

## Repositório ISCTE-IUL

---

**Deposited in *Repositório ISCTE-IUL*:**

2019-06-06

**Deposited version:**

Post-print

**Peer-review status of attached file:**

Peer-reviewed

**Citation for published item:**

Antonio, N., de Almeida, A., Nunes, L., Batista, F. & Ribeiro, R. (2018). Hotel online reviews: different languages, different opinions. *Information Technology and Tourism*. 18 (1-4), 157-185

**Further information on publisher's website:**

[10.1007/s40558-018-0107-x](https://doi.org/10.1007/s40558-018-0107-x)

**Publisher's copyright statement:**

This is the peer reviewed version of the following article: Antonio, N., de Almeida, A., Nunes, L., Batista, F. & Ribeiro, R. (2018). Hotel online reviews: different languages, different opinions. *Information Technology and Tourism*. 18 (1-4), 157-185, which has been published in final form at <https://dx.doi.org/10.1007/s40558-018-0107-x>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

---

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

---

# Hotel Online Reviews: Different Languages, Different Opinions

## 1 Introduction

Price, location, hotel facilities, loyalty programs, and design are some of the factors that drive a customer to choose a hotel (Anderson 2012; Cantallops and Salvi 2014). As the widespread use of the Internet has promoted the growth of user-generated content on social media platforms, online reviews have become another important driver of customers' hotel booking decisions (Anderson 2012; Duan et al. 2016; Cantallops and Salvi 2014; Ye et al. 2009). According to Duan et al. (2016), online reviews in the hospitality industry influence up to 50% of all hotel booking decisions and in turn influence hotels' business performance. Anderson (2012) estimated that a one percentage point increase in a social reputation index can lead to a 0.54% increase in hotel occupancy and a 1.42% increase in hotel revenue per available room. Öğüt and Onur Taş (2012) demonstrated that a one percentage point increase in a social reputation index can lead to an occupancy increase of up to 2.68%. Kim et al. (2015) and Torres et al. (2015) showed that as review ratings increase, customers' willingness to pay more for a room also increases.

Most studies on online reviews focus only on their quantitative summaries to represent user opinions (Duan et al. 2016). However, more recent studies are also focusing on the textual component of reviews (Bjørkelund et al. 2012; Duan et al. 2016; Han et al. 2016; Xiang et al. 2015; Xu and Li 2016). One of the reasons for this trend, as recognized by Han et al. (2016), could be that the analysis of textual components has the potential to identify "guests' true feelings".

Guests' online review feedback is an essential source of information for operations improvement. Review metadata, together with the corresponding review text, have the potential to allow hotels to assess global factors that affect review generation (such as guests' age, travel purpose, and language and the hotel category) and to assess guest behavior. For example, information about guests' language or nationality allows a better understanding of cultural differences (Cantallops and Salvi 2014). However, to our knowledge, only a handful of previous studies have examined the language of online reviews or the nationality of reviewers to understand how cultural differences affect reviews. With a sample of hotel guests from the U.S. and Singapore, Ayeh et al. (2016) studied the intention to use online reviews to make a purchase decision. The authors found that the antecedents of the intention to use customer-generated reviews differed between countries. Schuckert et al. (2015a) studied the rating behavior of English-speaking and non-English-speaking guests in Hong Kong hotels and, among other findings, discovered that the former awarded higher ratings than the latter. Hale (2016) studied the similarities in the ratings of London tourist attractions given by speakers of different languages, and the author ended up questioning the (common) practice of averaging the ratings of the reviews in different languages. There is a need for research that conducts an analysis across a large number of nationalities (Ayeh et al. 2016) and diverse international destinations (Schuckert et al. 2015a). Moreover, studying hotel reviews' quantitative ratings and their textual components by language could reveal more than each guest's opinion; it could explain the opinions of groups (Cantallops and Salvi 2014; Schuckert et al. 2015a; Han et al. 2016).

To understand how guests from different origins and in different international destinations assess hotels in online reviews, the present work uses data science tools such as statistics, data visualization, and natural language processing to study similarities and dissimilarities among reviews in different languages. Online reviews from 56 Portuguese hotels (29 city hotels and 27 resort hotels) published on the Booking and TripAdvisor websites were used in conjunction with occupancy data extracted from 8 of these 56 hotels' property management systems (PMS).

Although it is a relatively small country, Portugal is ranked 32<sup>nd</sup> worldwide in terms of the contribution of travel and tourism to its national gross domestic product (GDP) (World Travel & Tourism Council 2016), at 16.4%. Of the 19.1 million guests who stayed in tourist Portuguese lodging facilities in 2015, almost 71% resided in countries whose main official language is Portuguese, English, or Spanish (Instituto Nacional de Estatística 2016). For this reason, and because of the difficulty of processing text in a multitude of languages (Han et al. 2016), online reviews in these three languages were selected as the object of this study.

By using quantitative and qualitative data from two data sources, two types of hotels, and two distinct regions, this study answers previous studies' call for new research on online reviews with a fusion of data from different sources (Hale 2016; Han et al. 2016; Pacheco 2016), different cities (Hale 2016; Han et al. 2016; Pacheco 2016; Schuckert et al. 2015a), different hotel types (Kim et al. 2015), and different languages (Cantalops and Salvi 2014; Schuckert et al. 2015a). In summary, this study seeks to answer two main questions:

Q1) What is the relationship between the language in which online reviews are written and the ratings for the same hotel type/category?

Q2) What preferences/opinions does the textual component of the reviews reveal in each language?

This paper is organized as follows. Section 2 discusses the research background and related literature that identifies the state of the art on this topic. Section 3 describes the data, their collection, their processing, and the applied techniques. Section 4 presents and discusses the results. Section 5 presents the conclusion, implications, and limitations and provides directions for further research.

## 2 Research Background and Related Literature

### 2.1 Hotel Online Reviews

According to Surowiecki (2005), a crowd is a diverse collection of independent individuals that is better at making certain types of decisions or predictions than its individual members or even experts. Online reviewers may be considered a crowd because of their diversity of opinion, independence, decentralization, and aggregation. Thus, it is unsurprising that customers value hotels' online ratings more than hotels' official classifications or stars (Öğüt and Onur Taş 2012). Numerous studies describe the influence of online reviews on customers' purchasing decisions (Cantalops and Salvi 2014; Kim et al. 2015; Torres et al. 2015). However, according to Vermeulen and Seegers (2009), this influence is clear only when the customer is not familiar with the hotel; otherwise, the customer is less influenced by online reviews. Nonetheless, better online reviews allow hoteliers to exert greater pricing power (Anderson 2012; Kim et al. 2015; Torres et al. 2015); thus, they can reduce consumer sensitivity to price and enhance business performance, particularly in hotels with better official classifications (Öğüt and Onur Taş 2012).

Hotels should encourage guests to post both positive and their negative opinions online (Torres et al. 2015) since the positive impact of a positive review can overcome the harm caused by a negative one (Phillips et al. 2016; Vermeulen and Seegers 2009). This is true particularly with a high number of reviews (Melian-Gonzalez et al. 2013; Torres et al. 2015). Moreover, negative opinions can be constructive and enable hoteliers to channel their resources to improve less positive aspects of their hotels (Torres et al. 2015). Online reviews also have the capability to act as a tool for hotel proprietors to hold management teams accountable for the hotel's reputation (Torres et al. 2015), making online reviews more of an opportunity than a threat (Vermeulen and Seegers 2009).

## 2.2 Text Analytics of Online Reviews

Most hotel review websites break down the format of reviews into two components: quantitative (ratings summary) and qualitative (unstructured text) (European Commission 2014). Nevertheless, most studies use only quantitative features, such as ratings or total number of reviews (Duan et al. 2016). This approach can hide or distort the real value of reviews, especially when researchers use indexes that aggregate ratings from several sources (Bjørkelund et al. 2012) and do not consider how the rating scheme is created. The European Commission (2014) found that only 30% of hotel review websites in Europe explained their rating systems. Mellinas et al. (2015) also highlighted this fact by demonstrating that most studies assume that ratings for Booking are on a scale of 1 to 10 when in reality, the website uses a scale ranging from 2.5 to 10. Complementing the quantitative component with the qualitative component has the potential to provide a richer view of online reviews (Duan et al. 2016). This potential can be achieved through the application of text mining and text analytics. Text mining and text analytics rely on natural language-processing, pattern-discovery, and advanced presentation-layer elements such as visualization tools to extract meaningful information (Feldman and Sanger 2007), such as frequent terms, topics, sentiment polarity, or relationships, from the text.

The literature on the application of opinion mining and sentiment analysis to online reviews of products and services is extensive (Gupta et al. 2010; Wang et al. 2016). However, in the tourism and hospitality industries, a literature survey carried out by Schuckert et al. (2015b) revealed that only 8 (16%) of 50 articles published from 2003 to 2014 employed sentiment analysis. Sentiment analysis divides reviews into positive and negative ones, thereby guaranteeing an unbiased interpretation of text (Han et al. 2016; Kwok et al. 2017). A few recent studies have used sentiment analysis to explore the vast capability of the textual component as a feature that benefits hotel management: Duan et al. (2016) performed sentiment analysis to demonstrate the importance of customer preference for service quality and to evaluate service performance. Han et al. (2016) applied several natural language-processing algorithms to understand how hotels could improve their operations and better meet customer expectations. Xu and Li (2016) explored the determinants of customer satisfaction and dissatisfaction toward hotels, and Xiang et al. (2015) applied text analytics to deconstruct guest experiences and examine their association with ratings of reviews. Although this approach is not specified by the authors, it seems that all of the abovementioned studies except Han et al. (2016) target only reviews written in English. However, exploring the textual component of reviews and differentiating the reviews by the language they were written in encourages the discovery of insights specific to reviewers' cultural backgrounds: "text analysis across multiple languages presents methodological difficulties. However, when those issues are overcome, online reviews will potentially yield insights about cultural effects that can further aid hotel managers in improving their customer experiences" (Han et al. 2016, p. 17).

