amazed (11) informative (10) natural (10) powerful (10) refreshing (9) delighted (13) funny (13) familiar (12) excited (16)	unbelievable (9)	worried (19) unpleasant (10) nervous (9) unsafe (9) dangerous (8) disgusting (8) sick (8) annoyed (7) dark (40) hard (70) (not) favourite (3)		Emotional/sensory appeal (not aesthetic)

		honest (16) favourite (12) lucky (24) willing (28) glad (23) aware (18) smooth(17)				
	biggest (13) enormous (10)	larger (39) bigger (16) strong (17) wide (12)	tight (16)	narrow (18)		Size-strength related
Only salient EAs			retro (6)	quirky (12) artsy (5)		Aesthetic appeal
LIST (3)	stunning (41) marvelous (7) fab (11) excellent (423) coolest (6)	OK (109)	touristy (7) near (154) hot (102) private (98) sparkling (13)	overpriced (21) priced (26)	worst (38)	General appraisal
			decent (28)			Deviance
		quintessential (5) restful (6) tasty (13) yummy (9) airy (13) personable (7) favorable (5) approachable (6)	surprised (41) bustling (12)	annoying (13) skeptical (4)		Emotional/ sensory appeal (not aesthetic)
		roomy (10)	oversized (6) tiny (75)	claustrophobic (9) cramped (18)		Size-strength related

Description: As shown in Table 2, suggested categorisation of the evaluative adjectives in the corpus (with frequency of appearance included) from a triple perspective.

Since meaning in context is critical to any kind of discourse/terminological analysis, an immediate-context analysis of the corpus was performed to examine the placing of most of the instances in Table 2. Therefore, apart from their meaning *per se*, EAs were placed in Table 2 according to their most recurrent usage in corpus. If the EA analysed presented approximately the same number of overall positive and negative in-context interpretations or an adjective presented a weak evaluative force or was comparatively neutral or unmarked in nature, they were placed in the “neutral/subject to interpretation” column, despite acknowledging the fact that EAs are never totally neutral, otherwise they would lose their evaluative character.

From the placing and triple perspective analysis of the EAs in the corpus, the first big conclusion to be drawn is that, as expected, negative evaluations are outnumbered by positive or ultra-positive ones in top-ranked OCRs. On the whole, it may be stated that the evaluation categories “general appraisal” and “emotional/sensory appeal”

were the ones preferred by reviewers for shaping, depicting, and helping to form the image of the hotel reviewed. These categories are also the most common ones among the most representative EAs in the corpus (those belonging to list 1) and show a clear tendency toward extremes in evaluation. Deviance adjectives are also highly significant in the genre under analysis. The very same rationale of the genre allows for this assumed importance because deviance adjectives show precisely a deviance from an established or expected norm, and the accomplishment (or not) of expectations is one of the key aspects valued by both readers and reviewers. For instance, the qualification of a hotel as *cheap* indicates that it cost less than expected, the qualification of *dirty* indicates that it is less clean than expected/the accepted norm, and this kind of information is always critical because it is at the basis of compliance with basic standards and expectations.

As previously stated, there appears to be a clear tendency on the part of reviewers toward positivity or at least toward the use of positive EAs in their reviews of top-ranked hotels and this is especially noticeable in the “frequent and salient” category (list (1)). This “frequent and salient” category presents what can be considered the most representative adjectives of the domain and genre under study and in it we can appreciate a clear tendency toward positivity and ultra-positivity (hyperbole). It is interesting to notice that EAs from list (1) are normally instances of positive or extra-positive evaluations so that we can claim there is a deliberate tendency toward positive exaggeration in the corpus but originated in a non-interested position, which makes it more valuable and reliable. It is also interesting to note that “general appraisal” EAs such as *perfect*, *excellent*, *amazing*, *fantastic*, *wonderful*, or *incredible*, placed mainly at the “extremely positive evaluation” end of the continuum, are also the ones normally used in the titles of top-ranked hotel reviews. This can be explained by the fact that they show a general and very expressive overall evaluation, as shown by examples 9 and 10. The same happens with the few “general appraisal” EAs from list (2), which are also the ones normally used in the titles of OCRs for very negative appraisals. These EAs are mainly frequent but not salient, as the ones in list (1), except *worst* (see examples 11 and 12):

