Juan Pedro Mellinas
Universidad Politécnica de Cartagena

How Languages Influence Online Ratings in the Most Iconic Tourist Spots

Some previous studies have identified variations in online scores about hotels, depending on the nationality or language of the users, observing substantial differences in some cases. We use Tripadvisor reviews to identify differences by country and language in 15 of the most visited tourist spots in Spain, France and Italy, using a sample of more than one million reviews. Some of the differences in scores by language are similar to those of earlier studies about hotels, differing in others. We observe how language behaviors are repeated systematically; speakers of each language tend to rate all monuments in the same way, regardless of their typology or location. French, Spanish and Italians, value their national symbolic monuments very similarly to the way foreigners who visit them do.

Key words: reviews, tripadvisor, tourism, monuments, language

Juan Pedro Mellinas
Departamento de Economía de la Empresa
Facultad de Empresa
Universidad Politécnica de Cartagena
Calle Real, 3
30201 Cartagena (Murcia)
Spain
Email: losmellinas@yahoo.es

Juan Pedro Mellinas is collaborating professor at Universidad Politécnica de Cartagena. His research focuses on travelers behavior, using as source of information online reviews in websites like Booking.com and TripAdvisor.

Acknowledgments

Introduction

Online reviews in the tourism sector have gained great importance in recent years, resulting in dozens of publications related with this topic. Some of them identify certain characteristics of the users that make them assign a higher or lower score to the places visited, such as gender, type of trip (business, pleasure) or group (couple, family, friends, ...). Among those are the studies that have identified variations depending on the nationality or language of the users, observing substantial differences in some cases.

However, most of these studies have focused on hotels and, although the overall sample sizes may be adequate, they do not get large samples of users (who use different languages) in individual hotels. The purpose of this study is to identify differences by country and language in 15 of the most visited attractions of the 3 countries that receive the most tourists in Europe (Spain, France and Italy). This allows to obtain a sample of more than one million reviews, with tens of thousands of reviews in each place studied.

Additionally, we analyze whether Spaniards, French and Italians value monuments of their countries better than foreigners, to try to identify find a certain sense of nationalism or national pride.

Literature Review

During the last years, several studies have been carried out in the tourist field using online reviews, analyzing both the global valuations assigned to each establishment, as more specific aspects. In this sense, it is useful to obtain detailed information about users' language, that allows a better understanding of cultural differences (Cantallops & Salvi, 2014). It is well known that customers' different cultural backgrounds and languages can produce differences in perceptions to products and services (Chen, Cheung, & Law, 2012).

Some authors has suggested that travellers geographic and psychic distance can influence online hotel ratings, detecting differences based on that attribute (Phillips, Antonio, de Almeida, & Nunes, 2019). It has been also suggested that the use of domestic language exerts a positive impact on hotel online ratings in Moscow and Rome (Mariani, Borghi, & Kazakov, 2019). It could be explained by the intangible nature of services, that influence how travellers perceive and evaluate tourist services (Holmqvist & Grönroos, 2012).

An extensive literature review about academic literature on cultural differences when evaluating hotels were compiled, identifying four areas (Schuckert, Liu, & Law, 2015):

- Difference in perceiving a service (perception difference)

- Different expectations on what a service be can or cannot delivered (expectation difference),

- Emphasis/demand of a service (emphasis difference),

- Complaint difference, where language and culture plays a role, too.

 Online reviews databases allow to filter reviews by language, which has allowed to carry out diverse investigations in this regard. Table 1 shows the articles that to date have quantified these differences using various databases in different countries.

Table 1. Academic literature about differences in scores by language.

Bibliography	Sample location	Languages analysed	Key Research Findings
(Schuckert et al., 2015)	Hotels in Hong Kong	English and Non-English	English speaking customers tend to give higher ratings than non-English speaking customers.
(Antonio, de Almeida, Nunes, Batista, & Ribeiro, 2018)	Hotels in Portugal	ENG, SPA and POR	Average rating for English reviews is 79.8 (out of 100). Portuguese 76.1 and and Spanish is 75.1.

(Liu, Teichert, Rossi, Li, & Hu, 2017)	Hotels in China	ENG, GER, FRE, ITA, POR, SPA, JAP, CHI, RUS	Chinese: 3.96, English: 4.03, French: 3.95, German: 4.13, Italian: 3.99, Japanese: 3.81, Spanish: 4.02; Portuguese: 4.11, Russian: 4.14
(Hale, 2016)	Tourist attractions in London	ENG, GER, FRE, ITA, POR, SPA, JAP, CHI, RUS and others	DIFFERENCE TO AVERAGE RATING: Japanese: -0.17, Portuguese: 0.17, Russian: 0.29, English: -0.07, Italian: 0.02, Chinese: -0.10, Spanish: -0.01, French: -0.05, German: 0.02
(Radojevic, Stanisic, Stanic, & Davidson, 2018)	Hotels in 80 capital cities.		Highly developed countries are more demanding as hotel customers. A 100% increase in the GDP (gross domestic product) of the guest's country is associated with a 0.03 decrease in average overall ratings.
(Pacheco, 2016)	Hotels in Porto, Portugal	ENG, POR, FRE, SPA and POR-BRA	English: 3.97, Spanish: 3.87, French: 3.94, Portuguese: 4.00, Brazilian: 4.20

Methodology

TripAdvisor is the world's largest travel platform and it is used by 490 million travellers each month. It has been used has the main source of information in research when online reviews are involved because of the 760 million reviews about accommodations, restaurants and other tourist spots and locations (Tripadvisor, 2019).