## 2.3 Cultural Differences and Online Reviews

The European Commission is so concerned about consumers' reliance on online hotel reviews and the possible damage caused by biased or false reviews that it ordered a complete study on the subject (European Commission 2014). This comprehensive study recognizes that there are noticeable differences between European Union countries in terms of the importance placed on online hotel reviews (European Commission 2014). The study associates these differences with a shift away from "traditional" travel agencies and toward online agencies as well as with Internet use, online purchasing behavior, consumer travel habits, and consumers' place of origin/residence. In a more recent study, Ayeh et al. (2016) recognized differences in the adoption of online reviews and claimed that despite the importance of online reviews, "antecedents of the intention to use consumer-generated reviews may differ considerably from country to country" (p. 151). Hofstede (1984) and House et al. (2004) also described the effect of society and culture on

the conceptions of society members and how these conceptions affect behavior. Similarly, Chen et al. (2012) recognized that different cultural backgrounds and languages may be the source of many of the variations in consumers' perceptions of and reactions toward products and services. Despite these differences, few studies examine how reviews written in different languages rate the same hotel and what consumer values they reflect. Schuckert et al. (2015b) acknowledge that the discovery of cultural differences in online reviews or ratings is a promising line of research, but to our knowledge, only a few relatively recent works have studied this subject. Schuckert et al. (2015a) studied more than 86,000 online reviews of Hong Kong hotels on TripAdvisor. They concluded that English-speaking guests give higher ratings to hotels than non-English-speaking guests because of their cultural background and that hotel ratings with more non-English guests were negatively affected. Although Hale's (2016) work did not focus on hospitality, it studied how travelers rate London attractions in online reviews according to their language. The author examined more than 516,000 ratings from reviews published on TripAdvisor and revealed that ratings in English reviews were on average slightly (with very low statistical significance) lower than those in non-English reviews. Pacheco (2016) studied the ratings in 2,150 TripAdvisor reviews for 43 hotels from Oporto, Portugal, and concluded that these ratings differed by language groups. The highest ratings were given by Brazilian-Portuguese-speaking travelers, and the lowest were given by Spanish-speaking travelers. Liu et al. (2017) studied 412,784 TripAdvisor reviews for 10,149 hotels in Chinese cities to understand how travelers rated five hotel attributes (rooms, location, cleanliness, service, and value), and they concluded that ratings on these attributes differ substantially by language.

Different languages have different degrees of expressive power (Ravi and Ravi 2015). Language divides the world into concepts that are not defined by nature, and what a person finds "natural" depends on the conventions the person was taught (Deutscher 2010). Deutscher (2010) contends that habits of speech according to each person's mother tongue can generate habits of mind. Because language affects personal conduct, the simultaneous study of hotel online reviews' quantitative ratings and their textual components, by language, has the potential to reveal more than each guest's individual opinion (Cantalops and Salvi 2014; Han et al. 2016; Schuckert et al. 2015a).

### 3 Materials and Methods

To explore both quantitative and qualitative aspects of online reviews in diverse languages, data on approximately 56 hotels in Portugal were collected from Booking and TripAdvisor, two of the largest travel websites presenting hotel reviews (European Commission 2014). For 8 of the 56 hotels, the reviews were integrated with PMS data. This enabled a rarely adopted angle for studying online reviews, i.e., the examination of the similarities (or dissimilarities) between reviews in different languages. It was infeasible to do this for the multitude of languages spoken by guests who stayed in the studied hotels; thus, this work concentrated on reviews written in Portuguese, Spanish, and English.

Because both reviews and PMS data were stored in SQL server databases, the extraction process was performed by means of TSQL queries. The corresponding subsequent data analyses were carried out in R (*R Core Team* 2016).

#### 3.1 Data Collection

Previous studies have noted the lack of research that examines hotel online reviews by language, hotel type, city/region, and publication source (Cantalops and Salvi 2014; Hale 2016; Han et al. 2016; Kim et al. 2015; Schuckert et al. 2015a). Studies have also described that is very difficult to access hotel occupancy/sales data, and thus, authors often must rely on proxy measures to

extrapolate the total number of guests who post reviews. These proxy measures use metrics such as the total number of reviews, the total number of room sales (Öğüt and Onur Taş 2012; Ye et al. 2009), and the annual performance data averaged by month (Kim et al. 2015). To avoid this limitation and to explore the relationship between the number of online reviews and the actual number of rooms occupied, this research used a predefined set of hotels. This set consisted of eight hotels—four city hotels and four resort hotels—all of which were classified as upper-upscale/upscale hotels, i.e., had four or five stars according to Portugal’s official classification. The hotels agreed to provide access to their occupancy data for period of July 2015 to June 2016. Each hotel identified five direct competitors. Additional hotels were also selected to have a suitable sample considering the official classification levels and the hotel type. The city hotels were located in Lisbon, the Portuguese capital, and the resort hotels were located in Algarve, a well-known Portuguese resort region. The hotel set distribution (Table 1) is in line with the distribution of hotels by region, as noted in the 2015 Portuguese official statistics (Instituto Nacional de Estatística 2016).

**Table 1** Hotel set summary

Hotel classification	City		Resort	
	Hotels	Average rooms	Hotels	Average rooms
Two stars	5	52	4	32
Three stars	6	65	6	77
Four stars	12	127	12	202
Five stars	6	224	5	116
Total	29	117	27	106

Using custom-built web content extractors working in a temporal window of 6 months (January to June 2016), all reviews that were written in Portuguese, Spanish, and English and published between July 2015 to June 2016 were collected. Note that all reviews were collected in the original language they were written in and not in a machine-translated version. To avoid collecting machine-translated reviews (or reviews in other languages), the web extractors simulated a human user and selected a feature (present on both Booking and TripAdvisor websites) that enables users to view only reviews written in a certain language. Each hotel’s daily overall rating, rating by attribute, and total number of reviews were also collected. The initial total number of reviews was 23,353, and after duplicates and incorrect language classifications were removed, a total of 23,322 remained (Table 2).

**Table 2** Review frequency and distribution summary

	Booking		TripAdvisor		Total	
	N	%	N	%	N	%
<b>Hotel classification</b>	<b>15,772</b>	<b>67.6%</b>	<b>7,550</b>	<b>32.4%</b>	<b>23,322</b>	<b>100.0%</b>
Two stars	1,747	11.1%	296	3.9%	2,043	8.8%
Three stars	2,795	17.7%	727	9.6%	3,522	15.1%
Four stars	8,775	55.6%	4,896	64.8%	13,671	58.6%
Five stars	2,455	15.6%	1,631	21.6%	4,086	17.5%
<b>Hotel type</b>	<b>15,772</b>	<b>67.6%</b>	<b>7,550</b>	<b>32.4%</b>	<b>23,322</b>	<b>100.0%</b>

City	10,777	68.3%	3,053	40.4%	13,830	59.3%
Resort	4,995	31.7%	4,497	59.6%	9,492	40.7%
<b>Language</b>	<b>15,772</b>	<b>67.6%</b>	<b>7,550</b>	<b>32.4%</b>	<b>23,322</b>	<b>100.0%</b>
Portuguese	5,966	37.8%	1,647	21.8%	7,613	32.6%
Spanish	3,894	24.7%	835	11.1%	4,729	20.3%
English	5,912	37.5%	5,068	67.1%	10,980	47.1%

Although the data sources are similar, there are some important differences between them, as recognized by Bjørkelund et al. (2012). On one hand, both Booking and TripAdvisor have an overall rating and a textual component for each review. On the other hand, Booking ratings are presumed to be in a continuous range [1,10], while TripAdvisor ratings are in a discrete range {1, 2, 3, 4, 5}. Whereas TripAdvisor presents only one textual component, Booking presents two separate textual components: one for exposing positive aspects and another for negative aspects. Also important is that although both sources allow users to give ratings by attribute (cleanliness, location, comfort, etc.), Booking only shows aggregated results per hotel, whereas TripAdvisor shows results by attribute, per review. The reviewer's country is included in most Booking reviews, but country information is not mandatory for user identification for TripAdvisor and thus is mostly missing from reviews.

## 3.2 Data Preparation

### 3.2.1 Selecting, Cleaning, Merging, and Formatting of Data

Data preparation began by merging both Booking and TripAdvisor review data. Only common variables from both sources were included in the resulting dataset. Two new calculated variables, *review description* and *normalized rating*, were added. For TripAdvisor reviews, the description is just the transposition of the text component of the review. For Booking reviews, the description is the concatenation of both the positive and negative text fields. Because the two sources have different rating scales, it was necessary to create a normalized rating by applying min-max normalization, one of the most common normalization methods used to scale variables (Abbott 2014). Normalization was achieved with the formula  $x' = \frac{(x - \min(x))}{(\max(x) - \min(x))} \times 100$ , scaling ratings in the range of 1 to 100. Given the previously presented skewness of Booking review ratings, the minimum rate considered for these reviews was 2.5.

The occupancy data for the eight hotels integrate data on all hotels' PMS and includes the following variables: *date*, *HotelCommonID*, *total rooms*, *total rooms occupied by Booking*, *total rooms occupied by others*, *total room checkouts by Booking*, and *total room checkouts by others*. Additional variables containing information about the total number of room checkouts by Booking and by other agencies, per the main official language, were also included. These variables were calculated by crossing the reviewer's country of residence with the country's main official language (Central Intelligence Agency n.d.), which, although somewhat imprecise, acts as a heuristic measure.

### 3.2.2 Text Preprocessing

Text preprocessing is the most important step in transforming unstructured data (text) into a structured form (Feldman and Sanger 2007; Han et al. 2016); it allows for the retention of significant information and the removal of irrelevant information. This approach is based on the bag-of-words model, one of the most popular transformation processes used in text mining. This method involves the creation of a document-term matrix that considers each document (in this

case, each review textual component) a row and each term in each document a column displaying the frequency of its appearance in the document. Before using the bag-of-words approach, the following typical preprocessing steps must be executed (Feldman and Sanger 2007):

1. Transform all text to lowercase.
2. Normalize related entities: Transform to the same words that appear in different forms in all languages. For example, “wi-fi” and “wi fi” were converted to “wifi”, and a well-known location in Lisbon called Marquês de Pombal, which appeared in different forms, such as “marques do pombal”, “marquês do pombal”, “marquês de pombal”, and “marquês pombal”, was converted to “marquês-pombal”.
3. Perform stemming of common hospitality words such as “rooms”, “restaurants”, “bars”, and others that could be meaningful for data interpretation. Stemming means reducing words to their basic form—for example, removing the “s” from “restaurants”. This was done for each language.
4. For each language, normalize different spellings of the same words or expressions that could be written differently or could be misspelled. For example, in English, transform “didn’t” and “didnt” to “did not”.
5. Standardize domain-specific terms for each language. For example, in Portuguese, numerous words, such as “equipa” (team), “pessoal” (personnel), “funcionários” (employees), or “colaboradores” (collaborators), are used to describe hotel staff. Other examples related to guest origin also had to be taken into consideration. Brazilian Portuguese differs in some ways from the European Portuguese language, and because Brazil is an important market in Portugal, terms from Brazilian Portuguese such as “café da manhã”, “ônibus”, or “metrô” had to be transformed to the European Portuguese equivalents “pequeno-almoço”, “autocarro”, and “metro”, respectively (in English, “breakfast”, “bus”, and “metro”).
6. Remove punctuation, numbers, and stop words.