(9) *Wonderful! I would never stay anywhere else!*

(10) *Fabulous Hotel and Great Staff.*

(11) *The most horrible hotel I have ever stayed in.*

(12) *The worst Hotel ever; what a nightmare!*

It is also very common throughout the three lists to find numerous “emotional/sensory appeal” appraisals, normally at the “majority positive evaluation” end of the continuum: *friendly*, *helpful*, *relaxing*, *enjoyable*, and *romantic* are some of the most common ones. These kinds of adjectives are seen to be very important in our corpus, probably due to the fact that we are dealing with a highly subjective, opinionated genre which is constantly pervaded by the reviewers’ own emotions of fulfillment or, to a lesser extent in our corpus, disappointment.

The EAs in the category of “only-frequent” (list (2)) types were much more difficult to place in the table because in general they present more neutral semantic traits than other more key/salient adjectives (more significant in the genre). They are, so to speak, commoner or less marked adjectives and as such their meaning is not so strongly marked in the context of study as in the case of some extreme adjectives that are both salient and frequent, and thus undoubtedly characteristic of the genre under analysis. The adjectives in this category are essentially positive or neutral in nature and tend to provide “general appraisal” and “emotional/sensory appraisal”. Proportionally, “only-frequent” adjectives present more negative types than the adjectives in category (1). These adjectives pertaining to the negative part of the continuum are thus highly frequent (*awkward*, *awful*, *horrible*, *unpleasant*), but WST does not consider them as salient – and thus characteristic of hotel OCRs- as others such as *terrible*, *rude*, or *disappointing* in our corpus.

Appraisals regarding “size-strength related” negative qualities are proportionally common in EAs from list (3). That is, “negative size-strength related” EAs are especially common among the only salient/key set of adjectives – *claustrophobic*, *cramped*, *minuscule* – whereas for lists (1) and (2), these kinds of EAs are mainly positive or ultra-positive – *spacious*, *biggest*, *enormous*, and so forth. Despite their proportionally minority character, negative evaluations are critical in the context of top-ranked hotels. Many studies in psychology, political science, and marketing suggest that people often tend to grant negative information greater weight than positive information of an equal intensity (Ahluwalia 2002; Herr, Kardes & Kim 1991; among others). As Herr et al (1991) stated, negative attribute information is normally assigned greater weight than positive attribute information because it contrasts with expectations (Simpson & Ostrom 1976), something that is undeniable in the context analysed here. Accordingly, as Lalwani (2006) concedes, these and other similar findings have led researchers to conclude that negative information exerts a greater influence on judgments than equivalent

positive information because of its higher diagnosticity. This phenomenon has been termed the *negativity bias*. Hence, negative information normally exerts a greater influence than positive information because it is both *novel* and *unexpected*. Good appraisals are what readers expect when reading about top-ranked hotels, but it is negative information that more powerfully determines the final image or impression in the reader's subconscious. Moreover, negative information has been claimed to carry greater specific weight because it is normally less ambiguous than positive information (Wyer 1973). The analysis of negative evaluations in our corpus of study is important not just because this kind of evaluation normally has a greater impact on the reader's subconscious mind, but because they are more unexpected in a corpus of top-ranked hotel reviews. Negative appraisals in our corpus are especially significant as regards emotional/sensory appeal EAs (*noisy, unfriendly, rude, disappointing, picky, and stuffy*) and they mainly refer to staff work. Even "rooms" seems another of the aspects where negative appraisals appear more often, as the next section will show. Thus, we dare to state that due to the higher emotional impact of negative statements and due to the consciously or unconsciously assumed greater diagnostic character of negative evaluations, reviewers tend to amplify (exaggerate) positive evaluations through the use of extreme adjectives. The reason why they do this is to make their contributions as expressive as possible and to try to offset the effect of negativity, which is more easily achieved. The way these specific aspects are evaluated in the corpus is explained in the next section.

3.2 Collocational Behavior of the Main Evaluated Entities and Their Associated Evaluative Adjectives in Online Customer Reviews

In this section, the main aspects (evaluated entities) normally dealt with in hotel OCRs (location, cleanliness, staff, etc.) are analysed in terms of their linguistic combination and appearance with EAs. Their joint appearance in context has been analysed in order to determine their collocational behavior and thus the way EAs portray or convey the image of specific hospitality aspects dealt with in OCRs of top-ranked hotels.

As explained in previous sections (see Figure 2), the OCRs analysed deal mainly with the evaluated entities "value", "rooms", "location", "cleanliness", "service", and "sleep quality", which summarise the basic aspects to be considered and appraised in any hotel stay. On the basis of this list of assessable items proposed by Tripadvisor, we have created a simplified five-item list of evaluated entities regarding the concept "hotel": 1. Value, 2. Rooms, 3. Location, 4. Cleanliness, and 5. Service. The purpose of this part of the study was to analyse how the image of each of these key evaluated entities is formed according to both the EAs most frequently associated with them and the way these related EAs collocate. With this aim in mind, a maximum of four key/allusive words per evaluated entity were chosen and analysed in the corpus to study the way EAs and such allusive words create recurrent and significant linguistic patterns and thus contribute to the image-formation process. Except for the aspects "value" and "cleanliness", in which some adjectives were used as key/allusive words, the rest of those words were nouns, and all of them were analysed with the Concord tools "Concordance" and "Collocates". Accordingly, the list of evaluated entities and the associated key/allusive words chosen is as follows:

1. Value: *value, money, price*, worth** (Note 7)
2. Rooms: *room*, suite*, bed*, bathroom**
3. Location: *locat**
4. Cleanliness: *clean**
5. Service: *service, staff*

Additionally, the search word (key/allusive word) *hotel** was also submitted to WST analysis in order to study the most recurrent and significant collocates for this sort of hypernym in the corpus.

Obviously the selection of words was rather personal and subjective, and we are aware of the fact that there are many other possible (and maybe better) key/allusive words that could have been used for the analysis. Nonetheless, we considered the ones chosen a good first approach to the topic and well suited to be able to reach representative conclusions. Possible related key/allusive words of a too-generic nature or that allow for more than one interpretation (such as *place* for the category "location") were discarded to avoid imprecision and the same happened with words with a low frequency of appearance in the corpus (such as *tidy*). Obviously, the aspects analysed may also (and in fact very often do) appear mixed up or commented in an integrated way: rooms may be clean (mixing up the "rooms" and "cleanliness" categories), location may not be worth the money (mixing up "location" and "value"), all of which contributes to the richness of the study. From the analysis of these results, a series of general conclusions may initially be drawn:

Once again, the lists of collocates obtained reflected a clear and expected overall tendency toward the positive or

ultra-positive evaluation of specific hospitality aspects. In general (and proportionally), the most positively evaluated aspect is *location*, so location is, or can at least be expected to be, one of the most important and top-valued aspects among top-ranked hotels.