### 3.2.3 Sentiment Analysis

Because the analysis of huge volumes of opinions, as in the case of online reviews, can be a cumbersome process, sentiment analysis is a common technique used to obtain a generalized opinion summary (Han et al. 2016; Ravi and Ravi 2015). Sentiment analysis, or opinion mining, is the computational study of people’s opinions about entities, individuals, events, topics, and their attributes; it allows for the quantification of opinions according to their polarity (positive, negative, or neutral) (Liu and Zhang 2012). By assigning each review a polarity value based on the textual component, it is possible to compare what is mentioned in the textual component of reviews with their rate.

Sentiment analysis was performed for each document (review) to understand each review’s global opinion polarity, analogously to the approach of Han et al. (2016) and Bjørkelund et al. (2012). It was also performed for each sentence, which allowed the measurement of opinions in terms of more particular aspects (Duan et al. 2016). A dictionary-based approach, also known as a lexicon-based approach, was adopted. Dictionaries are a collection of opinion words with a polarity classification (Ravi and Ravi 2015). Dictionary selection is an important methodological consideration with respect to its adequacy to the domain of the text (Han et al. 2016). In this case, because no specific dictionary for the hospitality domain was found for any of the languages, the criterion to choose dictionaries was based on relatively easy transformation, completeness (dictionaries had to have an extensive range of words), and openness (dictionaries should not be specific to any type of domain). Thus, the SentiLex-PT 02 sentiment lexicon (Silva et al. 2012) was chosen for Portuguese, the ElhPolar dictionary (Saralegi and San Vicente 2013) was chosen for Spanish, and the well-known Opinion Lexicon from Hu and Liu (2004) was chosen for English.



Sentiment strength was calculated by counting positive and negative words and then applying the formula of Bjørkelund et al. (2012):

$$sentiment\ strength = \frac{\sum positive\ words}{\sum positive\ words + \sum negative\ words'}$$

where 0 is perfectly negative, and 1 is perfectly positive. Finally, a column with the sentiment strength value of each review was added to the dataset.

### 3.3 Data Visualization

The textual component of online reviews has the potential to reveal what guests like and dislike about their overall travel experience, without being limited to the number of attributes that quantitative ratings impose (cleanliness, value for money, location, etc.). Therefore, exploring the text may lead to additional insights about complaints and pinpoint what customers really value and view as important. This is especially useful for understanding the differences between customers from diverse nationalities.

As advocated by Feldman and Sanger (2007), because data visualization is an important tool to extract meaningful information, a visualization presentation was devised to illustrate the main aspects that reviewers mentioned in each language and how reviewers classified them. To achieve this, chord diagrams (Holten 2006), a graphical method used to display interrelationships among points typically used with network graphs, was employed. These diagrams portray the relationships among words, and they display the interrelationship between more frequent nouns and the adjectives used to characterize them. Edges (lines) connecting the nodes (words) vary in color and thickness. The color varies according to the sentiment polarity (red for negative, blue for neutral, and green for positive), and the thickness varies according to the correlation (calculated based on the number of times both the noun and the adjective appeared in the same sentence).

The nodes and edges of each graph were created based on the following algorithm:

1. Elaborate a list of terms to exclude. These are terms that could be both nouns and adjectives but are not very commonly used in this domain (e.g., “nice” is a common adjective used in English online reviews, but it is also a noun, i.e., the French city of Nice).
2. Obtain the top 25 most frequent nouns. Each noun will be a node in the graph.
3. Obtain all adjectives that appear in the same sentence as nouns with a correlation equal or superior to a given threshold. (A threshold of 0.06 was defined by testing different values until a suitable number of adjectives was returned.)
4. If an adjective has not appeared before it, then it should be added to the graph as a new node.
5. Calculate the sentiment strength for each sentence where both a noun and an adjective appear.
6. Add graph edges for each tuple <noun, adjective>.
7. For each of the edges, calculate the sentiment strength based on the formula

$$tuple\ strength = \frac{\sum sentences\ strength}{\sum sentences}$$

Because the presentation should show a categorical value (positive, neutral, or negative), binning was employed; this is a common technique used to convert numeric variables to categorical ones (Abbott 2014). Values from 0 to 0.33 were considered negative, from 0.33 to 0.66 were considered neutral, and above 0.66 were considered positive.

To identify the nouns and adjectives, several ontological lexical databases were used, i.e., WordNet 3.0 for English (Miller 1998), Onto.PTv0-4 for Portuguese (Oliveira and Gomes 2014), and WordNet LMF-ES for Spanish (Euskal Herriko Unibertsitatea. IXA Taldea 2014). Part-of-speech tagging was not performed; therefore, apart from the noun exclusion list, no other disambiguation method was employed.

For the readability and validity of the presentation, only the top 100 edges and edges with a minimum of 10 appearances are displayed in the diagrams.

## 4 Results and Discussion

The results obtained from the analysis of the reviews' quantitative and textual components clearly reveal different behaviors toward online reviews depending on language. These results are in line with the findings of Schuckert et al. (2015a), Hale (2016), Pacheco (2016), and Liu et al. (2017) but extend their results by demonstrating there are differences in the quantitative and qualitative (textual) components.

### 4.1 Quantitative Analysis

It can be observed that the distribution of the number of reviews by hotel type, hotel classification, and language is quite distinct, especially between English and non-English reviews (Fig. 1). Portuguese and Spanish reviewers seem to prefer to publish their reviews on Booking rather than TripAdvisor. In contrast, English reviewers seem to prefer TripAdvisor over Booking. The only exception for this is in five-star city hotels, where English reviewers seem to prefer Booking.

Although this behavior similar in resort hotels, it changes in four-star hotels, where TripAdvisor reviews far outnumber Booking reviews. The variation for five-star hotels may be due to their popularity among corporate guests, who may use other channels to make a reservation. The difference for four-star hotels could derive from the fact that in resorts, the influence of traditional travel operators from the United Kingdom, Germany, and other important markets is still preponderant, as confirmed in Fig. 2. Booking accounted for only 23.4% of all room checkouts in the four resort hotels but 32.8% in the four city hotels. In contrast to TripAdvisor, only users who booked through Booking are allowed to publish reviews on its website (Bjørkelund et al. 2012). Thus, customers from traditional travel operators publish their reviews on websites that do not require a booking, such as TripAdvisor or similar websites.

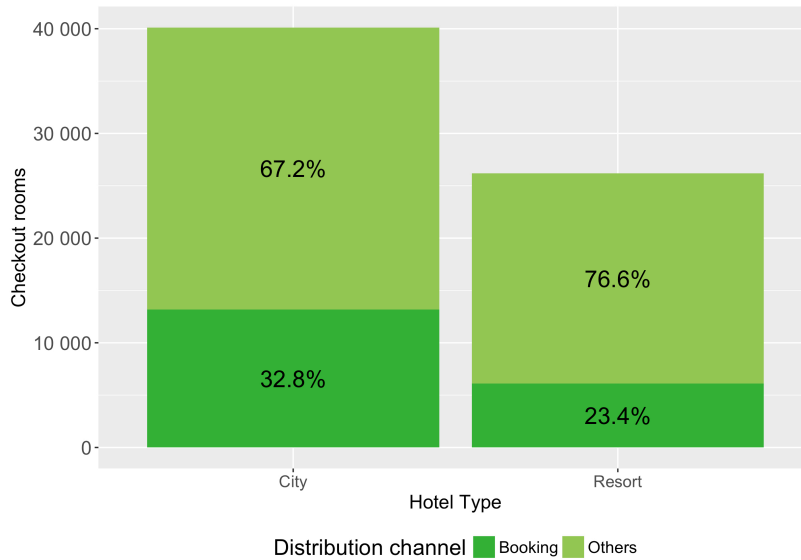
Notably, the ratio of English-speaking reviewers (line and percentage in each bar in Fig. 3) is much higher than that of Portuguese- or Spanish-speaking reviewers. The ratio of reviews published by Spanish-speaking guests is similar for city and resort hotels, but the ratio for Portuguese-speaking guests is very different. This suggests that Portuguese-speaking guests at resorts are more willing to share their experiences than those who stay at city hotels.

Some guests may prefer to write their reviews in English rather than in their mother tongue because of the universality of the English language, which might contribute to the higher number of reviews in English. This is perceptible in resort hotels, where nearly 98% of reviews are written in English. The preference to write in English is depicted in Fig. 4, where a significant number of reviews written in English come from users whose main official language in their countries of residence is not English (users from Portugal, Spain, Brazil, Switzerland, Germany, and the Netherlands, among others). A similar phenomenon occurs with other languages but to a much less significant degree. For example, there are users from Switzerland, France, and Belgium who write reviews in Portuguese. Details of these reviews show that the user often has a Portuguese name or surname but resides in one of those countries, most likely working there (Portugal has a large emigrant community that returns to the homeland for holidays.)