One of the aspects allowing for more negative evaluations is the evaluated entity “Room”, and more specifically the key/allusive aspects *bed**, *room**, and *bathroom**. The negative collocates of such words are normally “size/strength-related” EAs referring to the inappropriate or disappointing size of the room/bathroom (*small*, *smaller*, *little*, *tiny*), together with general appraisal EAs (*terrible*) or emotional/sensory appeal adjectives such as *uncomfortable* or *hard*, mainly when referring to beds. The second aspect allowing for more negative evaluations (also proportionally) is “Service”, whose key/allusive words tend to collocate in the corpus with *poor*, *bad*, *unfriendly*, and *awesome*. Hence, despite the proportionally small number of negative appraisals, we can conclude that according to our corpus analysis, the evaluated entity “Room*” is the one that is most often negatively evaluated in top-ranked hotel OCRs, followed by that of “Service”.

The general evaluated entity “Hotel” is negatively assessed in the corpus by means of EAs such as *terrible*, *expensive*, or *small* despite the evident and already-mentioned tendency toward positive or ultra-positive appraisals of a rather generic nature. Hyperbole is especially noticeable in this evaluated entity, in which “general appraisal” extreme EAs such as *wonderful*, *fantastic*, *fabulous*, *gorgeous*, and so on are frequent. “Emotional/sensory appeal” adjectives are also especially common in the appraisal of hotels in general, examples being *romantic*, *friendly*, *favorite*, and *welcoming*. It is interesting to note that the aspects that most often contain negative appraisals in our corpus are, curiously, also the ones that are most often positively appraised with extremely positive EAs. For instance, apart from the rather widely extended complaints about size of rooms and bathrooms, the aspect “Room*” frequently shows key collocations referring to the *excellent*, *fantastic*, *amazing*, *wonderful*, *outstanding*, or *superb* nature of its associated key/allusive words.

Finally, “Concordance” and “Collocates” were used to retrieve some of the most common or simply significant collocations used by reviewers when reviewing top-ranked hotels for each of the evaluated entities under study. By so doing, we believe a complete picture of the image-formation mechanisms provided by OCRs through EAs can be obtained. Obviously, just a brief illustrative sample of collocations has been included because of space limitations (Note 8). For instance, we find many collocations carrying a super-positive evaluation of the different aspects:

VALUE: *Great/very good/excellent/fantastic VALUE // Excellent/great/superb/best/amazing VALUE for (the) MONEY // WORTH every penny/dollar/euro //absolutely WORTH it // well/totally WORTH it // WORTH a try // expensive but WORTH it // WORTH the PRICE for (city) // WORTH the MONEY // WORTH paying/trying //...*

ROOMS: *Excellent/fantastic/wonderful ROOM(S) // ROOM(S) was/were amazing // stunning/fabulous/extremely comfortable SUITE // the SUITE is beyond belief // cloud-like BED // the most heavenly BEDDING // the BEDROOM is magical // magic BED // the BED was supremely comfortable // the BEDROOM is exceptional // the BEDROOMS were impressive // huge /wonderful BATHROOM // the BATHROOM was simply amazing// BATHROOMS are fabulous// immaculate CLEANLINESS //...*

LOCATION: *excellently LOCATED // LOCATION is brilliant/ideal/great /fabulous // the best/magnificent/amazing/fantastic/terrific/prime/unbelievable/ perfect LOCATION // LOCATION couldn't be better//...*

CLEANLINESS: *impeccably CLEAN // immaculately/exceptionally/extremely CLEAN //...*

SERVICE: *excellent/unforgettable/wonderful STAFF// STAFF were(was) very helpful/(extremely) helpful and eager/amazing/wonderful/incredibly friendly and helpful// excellent and flawless STAFF // SERVICE is impeccable // SERVICE was excellent // perfect/wonderful/outstanding/brilliant/superb/great/first-class SERVICE // SERVICE was good/top-notch//...*

Among those providing a more average or even neutral evaluation, we find:

VALUE: *OK/good/fair for the PRICE // Good VALUE for (the) MONEY // reasonably PRICED for (CITY) // reasonable PRICE // well PRICED // modest/fair/standard/normal/cheap PRICE / VALUE for the PRICE // comfortable / lovely BED(S) // the BED was clean // BED(S) were cosy//...*