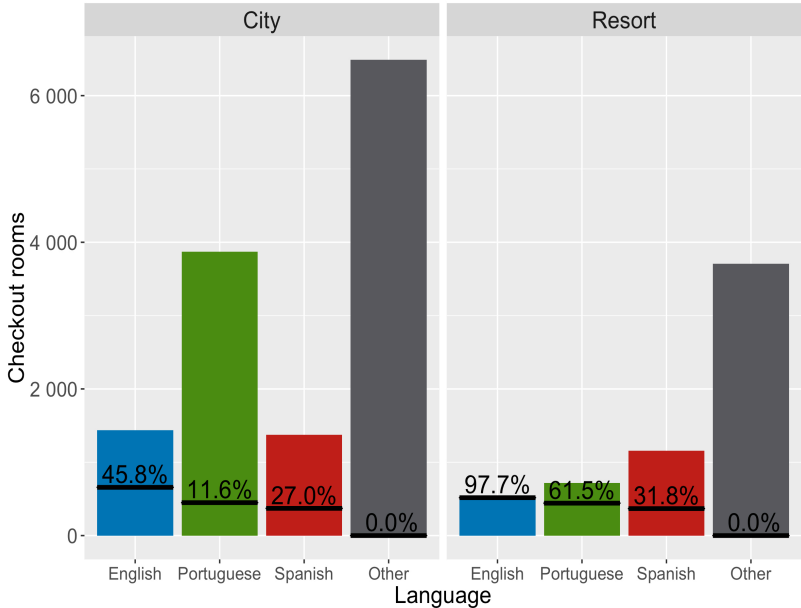
**Fig. 1** Number of reviews by hotel type, hotel classification, and language



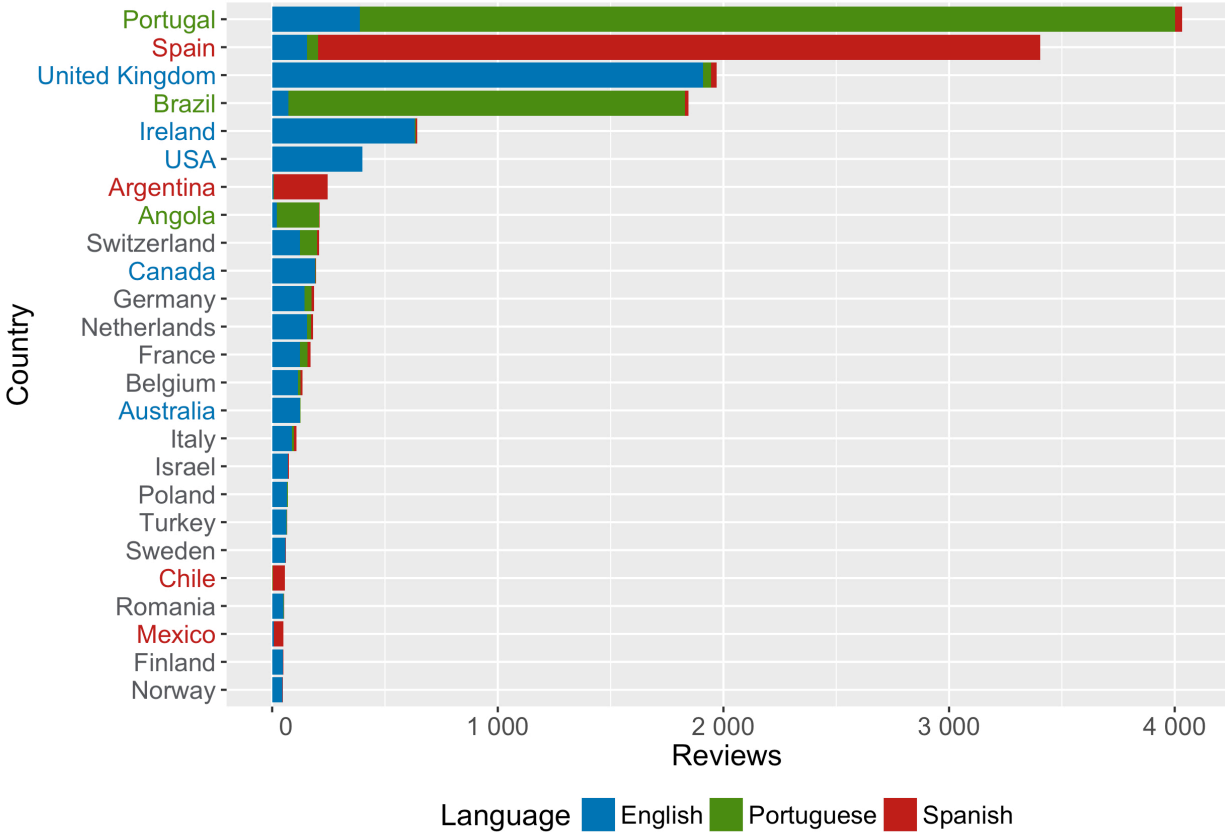
**Fig. 2** Checkouts by distribution channel



**Fig. 3** Booking customers' reviews ratio per checkouts



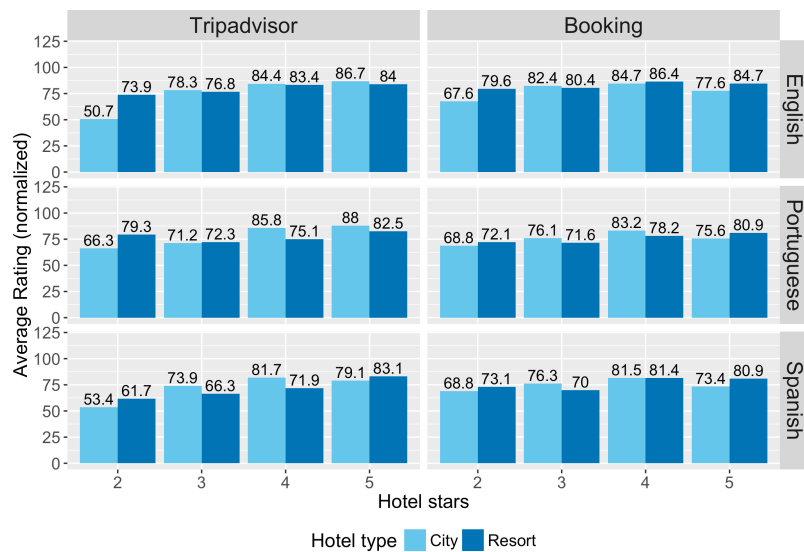
**Fig. 4** Top 25 published reviews on Booking by country and official language



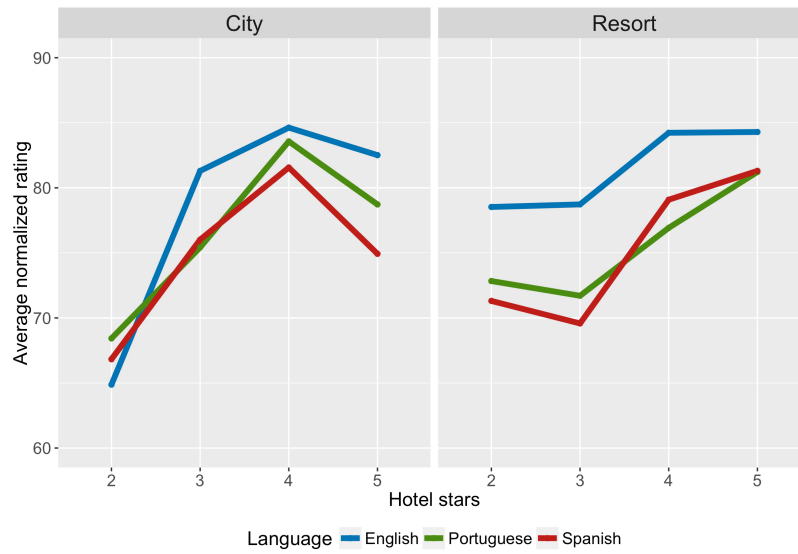
Figs. 2-4 only use Booking findings because TripAdvisor allows anyone to publish reviews. Thus, customers can publish more than one review, and not all reviews may come from real hotel customers, which prevents the counting of bookings. Furthermore, in TripAdvisor, the location field of the reviewer is neither mandatory nor validated, which imposes limits on assessing the country of most of TripAdvisor reviewers. Therefore, analyses that involve the country of the reviewer cannot include TripAdvisor reviews. Because Booking acts as a travel agency, using PMS data allows accounting for the number of bookings from Booking (i.e., room checkouts). Additionally, it is possible to calculate the ratio of Booking guests who posted reviews.

As depicted in Fig. 5, language differences are especially noticeable in terms of ratings. The normalized average rating for English reviews is 79.8 (out of 100) overall, whereas that for Portuguese and Spanish reviews is 76.1 and 75.1, respectively. This difference is even clearer in Fig. 6, where English reviews present higher ratings than Portuguese and Spanish reviews, except for two-star city hotels. For resort hotels, English review ratings surpass Portuguese and Spanish review ratings by 3 points in five-star hotels and up to 7 points in three-star hotels, with a difference of approximately 5 points for the remaining types.

**Fig. 5** Normalized ratings by source, language, and hotel type



**Fig. 6** Ratings by language, hotel type, and hotel classification



This differentiation of ratings per language, hotel type and hotel classification is verifiable in the analysis of the ratings’ means presented in Table 3. Because ratings are negatively skewed—that is, there are many more high ratings than low ratings—this dataset does not present a normal distribution. The analysis is thus performed using the Kruskal-Wallis method; this method is considered a nonparametric equivalent of the one-way analysis of variance, which does not make a distributional assumption. In addition, Kruskal-Wallis compares the means between groups and tests the probability that a random observation from each group is likely to be a random observation from another group. This analysis was executed in R using the package “pgirmess” (Giraudoux 2016), with a significance threshold set at 0.05. The results show a *p*-value below the significance value (denoted with (\*)) in the table) for both the *normalized rating* and the *sentiment strength* for all groups (language, hotel type, and hotel classification), thus confirming a significant difference in the ratings between groups.

**Table 3** Kruskal-Wallis ratings test results

	Normalized Rating			Sentiment Strength		
	Chi-squared	Df	P-value	Chi-squared	Df	<i>p</i> -value
<b>Language</b>	528.3000	2	(*) 1.9E-12	921.8769	2	(*) 6.6E-20
<b>Hotel type</b>	6.6945	1	(*) 9.7E-03	30.8150	1	(*) 2.8E-08
<b>Hotel classification</b>	1,051.878752	3	(*) 1.0E-23	403.1631	3	(*) 4.6E-87

To better understand these differences, a post hoc analysis was performed using pairwise comparisons for each combination of language, hotel type, and hotel classification. The analysis was executed in R using the chi square distribution and the Nemenyi approach, as implemented in the package “PMCMR” (Pohlert 2014). Table 4’s *p*-values show that for all combinations of hotel type and classification, the *normalized ratings* of English reviews differ significantly (*p*-value < 0.05) from those of Spanish and Portuguese reviews, except for two-star city hotels. In terms of the *sentiment strength*, this difference prevails for all hotel type and classification combinations. On the other hand, although some differences exist between Spanish and Portuguese review ratings, the differences are statistically significant only in four-star city hotels for *normalized rating*

and in four-star resort hotels for *sentiment strength*. Additionally, these results show an interesting similarity between ratings calculated from the quantitative component of reviews (*normalized rating*) and ratings calculated from the qualitative component (*sentiment strength*), with differences observed only in two-star city hotels (English/Portuguese and English/Spanish) and in four-star city and resort hotels (Portuguese/Spanish).

Considering the possible impact of a 1-point increase in ratings on hotel occupancy rates and revenue per available room (Anderson 2012; Ögüt and Onur Taş 2012), this difference among English, Portuguese, and Spanish reviews should be something hoteliers (especially those at resort hotels) seriously consider in their market-mix strategies.

Another interesting perspective comes from looking at the average ratings of Booking reviews in the three languages but according to the main official language of the guest's country (Table 5 and Fig. 7). Guests whose country of residence has the main official language of English tend to rate hotels better than guests of non-English-speaking countries. On the opposite side are guests who live in a country whose main official language is Portuguese or write reviews in Portuguese; these guests tend to give the worst ratings, independently of the language in which the reviews are written. The quantity of users who are from countries where the main official language is not English but write their reviews in English could help explain ratings differences. Only 57.7% of English-written reviews are from reviewers whose country's official language is English, whereas for Portuguese and Spanish, these percentages are 94.4% and 95.2%, respectively. The Kruskal-Wallis test reveals a significant difference between the average ratings by official country language (with a chi square of 392.33, 3 degrees of freedom, and a  $p$ -value of 1.017127e-84, which is below the 0.05 significance value). However, the post hoc analysis (see Table 6) shows that this difference is not statistically significant between reviewers who have Portuguese and Spanish as their official language. This could be explained by the fact that the clear majority of Portuguese- and Spanish-speaking guests in Portuguese hotels come from Portugal and Spain. Because of their proximity, reviewers from these countries do not differ substantially in terms of culture, and they tend to have similar points of view.

**Table 4** Post hoc ratings pairwise results ( $p$ -values)

			Normalized Rating		Sentiment Strength		
			English	Portuguese	English	Portuguese	
City	**	Portuguese	0.40985		(*) 0.00374		
		Spanish	0.81670	0.80602	(*) 0.00005	0.35327	
	***	Portuguese	(*) 4.4E-13		(*) 0.0E+00		
		Spanish	(*) 8.9E-09	0.86322	(*) 7.6E-12	0.16386	
	****	Portuguese	(*) 0.01749		(*) 0.0E+00		
		Spanish	(*) 4.3E-12	(*) 0.00006	(*) 0.0E+00	0.47923	
	*****	Portuguese	(*) 1.7E-08		(*) 0.0E+00		
		Spanish	(*) 1.6E-11	0.07442	(*) 0.0E+00	0.98873	
	Resort	**	Portuguese	(*) 0.00310		(*) 0.00444	
			Spanish	(*) 0.00080	0.84677	(*) 0.00176	0.90611
***		Portuguese	(*) 4.4E-10		(*) 1.7E-13		
		Spanish	(*) 3.3E-07	0.89859	(*) 5.3E-14	0.13119	
****		Portuguese	(*) 0.0E+00		(*) 0.0E+00		
		Spanish	(*) 1.3E-14	0.05061	(*) 0.0E+00	(*) 0.00023	
*****		Portuguese	(*) 0.00066		(*) 4.3E-09		
		Spanish	(*) 0.00263	0.91168	(*) 3.5E-12	0.08489	

**Table 5** Booking review distribution by language and official language of users

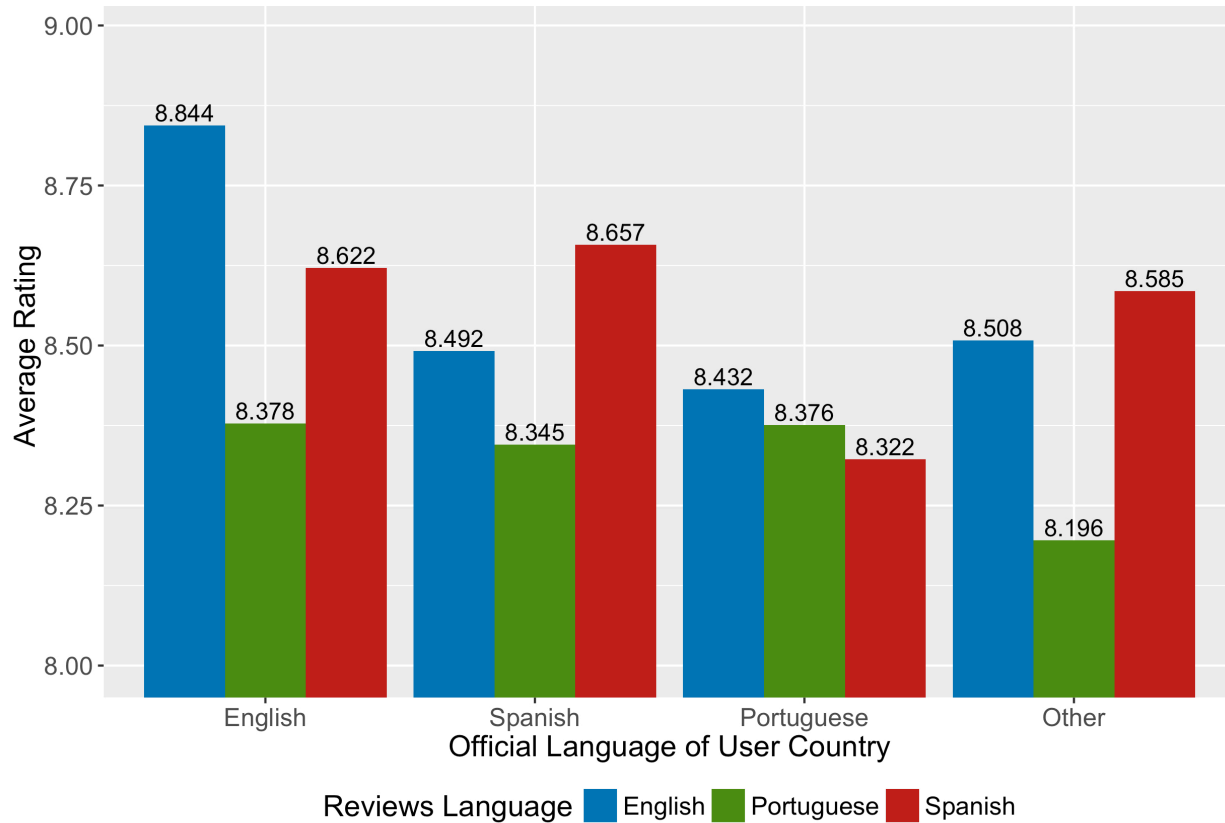
		Language of Reviews					
		English		Portuguese		Spanish	
		N	%	N	%	N	%
Country official language	English	3,391	57.7%	51	0.9%	41	1.1%
	Portuguese	488	8.3%	5,590	94.4%	49	1.3%
	Spanish	180	3.1%	54	0.9%	3,672	95.2%
	Other	1,813	30.9%	228	3.8%	97	2.5%
TOTAL		5,872	100.0%	5,923	100.0%	3,859	100.0%

**Table 6** Post hoc pairwise results ( $p$ -values) per official language

	English	Portuguese	Spanish
Portuguese	(*) 0.0E+00		
Spanish	(*) 0.0E+00	0.67605	
Other	(*) 0.0E+00	(*) 7.3E-07	(*) 0,00037



**Fig. 7** Booking average review ratings by language and official language of users



## 4.2 Text Analysis

The analysis of the textual component of reviews revealed interesting results regarding the correlation between what is written and the respective quantitative rating and what guests from different origins value most.

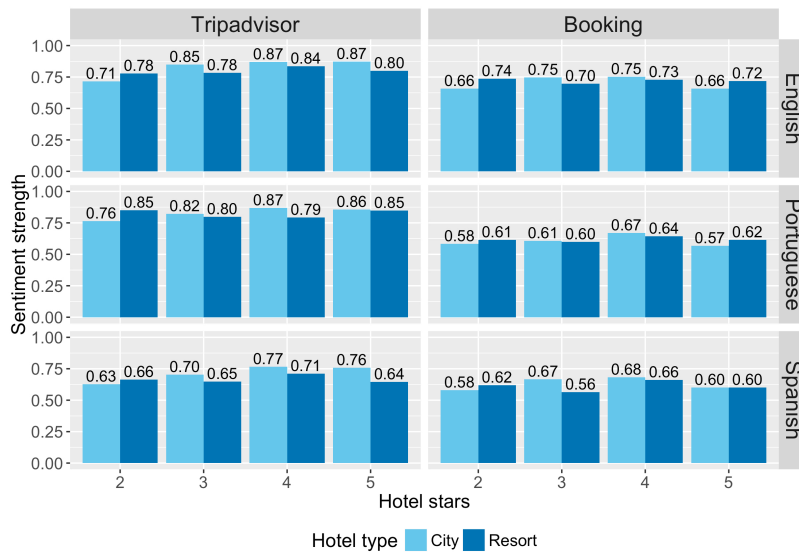
As recognized by Bjørkelund et al. (2012), because of Booking's separate fields for users to comment on positive and negative aspects, users tend to write more summarized texts on Booking than on TripAdvisor; on average, reviews on the latter have more words than reviews on the former, independently of the language used. Again, an analysis of the review sources reveals differences among languages. On average, English reviews have 67 words on TripAdvisor and 14 words on Booking, whereas Spanish reviews have 62 and 13 words, respectively. Reviews in Portuguese are the least wordy, with 49 and 11 words, respectively. The results reveal a negative association between the number of words in reviews and the ratings in the three languages and two sources, meaning that as the number of words increases, ratings decrease. A similar result was previously acknowledged by Han et al. (2016) for English reviews on TripAdvisor.

The new results validate the findings about the association between sentiment strength and review ratings (Table 3 and Table 4); they extend that association for Portuguese and Spanish reviews on both review sources. The sentiment strength ranking by language is the same as the rating ranking. English reviews rank highest, with an overall sentiment strength of 0.76, followed by Portuguese (0.72) and Spanish (0.66) reviews. This association between sentiment strength and review ratings occurs for all review sources, languages, hotel types, and hotel classifications (see Figs. 5 and 8). The association between normalized average ratings and average sentiment strength by written language and source presents moderate/high correlation values (0.63 in Spanish reviews from TripAdvisor) to strong correlated values (0.90 in Portuguese reviews from TripAdvisor). From the perspective of the review source, the association between average ratings

and average sentiment strength is higher for Booking reviews (0.73 correlation) than for TripAdvisor (0.68 correlation).