ROOMS: *Good/great/spacious/nice/comfortable/lovely/large/beautiful ROOM(S) // adequate ROOM(S) // lovely SUITE(S) // the SUITE(S) are elegant // comfortable / lovely BED // the BED was clean // BEDS were cosy // the BEDROOMS are clean and tidy // the BATHROOM was clean // clean/well-cleaned/lovely BATHROOM // the BATHROOM was roomy // BATHROOM is okay//...*

LOCATION: *good/quiet/convenient/strategic LOCATION // LOCATION was okay // ...*

CLEANLINESS: *CLEAN and tidy room(s)/hotel// CLEAN room(s)/hotel/ bathroom/toilet // CLEAN and well-maintained hotel // CLEAN and warm room // rooms are reasonably CLEAN // CLEAN towels // ...*

SERVICE: *well-trained/professional/attentive/charming/friendly STAFF// the staff was lovely // attentive/welcoming/friendly/good/efficient/professional SERVICE //...*

Finally, although in a much smaller proportion, Concordance also shows recurrent negative or super-negative evaluations:

VALUE: *over-PRICED hotel // extremely poor/very poor VALUE for money // (not) worth the PRICE // (not) worth the PRICE for (city) //...*

ROOMS: *tiny ROOM(S) // ROOM(S) was/were dark // bad ROOM(S) // expensive ROOM(S) // the BED was awful // the worst BED // the BED was the hardest // the most uncomfortable BED // the worst BEDS ever//...*

LOCATION: *terrible LOCATION // LOCATION is horrible // LOCATION is not great //...*

CLEANLINESS: *horrible cleaning STAFF //...*

SERVICE: *horrible cleaning STAFF // STAFF were not very friendly// STAFF is really rude // miserable STAFF // poor (customer) SERVICE //...*

4. Conclusion

It is commonplace these days to acknowledge the enormous potential the Internet has as regards its communicative capabilities. The Internet aids active linguistic exchanges, increased interactivity, and language revitalisation by enabling all sorts of speakers all over the world to share information and maintain a virtual contact that is more or less synchronic depending on the digital genre chosen. As Crystal (2005) stated, we are at the beginning of an era of linguistically based online applications, in which OCRs are only an example. In the specific case of top-ranked hotel OCRs, a clear and expected tendency toward reporting positive aspects and thus using positive evaluations was observed despite not being an interested genre such as promotional brochures or webpages. These positive or super-positive evaluations are especially significant in both frequent and salient EAs from the corpus of study, but this very same purposeful bias of the corpus toward positive aspects simultaneously increases the informative character of negative aspects, which take on a new decisive and inherent force due to their unexpectedness and diagnostic nature. EAs have proven to be critical in the interpretation of subjective expected values, and they show great potential in destination image formation, to which the possibilities of the WorldWideWeb also contribute enormously.

It is rather common to detect a mismatch between what hotels say to offer, the way they present it and what customers perceive when in the hotel. This paper has thus aimed to show the way EAs are used by real online reviewers and the effect such evaluations may have on the image-formation process of top-ranked hotels on the part of the readers. Hence, from a very general perspective, the study has shown, among many other aspects: (1) that there exists a considerable lexical richness in the genre analysed as regards the EAs utilised; (2) that ultra-positive evaluations are highly common (there is a tendency toward the use of positive extremes in appraisals), at times displaying a rather exaggerated or hyperbolic use of evaluations; (3) that negative appraisals carry a specific added force in this context; and (4) that the EAs employed in the review titles tend to use general appraisal EAs to summarise the kind of evaluation developed later in the review itself. Moreover, the distribution of EAs was analysed on a positivity continuum, but also according to frequency and/or saliency criteria, and on the basis of the functional/semantic and pragmatic value of the adjectives. Finally, a linguistic analysis of the way the main aspects appraised and their corresponding EAs collocate in discourse has contributed to the understanding of how EAs may act as opinion-formers/providers and thus in image formation processes. We thus