Other interesting findings are revealed by a simple analysis of word frequency per language and per hotel type, as displayed in Table 7. For city hotels, “room” and then “hotel” are the two words most frequently used in English reviews, validating the findings of Xiang et al. (2015) and Han et al. (2016). The same applies to Portuguese and Spanish reviews, with “habitación” (Spanish for “room”) and “quarto” (Portuguese for “room”) appearing in the same order. These same words/concepts are also the most frequent in resort hotels but in reverse order for all three languages. One possible explanation could be that resort hotel guests have different motivations to book a hotel than those staying in city hotels; thus, they ascribe more importance to the overall “experience” and to the hotel overall than to the room.

**Fig. 8** Sentiment strength by source, language, hotel type, and hotel classification



**Table 7** Top 20 words by hotel type and language

CITY			RESORT		
English	Spanish	Portuguese	English	Spanish	Portuguese
room	Habitación	quarto	hotel	hotel	hotel
hotel	Hotel	hotel	room	habitación	quarto
staff	Desayuno	localização	staff	desayuno	bom
breakfast	Personal	pequeno-almoço	good	ubicación	piscina
good	Ubicación	bom	pool	personal	localização
location	Bien	funcionário	great	bien	pequeno-almoço
great	bueno	excelente	view	playa	simpatia
nice	Cama	bem	breakfast	piscina	praia
friendly	Metro	conforto	beach	bueno	funcionário
helpful	Centro	atendimento	restaurant	vistas	excelente
clean	Baño	tudo	food	cama	vista
excellent	Excelente	cama	lovely	wifi	bem
lisbon	Atención	simpatia	nice	zona	tudo
restaurant	aparcamiento	metro	location	limpieza	qualidade
stay	Lisboa	lisboa	friendly	tranquilidad	limpeza
comfortable	Amable	pequeno	clean	calidad	conforto
metro	Pequeño	ótimo	Bar	precio	ter
one	Precio	restaurante	one	buffet	ser
view	Buen	nada	just	anque	restaurante
city	Servicio	vista	stay	excelente	mar

“Staff” (“personal” in Spanish and “funcionário” in Portuguese), “breakfast” (“desayuno” in Spanish and “pequeno-almoço” in Portuguese), and “location” (“ubicación” in Spanish and “localização” in Portuguese) are among the more frequent words in all three languages, with little changes in ranking among hotel types and languages.

Table 7 also illustrates the differences in what customers write about for city and resort hotels. For example, in all three languages, “metro” (subway) or “Lisbon” (“Lisboa” in Portuguese and Spanish) appear in the top 20, whereas in resort hotels, these words are replaced by words such as “pool” (“piscina” in Spanish and Portuguese), “view” (“vista”/“vistas” in Spanish and Portuguese), or “beach” (“playa” in Spanish and “praia” in Portuguese).

Word frequencies are helpful to understand which topics guests mention by language, but they do not expose what guests feel or think. Reading each review where the word/topic is mentioned to understand what guests are saying would be a very time-consuming task. This is where data visualization can help. As expressed by Ware (2009), “one of the great benefits of data visualization is the sheer quantity of information that can be rapidly interpreted if it is presented well” (p. 2). The chord diagrams shown in Figs. 10, 11, and 12 express the power of data visualization. By displaying the polarity of sentences where nouns are more frequently mentioned, it is possible to identify which adjectives are used in conjunction with the nouns and, at a glance, to understand guest sentiment concerning the most frequent topics. According to the previously explained chord construction algorithm, these chord diagrams are composed of the following:

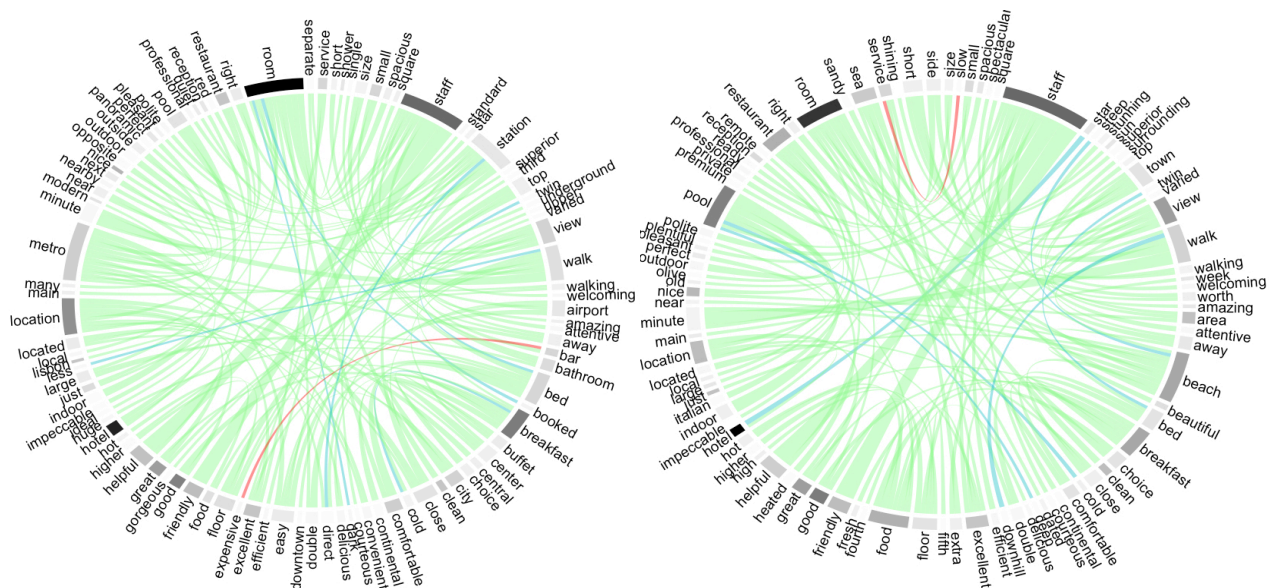
1. Labels (or nodes in graph terminology): words around the chart, which are the top nouns and adjectives mentioned in reviews.
2. Bar below labels: the color of the bar illustrates the word’s overall frequency in reviews. The darker the color, the more frequent the word.
3. Lines connecting words (or edges in graph terminology): a line connecting a noun and an adjective shows that both are mentioned in the same sentence, meaning that there is a correlation between them.

4. Line thickness: the thickness of the line illustrates the correlation value between nouns and adjectives (i.e., the higher the correlation, the thicker the line).
5. Line color: the color illustrates the average sentiment polarity of the sentences where both noun and adjectives were mentioned, with red for negative, blue for neutral, and green for positive.

The visualizations confirm that words such as “room”, “hotel”, and “breakfast” are the most frequent in both hotel types and in all languages, and they confirm the differences between city and resort hotels presented in Table 7. Most importantly, however, these visualizations show the differences among languages.

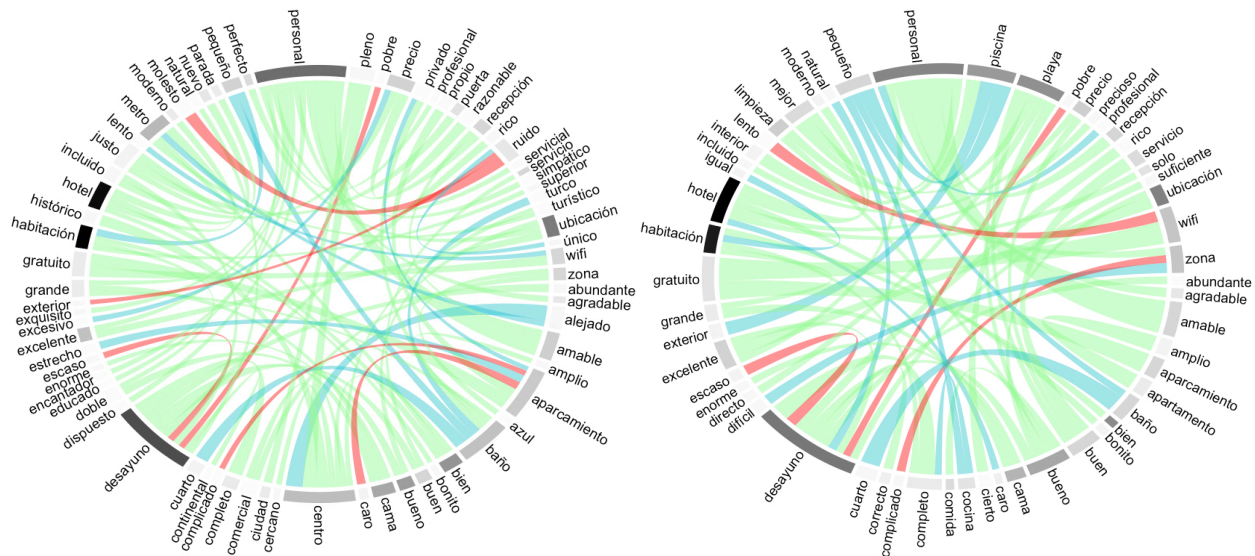
Fig. 9 shows a visualization of the interrelationships among words for English reviews. It exposes some neutral opinions but only one negative topic per hotel type. For city hotels, the negative topic is shown in the edge “bar—expensive;” for resort hotels, in the edge “service—slow”. Neutral opinions are contrasted among reviewers and contradictory sentiments. Taking city hotels as an example, by exploring what was written about “station—direct” or “room—dark”, it is observed that one user mentioned “room small dark upsetting”, which converts to a sentiment of null strength. Another user mentioned “room clean bit dark”, which resulted in a sentiment strength of 0.5, corresponding to a neutral sentiment. The sentence “although understandably dark room in fact spacious met requirements perfectly” corresponds to a positive sentiment strength of 0.67. For resort hotels, neutral relations also appear in topics such as “pool—deep”, “pool—cold”, and “beach—steep” (complaining about how stairs to the beaches are steep but simultaneously relating that the beaches were good). For positive edges, although some are common for both hotel types, such as “staff—friendly”, “staff—helpful”, and “breakfast—good”, others are clearly related to the hotel type. City hotel users’ mentions of location, such as “metro—close”, “metro—easy”, and “location—center”, are replaced with mentions such as “walk—minute”, “view—spectacular”, and “beach—close” for resort hotels.