believe that this kind of analysis sheds some light on the linguistic evaluative strategies used by reviewers through EAs and on the influence they may have on readers and on their image of the hotel reviewed as a whole. Accordingly, a proper analysis of image formation processes as carried out in tourist-to-tourist online interaction could be useful to foster the potential of hotel promotional discourse, for instance through webpages, which could be improved to attain increased reliability. This could be translated into a greater and better effect on prospective customers by enhancing the psychological and symbolic values that tourists – both reviewers (former clients) and review readers (potential customers) value and appreciate when choosing a hotel.

References

- Ahluwalia, R. (2002). How Prevalent Is the Negativity Effect in Consumer Environments? *Journal of Consumer Research*, 29(2), 270-279. <http://dx.doi.org/10.1086/341576>
- Baker, P. (2006). "The question is, how cruel is it?" Keywords, Foxhunting and the House of Commons. In AHRC ICT Methods Network Expert Seminar on Linguistics. Lancaster University, UK.
- Baloglu, S., & McCleary, K. (1999). A model of destination image formation. *Annals of Tourism Research*, 26(4), 868-897. [http://dx.doi.org/10.1016/S0160-7383\(99\)00030-4](http://dx.doi.org/10.1016/S0160-7383(99)00030-4)
- Barton, E. L. (1993). Evidentials, argumentation, and epistemological stance. *College English*, 55(7), 745-771. <http://dx.doi.org/10.2307/378428>
- Berber Sardinha, T. (2000). Comparing corpora with WordSmith Tools: How large must the reference corpus be? In A. Kilgarriff, & T. Berber Sardinha (Eds.), *Proceedings of The Workshop on Comparing Corpora* (pp. 7-13). <http://dx.doi.org/10.3115/1604683.1604687>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Bosque Muñoz, I. (1990). *Las categorías gramaticales*. Madrid: Editorial Sintesis SA.
- Cai, L. (2002). Cooperative branding for rural destinations. *Annals of Tourism Research*, 29(3), 720-742. [http://dx.doi.org/10.1016/S0160-7383\(01\)00080-9](http://dx.doi.org/10.1016/S0160-7383(01)00080-9)
- Calvi, M. V. (2010). Los generos discursivos en la lengua del turismo: una propuesta de clasificación. *Ibérica*, 19(9-32).
- Conrad, S., & Biber, D. (2000). Adverbial Marking of Stance in Speech and Writing. In S. Hunston, & G. Thompson (Eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse* (pp. 56-73). Oxford: OUP.
- Crystal, D. (2002). *El lenguaje e Internet*. Madrid: Cambridge University Press.
- Crystal, D. (2005). The Scope of Internet Linguistics. Paper delivered online to the American Association for the Advancement of Science Meeting. Retrieved from http://www.davidcrystal.com/DC_articles/Internet2.pdf
- Echtner, C. M., Ritchie, J., & Brent, R. (2003). The meaning and measurement of destination image. *The Journal of Tourism Studies*, 14(1), 37-48.
- Edo Marzá, N. (2012). Páginas web privadas e institucionales: El uso de la adjetivación en un corpus inglés-español de promoción de destinos turísticos. In J. Sanmartín Sáez (Ed.), *Discurso turístico e Internet: normas y usos*. Editorial Veruvert.
- Fellbaum, C., Gross, D., & Miller, K. (1993). Adjectives in WordNet. In G. Miller et al., *Five Papers on WordNet*. CSL Report 43.
- Fish, S. (1980). *Is There a Text in This Class?* Cambridge, MA: Harvard University Press.
- Gartner, W. C. (1993). Image formation process. *Journal of Travel & Tourism Marketing*, 2, 191-215. http://dx.doi.org/10.1300/J073v02n02_12
- Garzone, G. (2009). Identity in Tourist Communication on the Internet: Italy's and Spain's Websites Contrasted. In G. Garzone, & P. Catenacio (Eds.), *Identities across Media & Modes: Discourse Perspectives*. Peter Lang: Bern.
- Gea Valor, M. L. (2008). Epistemic and Deontic Modality in Online Customers' Reviews. In R. Monroy, & A. Sánchez (Eds.), *25 Años de Lingüística Aplicada en España: Hitos y Retos / 25 Years of Applied Linguistics in Spain: Milestones and Challenges* (pp. 217-223). Murcia: Servicio de Publicaciones de la Universidad de Murcia.