**Fig. 9** Interrelationships among words in English (city hotels on the left, resort hotels on the right)



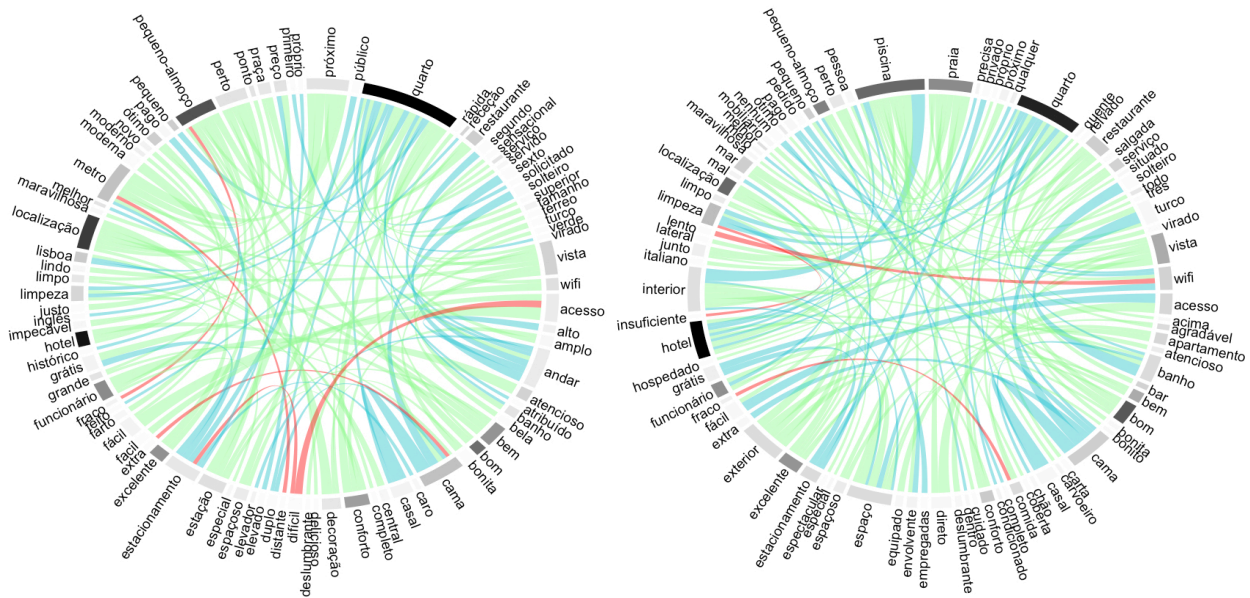
Spanish reviews (Fig. 10) show a contrasting visualization because of the smaller number of edges (which is related to the smaller number of reviews in Spanish) and because of the minimum frequency cutoff defined in the algorithm. Nevertheless, this diagram immediately draws attention to the differences between Spanish and English reviews. For both city and resort hotels, Spanish reviewers showed their discontentment toward breakfast, with edges such as “desayuno—escaso” (“breakfast—scarce”) and “desayuno—pobre” (“breakfast—lack of variety”). They also mentioned words that did not appear in English reviews, such as “aparcamiento” (“car parking”) and “baño” (“bathroom”). While the former is understandable because of Spain’s proximity to Portugal (which leads many people to travel to Portugal by car), the latter is somewhat unexpected. Regarding parking, guests in city hotels complained that it was expensive and complicated (“aparcamiento—caro” and “aparcamiento—complicado”). Regarding bathrooms, in both types of hotels, guests complained about the size: “baño—pequeño” (“bathroom—small”). Spanish reviewers also complained of “noise” (“ruido”) in city hotels and of “wifi” slowness (“lento”) in resort hotels.

**Fig. 10** Interrelationships among words in Spanish (city hotels on the left, resort hotels on the right)



The diagrams for Portuguese reviews (Fig. 11) again reveal similarities and dissimilarities among languages. As with Spanish reviews, Portuguese reviews present a higher ratio of negative and neutral sentences than English reviews. Portuguese reviewers also have a negative or not-so-good (neutral in average) sentiment toward breakfast in city hotels, mentioning “pequeno-almoço-fraco” (“breakfast-weak”). In resort hotels, this neutral sentiment relates to breakfast in the room, “pequeno-almoço-quarto” (“breakfast-room”), and is negative toward the food, “comida-fraca” (“food-weak”). Portuguese reviews also mentioned car parking but did not describe it as negatively as Spanish reviewers. For city hotels, “estacionamento—difícil” (“parking—difficult”) is also mentioned, but there are neutral and contradictory sentences on parking price: “estacionamento-pago” (“parking-paid”) and “estacionamento-grátis” (“parking-free”). In the sentences where this association occurs, most only mention “estacionamento-gratis”, which translates into a sentiment strength of 0.5 (neutral). Portuguese reviews of city hotels also complained about extra beds (which is also mentioned in resort hotels in a neutral way) and distance to the metro, difficult access to the hotel, or hotel parking. As Spanish reviewers did, Portuguese reviewers also complained about “wifi” slowness in resort hotels.

**Fig. 11** Interrelationships among words in Portuguese (city hotels on the left, resort hotels on the right)



At a glance, the three previous diagrams show which topics related to location, comfort, facilities, and food are most mentioned by customers in any of the three languages. In a clear and comprehensible way, they stress what Xu and Li (2016) mentioned as customer satisfaction and dissatisfaction factors, namely, “view”, “location”, “wifi”, “parking”, “bathroom”, and “pool”. Furthermore, the diagrams illustrate that those factors are weighted differently by reviewers in different languages. For example, “wifi” and “parking” are among the top-mentioned words by Portuguese and Spanish reviewers but not by English reviewers. Other examples include “bathroom”, which is widely mentioned and negatively criticized in Spanish reviews but not in English or Portuguese reviews. Although it is widely mentioned in all the studied languages, “breakfast” is only particularly criticized in Spanish reviews.

## 5 Conclusions

This study uses online reviews that are written in Portuguese, Spanish, and English, are published on the Booking and TripAdvisor websites for city and resort hotels, and refer to different official categories (stars). It integrates this information with hotel occupancy data and shows that both ratings and textual components of reviews differ according to the language in which they are written. In addition to confirming what Schuckert et al. (2015a), Pacheco (2016), and Liu et al. (2017) discovered about differences in ratings between English and non-English reviews, this study’s findings support other authors’ suspicions that the textual component of reviews can reveal even more about the influence of guests’ cultural backgrounds on their preferences, likes, and dislikes (Cantalops and Salvi 2014; Han et al. 2016; Schuckert et al. 2015a, 2015b).

This study’s contributions are supported in the following research findings:

1. Although Booking is a very important player in the market, its market share differs for city hotels (32.8%) and resort hotels (23.4%). Therefore, collecting reviews only from Booking might not provide a representation of the different types of guests who write reviews. These findings highlight that online reviews research must use more than one source of data.
2. By using two sources of online reviews, this study validates the possibility of combining reviews from multiple sources that have different textual component schemes and different rating scales. Moreover, this study addresses the question of how aggregate indexes or

normalized ratings are created, in alignment with findings by Bjørkelund et al. (2012) and Mellinas et al. (2015). In addition, the use of Booking, which allows only guests who book through its website to post reviews, produced results that were not otherwise possible to obtain. These results include the finding that the ratios of reviews published in the official language of guests' countries of residence may be similar for the same language but differ among languages.

3. By combining online review data with hotel occupancy data, this study was able to expose Booking's market share by hotel type, and it found that guests' behavior differs with their country of residence. For example, proportionally to the number of reviews in their country's main official language, Portuguese and Dutch guests publish much more reviews in English than guests from other countries.
4. This study expands Schuckert et al.'s (2015a) findings that English reviews give higher ratings than non-English reviews by showing the distinct differences between languages (in this case, Portuguese and Spanish), and it corroborates the findings of Pacheco (2016) and Liu et al. (2017). This study adds new information to previous findings by demonstrating that the difference among ratings is even greater in resort hotels, and the same was found in the sentiment analysis of the textual component of reviews.
5. This study confirms that sentiment analysis is a useful tool to quantify the opinions given by guests in the textual component of reviews. It finds that this is true not only for English but also for Portuguese and Spanish, thereby extending the findings of Bjørkelund et al. (2012), Han et al. (2016), and Duan et al. (2016). The negative association between the number of words in a review and sentiment strength is confirmed not only for English (Han et al. 2016) but also for Portuguese and Spanish reviews. Moreover, the study asserts that average sentiment strength is highly associated with average ratings.
6. By presenting an algorithm to create a chord diagram for a clear visualization of existing relations among highly used nouns and the adjectives used to classify them, this study demonstrates that it is possible to efficiently analyze a vast amount of online reviews and extract knowledge from them. The visualization highlights similarities between reviews in different languages, such as the predominance of certain terms for all languages. It also shows that there are dissimilarities among hotel types and languages.

## 5.1 Implications

From a research point of view, the present study demonstrates that examining online reviews by language has the potential to uncover similarities and dissimilarities among guests with different cultural backgrounds. Thus, more resources should be employed in research on online reviews. This study draws attention to the use of multiple data sources and highlights the need for the prudent execution of online review data aggregation, as different sources have different formats and different rating scales. It also suggests that researchers should employ some caution when generalizing the findings from reviews in one language to reviews in other languages and to reviews that differ in terms of the classification and type of hotels to which they refer.

This study also has managerial implications. First, understanding what guests with different cultural backgrounds criticize or value, such as facilities, amenities, or service, allows hoteliers to act to improve their hotels' social reputation (e.g., have special packages for Portuguese or Spanish customers that include parking or understand what they can do to prevent Spanish-speaking customers from complaining about breakfast). Second, this understanding allows hoteliers to better direct their marketing efforts. Differentiating website content by language could be a strong selling strategy. For example, in the Spanish version of a hotel website, a hotel that has new and spacious bathrooms or rooms with remarkable views should highlight these points to gain a competitive advantage. Third, by stressing the importance of online reviews, this study highlights the need for hoteliers to devise tactics and measures that motivate guests to write

online reviews. Hoteliers can send a post-checkout survey and suggest its automatic publication on TripAdvisor or another website. Fourth, knowing the possible impact of a hotel's social reputation on its performance (Anderson 2012; Kim et al. 2015; Ögüt and Onur Taş 2012; Torres et al. 2015), revenue managers could guide their market-mix strategies to appeal to customers from countries who speak languages associated with higher online ratings or to discourage customers who rate negatively certain aspects that the hotel cannot change (e.g., difficult parking).