- Gylling, L. (2004). An analysis of a destination's image and the language of tourism (Doctoral dissertation). The Aarhus School of Business. Retrieved from <http://pure.au.dk/portal-asb-student/files/2156/000134297-134297.pdf>
- Hall, C. M. (Ed.). (2003). *Introduction to Tourism: Dimensions and Issues*. Australia: Pearson Education. <http://dx.doi.org/10.1080/13683500308667952>
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: Arnold.
- Herr, P. M., Kardes, F., & Kim, J. (1991). Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective. *Journal of Consumer Research*, 17(4), 454-462. <http://dx.doi.org/10.1086/208570>
- Hewings, M. (2004). An 'important contribution' or 'tiresome reading'? A study of evaluation in peer reviews of journal article submissions. *Journal of Applied Linguistics*, 1(3), 247-274. <http://dx.doi.org/10.1558/japl.2004.1.3.247>
- Hunston, S., & Gill, F. S. (2000). *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: Benjamins.
- Hunston, S., & Thompson, G. (2000). *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- Kerbrat-Orecchioni, C. (1980). *L'enonciation. De la subjectivité dans le langage*. Paris: Armand Colin.
- Lalwani, A. (2006). Negativity and Positivity Biases in Product Evaluations: The Impact of Consumer Goals and Prior Attitudes (Doctoral dissertation). ZIBS Technical Report. Urbana, Illinois. Retrieved from <http://www.zibs.com/techreports/Negativity%20and%20Positivity%20Biases%20in%20Product%20Evaluations.pdf>
- Martin, J. R. (2000). Beyond exchange: APPRAISAL systems in English. In S. Hunston, & G. Thompson (Eds.), *Evaluation in Text*. Oxford: Oxford University Press.
- Martin, J. R., & White, P. R. R. (2005). *The Language of Evaluation, Appraisal in English*. London and New York: Palgrave Macmillan.
- McCarthy, M., & Carter, R. (2004). "There's millions of them": Hyperbole in everyday conversation. *Journal of Pragmatics*, 36(2), 149-184. [http://dx.doi.org/10.1016/S0378-2166\(03\)00116-4](http://dx.doi.org/10.1016/S0378-2166(03)00116-4)
- Morgan, N., & Pritchard, A. (1998). *Tourism Promotion and Power – Creating Images, Creating Identities*. Wiley.
- Norricks, N. R. (2004). Discussion note: Hyperbole, extreme case formulation. *Journal of Pragmatics*, 36, 1727-1739. <http://dx.doi.org/10.1016/j.pragma.2004.06.006>
- Pharo, N. (2008). "Web Documents and Genre." Digital Documents Lecture. Presentation given in Høgskolen i Oslo, Oslo, Norway. 3 Nov. 2008.
- Posteguillo, S. (2002). Netlinguistics and English for Internet Purposes. *Ibérica*, 4, 21-38.
- Posteguillo, S. (2003). *Netlinguistics. An Analytical Framework to Study Language, Discourse and Ideology in Internet*. Castellón: Publicacions de la Universitat Jaume I.
- Quirk, R., & Greenbaum, S. (1979). *A university grammar of English* (9th impression). London: Longman Group Limited.
- Samson, C. (2006). "... Is different from...": A corpus-based study of evaluative adjectives in economics discourse. *IEEE Transactions on Professional Communication*, 49(3), 236-245. <http://dx.doi.org/10.1109/TPC.2006.880741>
- Scott, M., & Oxford University Press. (1998). *WordSmith Tools Manual*. Retrieved from <http://www.lexically.net/downloads/version5/HTML>
- Shaw, P. (2006). Evaluative language in evaluative and promotional genres. In G. D. Lungo, M. Dossena, & B. Crawford (Eds.), *Variation in Business and Economics Discourse: Diachronic and Genre Perspectives* (pp. 152-165). Rome: Officina Edizioni.
- Sidnell, J. (2004). 'There's risks in everything': Extreme-case formulations and accountability in inquiry testimony. *Discourse & Society*, 15(6), 745-766. <http://dx.doi.org/10.1177/0957926504046503>