From the point of view of hotel proprietary companies, knowing what guests of different backgrounds like and dislike allows companies to make better-informed decisions on where to build hotels and how to improve hotel segmentation according to the countries of the guests they are supposed to attract.

## 5.2 Limitations and Future Work

As with any other study, this one presents some limitations that can also be considered directions for further research. First, because of the difference between the Booking and TripAdvisor review formats, some of the analysis had to be performed using data from only one of the sources. To overcome this limitation, further research should use data from several sources to ensure that at least two of them can be used together. Further research on this topic should also target other languages.

The combination of PMS data with online review data is one of the strengths and innovations of this study. However, the PMS data are available for only four city hotels and four resort hotels, all of which are classified as upscale/upper-upscale hotels. Future research should use data from more hotels and other categories.

The use of chord diagrams is an important visualization novelty introduced in this work. In terms of future work, these diagrams and the algorithm behind their creation have the potential to become a useful tool when exploring larger numbers of online reviews. This potential should be coupled with the creation of dictionaries (by language) of domain terms to look for, which could then be used for text preprocessing. As an example, in case of a hotel in a city, an English customized dictionary would convert the words "metropolitan", "tube", and "subway" to "metro" and convert "conference," "party," "gala dinner," and "seminar" to "event". In this way, instead of looking for top-mentioned nouns, the algorithm could look only for the converted terms and display the adjectives used to classify them. This algorithm could be enhanced with part-of-speech disambiguation, which might improve sentiment analysis performance.

Though interesting to explore, this study did not address the reasons why reviewers writing in English give higher ratings than reviewers writing in other languages. Future research could try to justify these differences in ratings and in the topics covered by reviews in the different languages according to Hofstede's cultural dimensions (Hofstede 1984) or according to psychic distance (Dow and Karunaratna 2006).

## References

- Abbott D (2014) Applied predictive analytics: Principles and techniques for the professional data analyst. Wiley, Indianapolis
- Anderson CK (2012) The impact of social media on lodging performance. *Cornell Hosp Rep* 12:4–11
- Ayeh JK, Au N, Law R (2016) Investigating cross-national heterogeneity in the adoption of online hotel reviews. *Int J Hosp Manag* 55:142–153. doi: 10.1016/j.ijhm.2016.04.003
- Bjørkelund E, Burnett TH, Nørvag K (2012) A study of opinion mining and visualization of hotel reviews. In: *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services*. ACM, pp 229–238
- Cantalops AS, Salvi F (2014) New consumer behavior: A review of research on eWOM and



hotels. *Int J Hosp Manag* 36:41–51. doi: 10.1016/j.ijhm.2013.08.007

Central Intelligence Agency (n.d.) The world factbook: Field listing: Languages. <https://www.cia.gov/library/publications/the-world-factbook/fields/2098.html>. Accessed 7 Feb 2016

Chen RXY, Cheung C, Law R (2012) A review of the literature on culture in hotel management research: What is the future? *Int J Hosp Manag* 31:52–65. doi: 10.1016/j.ijhm.2011.06.010

Deutscher G (2010) *Through the language glass: Why the world looks different in other languages*. Macmillan, New York

Dow D, Karunaratna A (2006) Developing a multidimensional instrument to measure psychic distance stimuli. *J Int Bus Stud* 37:578–602

Duan W, Yu Y, Cao Q, Levy S (2016) Exploring the impact of social media on hotel service performance: A sentimental analysis approach. *Cornell Hosp Q* 57:282–296. doi: 10.1177/1938965515620483

European Commission (ed) (2014) *Study on online consumer reviews in the hotel sector. Final report*. European Commission, Brussels, Belgium

Euskal Herriko Unibertsitatea. IXA Taldea (2014) Spanish WordNet-LMF. <http://hdl.handle.net/10230/22923>. Accessed 6 Feb 2016

Feldman R, Sanger J (2007) *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge University Press, New York

Giraudoux P (2016) *Pgirmess: Data analysis in ecology*

Gupta N, Di Fabrizio G, Haffner P (2010) Capturing the stars: Predicting ratings for service and product reviews. In: *SS '10 Proceedings of the NAACL HLT 2010 Workshop on Semantic Search*. Association for Computational Linguistics, Stroudsburg, pp 36–43

Hale SA (2016) User reviews and language: How language influences ratings. *Hum-Comput Interact* 1208–1214. doi: 10.1145/2851581.2892466

Han HJ, Mankad S, Gavirneni N, Verma R (2016) What guests really think of your hotel: Text analytics of online customer reviews. *Cornell Hosp Rep* 16:3–17

Hofstede G (1984) *Culture's consequences: International differences in work-related values*. SAGE, London, UK

Holten D (2006) Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Trans Vis Comput Graph* 12:741–748

House RJ, Hanges PJ, Javidan M, et al (eds) (2004) *Culture, leadership, and organizations: The GLOBE study of 62 societies*. SAGE Publications, Inc., Thousand Oaks

Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Kim W, Kohavi R (eds) *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, pp 168–177

Instituto Nacional de Estatística (2016) *Tourism statistics-2015*. [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_publicacoes&PUBLICACOESpub\\_boui=265858123&PUBLICACOESstema=55581&PUBLICACOESmodo=2](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=265858123&PUBLICACOESstema=55581&PUBLICACOESmodo=2). Accessed 1 Sep 2016

Kim WG, Lim H, Brymer RA (2015) The effectiveness of managing social media on hotel performance. *Int J Hosp Manag* 44:165–171. doi: 10.1016/j.ijhm.2014.10.014

Kwok L, Xie KL, Richards T (2017) Thematic framework of online review research: A systematic analysis of contemporary literature on seven major hospitality and tourism journals. *Int J Contemp Hosp Manag* 29:307–354. doi: 10.1108/IJCHM-11-2015-0664

Liu B, Zhang L (2012) A survey of opinion mining and sentiment analysis. In: Aggarwal CC, Zhai CX (eds) *Mining text data*. Springer, Boston, MA, pp 415–463

Liu Y, Teichert T, Rossi M, et al (2017) Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tour Manag* 59:554–563. doi: 10.1016/j.tourman.2016.08.012

Melian-Gonzalez S, Bulchand-Gidumal J, Gonzalez Lopez-Valcarcel B (2013) Online customer reviews of hotels: As participation increases, better evaluation is obtained. *Cornell Hosp Q*

54:274–283. doi: 10.1177/1938965513481498

Mellinas JP, María-Dolores S-MM, Bernal García JJB (2015) Booking.com: The unexpected scoring system. *Tour Manag* 49:72–74. doi: 10.1016/j.tourman.2014.08.019

Miller GA (1998) WordNet: A lexical database for English. In: *Communications of the ACM*. MIT Press, Cambridge, pp 39–41

Ögüt H, Onur Taş BKO (2012) The influence of Internet customer reviews on the online sales and prices in hotel industry. *Serv Ind J* 32:197–214. doi: 10.1080/02642069.2010.529436

Oliveira HG, Gomes P (2014) ECO and Onto.PT: A flexible approach for creating a Portuguese wordnet automatically. *Lang Resour Eval* 48:373–393. doi: 10.1007/s10579-013-9249-9

Pacheco L (2016) An analysis of online reviews by language groups: The case of hotels in Porto, Portugal. *Eur J Tour Res* 14:66–74

Phillips P, Barnes S, Zigan K, Schegg R (2016) Understanding the impact of online reviews on hotel performance: An empirical analysis. *J Travel Res* 56:235-249. doi: 10.1177/0047287516636481

Pohlert T (2014) The pairwise multiple comparison of mean ranks package (PMCMR)

R Core Team (2016) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>

Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowl-Based Syst* 89:14–46. doi: 10.1016/j.knosys.2015.06.015

Saralegi X, San Vicente I (2013) Elhuyar at TASS 2013. In: Esteban AD, Loinaz IA, Román JV (eds) *Proceedings of XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural*. Madrid, El Congreso Español de Informática, pp 143–150

Schuckert M, Liu X, Law R (2015a) A segmentation of online reviews by language groups: How English and non-English speakers rate hotels differently. *Int J Hosp Manag* 48:143–149. doi: 10.1016/j.ijhm.2014.12.007

Schuckert M, Liu X, Law R (2015b) Hospitality and tourism online reviews: Recent trends and future directions. *J Travel Tour Mark* 32:608–621. doi: 10.1080/10548408.2014.933154

Silva MJ, Carvalho P, Sarmiento L (2012) Building a sentiment lexicon for social judgement mining. In: Caseli H, Villavicencio A, Teixeira A, Perdigão F (eds) *Computational processing of the Portuguese language*. Springer Berlin Heidelberg, New York, pp 218–228

Surowiecki J (2005) *The wisdom of crowds*, reprint edition. Anchor, New York

Torres EN, Singh D, Robertson-Ring A (2015) Consumer reviews and the creation of booking transaction value: Lessons from the hotel industry. *Int J Hosp Manag* 50:77–83. doi: 10.1016/j.ijhm.2015.07.012

Vermeulen IE, Seegers D (2009) Tried and tested: The impact of online hotel reviews on consumer consideration. *Tour Manag* 30:123–127. doi: 10.1016/j.tourman.2008.04.008

Wang B, Huang Y, Li X (2016) Combining review text content and reviewer-Item rating matrix to predict review rating. *Comput Intell Neurosci Article ID 5968705 2016:1-11*. doi: 10.1155/2016/5968705

Ware C (2009) *Information visualization: Perception for design*, 2nd edn. Elsevier [u.a.], Amsterdam, Netherlands

World Travel & Tourism Council (2016) *Travel & tourism: economic Impact 2016 Portugal*. World Travel & Tourism Council, London

Xiang Z, Schwartz Z, Gerdes JHJ, Uysal M (2015) What can big data and text analytics tell us about hotel guest experience and satisfaction? *Int J Hosp Manag* 44:120–130. doi: 10.1016/j.ijhm.2014.10.013

Xu X, Li Y (2016) The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. *Int J Hosp Manag* 55:57–69. doi: 10.1016/j.ijhm.2016.03.003

Ye Q, Law R, Gu B (2009) The impact of online user reviews on hotel room sales. *Int J Hosp Manag* 28:180–182. doi: 10.1016/j.ijhm.2008.06.011