- Simpson, D. D., & Ostrom, T. M. (1976). Contrast Effects in Impression Formation. *Journal of Personality and Social Psychology*, 34(4), 625-629. <http://dx.doi.org/10.1037/0022-3514.34.4.625>
- Stubbs, M. (1998). Language and the Mediation of Experience: Linguistic Representation and Cognitive Orientation. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 358-373). Oxford: Blackwell Publishers.
- Swales, J. M. (1990). *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M., & Burke, A. (2003). 'It's really fascinating work': Differences in evaluative adjectives across academic registers. In P. Leistyna, & C. F. Meyer (Eds.), *Corpus Analysis: Language Structure and Language Use* (pp. 1-18). Amsterdam: Rodopi.
- Taljard, E., & de Schryver, G. M. (2002). Semi-automatic Term Extraction for the African Languages, with Special Reference to Northern Sotho. *Lexikos*, 12, 44-74.
- Tripadvisor.com: <http://www.tripadvisor.es/>
- Tucker, G. H. (1997). A functional lexicogrammar of adjectives. *Functions of Language*, 4(2), 215-250. <http://dx.doi.org/10.1075/foL4.2.04tus>
- Wyer, R. S. (1973). Category Ratings as "Subjective Expected Values": Implications for Attitude Formation and Change. *Psychological Review*, 80(6), 446-467. <http://dx.doi.org/10.1037/h0035455>
- Yus, F. (2001). *Ciberpragmática. El uso del lenguaje en Internet*. Barcelona: Ariel.

Notes

Note 1. In this case, the OCRs selected have been circumscribed to reviewers from English-speaking countries in an attempt to guarantee as much as possible the reliability of the language under analysis.

Note 2. "Types" refers to the number of different words appearing in the corpus or in a part of it, whereas "Tokens" refers to the total number of words in the corpus or in a part of it.

Note 3. WST results will always depend on the settings established in the program for the analysis.

Note 4. We are aware of the fact that the categories themselves and the resulting categorisation of EAs are the result of rather subjective processes in which certain reservations may always exist. In fact, many adjectives could be interpreted in different ways to the one proposed in the classification.

Note 5. Deviance adjectives, according to Swales and Burke (2003), denote how closely related something is to what one would expect it to be. Hence, deviance indicates how markedly different something is from an accepted norm.

Note 6. These adjectives (as happens with many other "slippery" adjectives from the table that could be interpreted from different perspectives) could also be placed in the "deviance" category. In these cases in which more than one interpretation was possible, the most frequent usage or interpretation according to corpus analysis was chosen.

Note 7. The asterisk reproduces the way key/allusive words were introduced into the WST program. Such syntax allows the program to retrieve not only "price" but also all the instances containing the letters preceding the asterisk, that is, "priced", "pricey", "prices", etc. The same is indicated by the rest of the asterisks for the rest of the key/allusive words.

Note 8. Collocations were chosen on the basis of the most frequent position of the collocates with respect to the search word, this meaning that if *spacious room* is included, the *room was spacious* will also probably appear in the corpus, but the former will be chosen because of its more frequent character.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).