When carrying out the sample design, it is pretended to obtain places that have huge number of reviews in several languages, not only in the most popular ones. That is why we choose the places with the highest number of visitors in the three countries chosen (France, Italy and Spain). In the case of France, all places are in Paris and surroundings, but in the cases of Spain and Italy, we choose points of interest in different cities. We chose the same

nine languages analyzed in a previous study on hotels in China (Liu et al., 2017): English, Spanish, French, German, Portuguese, Italian, Russian, Japanese and Chinese.

We obtain a database of 1,132,141 reviews, ranging from 10,003 of the Santiago's Cathedral to 155,626 of Sagrada Familia. Moreover, in all cases, there are at least 100 reviews per language in each place, except for the 74 of Santiago's Cathedral in Russian. The languages with more reviews are Spanish (403,498) and English (401,488).

In the case of France and Italy, more than 95% of the reviews in French and Italian correspond to residents in those countries. With these data, language can be identified with nationality and cultural background in almost all cases. However, the population in Spain is around 46 million, while there are 400 million Spanish speakers in Latin America. This results in that for the places in Spain, about 80% of the reviews in Spanish come from locals, while in tourist spots of France and Italy, users from Spain are only 66% of those written in Spanish. For this reason, it is not appropriate to identify language with nationality and cultural background in a manner as clear as in the case of Italians and French.

In the case of other languages: Chinese, German, Japanese and Russian, the language is identified with the country in more than 90% of cases, so the identification language-nationality can be applied. However, in the case of Portuguese, we see how more than 80% of the reviews in that language come from Brazil (210 million population) and only 20% from Portugal (10 million population). The case of the English language is a great mix of countries of origin, without any reaching 50% of the reviews in that language.

Previous studies have identified trends in scores, which in certain nationalities tend to be higher or lower. We seek to find out if these differences exist and if they are similar using a differentiated sample. On the other hand, we test whether tourist spots are better or worse valued by locals. This could reveal feelings of nationalism, national pride or a better adaptation to local tastes.

Findings

In table 2 we can see the average score of each place of interest, together with the average score obtained in each language. Shaded in gray are the language scores that are higher than the average score of the rated place.

Table 2. Tourist spots scores

C.	City	Place	Total	English	Spanish	French	Italian	Portug.	German	Chinese	Japanese	Russian
FR	Paris	EIFEL TOWER	**4.597**	4.55	4.62	4.44	4.63	4.80	4.46	4.62	4.37	4.68
FR	Paris	LOUVRE MUS.	**4.594**	4.52	4.57	4.57	4.64	4.79	4.47	4.67	4.51	4.71
FR	Paris	NOTRE DAME	**4.615**	4.61	4.61	4.53	4.63	4.69	4.47	4.61	4.39	4.84
FR	Paris	VERSAILLES	**4.229**	4.15	4.27	4.00	4.11	4.70	3.98	4.32	4.35	4.41
FR	Paris	DISNEYLAND	**4.034**	3.95	4.07	3.88	4.30	4.45	3.61	4.09	4.02	4.48
SP	Sevilla	CATHEDRAL	**4.611**	4.57	4.61	4.60	4.62	4.72	4.56	4.59	4.53	4.86
SP	Granada	ALHAMBRA	**4.707**	4.68	4.73	4.66	4.70	4.82	4.59	4.63	4.69	4.69
SP	Barcelona	SAGRADA FAM.	**4.692**	4.73	4.72	4.57	4.66	4.76	4.56	4.79	4.66	4.77
SP	Santiago	CATHEDRAL	**4.649**	4.65	4.65	4.52	4.61	4.77	4.49	4.59	4.56	4.85
SP	Madrid	PRADO MUS.	**4.664**	4.63	4.72	4.45	4.55	4.67	4.57	4.59	4.50	4.81
IT	Rome	COLISEUM	**4.643**	4.64	4.66	4.51	4.73	4.77	4.48	4.58	4.45	4.67
IT	Pisa	TOWER	**4.472**	4.46	4.48	4.40	4.47	4.62	4.28	4.49	4.49	4.66
IT	Rome	VATICAN MUS.	**4.404**	4.25	4.27	4.20	4.44	4.65	4.00	4.60	4.59	4.66
IT	Florence	CATHEDRAL	**4.618**	4.60	4.63	4.45	4.64	4.72	4.42	4.67	4.64	4.90
IT	Naples	POMPEII	**4.565**	4.60	4.60	4.44	4.48	4.76	4.47	4.52	4.62	4.77
		TOTAL	**4.539**	**4.51**	**4.55**	**4.42**	**4.55**	**4.71**	**4.36**	**4.56**	**4.49**	**4.72**

We see how Spain has the highest scores in its monuments, with the 4 best rated in this country. We do not detect significant differences by type of place (Museum, Palace, Religious building, etc.). We also observe how language behaviors are repeated systematically; speakers of each language tend to rate all monuments in the same way, regardless of their typology or location.

The scores in the two most used languages (English and Spanish) are very similar, although slightly higher in Spanish. Italians also provide average scores very similar to those in English and Spanish.

The lowest scores are clearly those from German speakers (4.359) and French speakers (4.41), rating under the average in all cases. Differences are quite significant, being higher than one tenth in most of the cases, especially in Versailles and Vatican Museum. Portuguese and Russians speakers are the most enthusiastic with average values of 4.712 and 4.718, The differences are more than one tenth in most cases, reaching 4 tenths in the cases of Disneyland and Versailles.

Despite the apparent cultural differences with the rest of the sample, Chinese scores offer minimal variations from the average. The only exception is Vatican, where they assign almost two tenths more and Versailles that assign almost an additional point. Japanese value with a lower than average score, but without reaching the French and German figures. Curiously, they have in common with the Chinese the high relative score assigned to Versailles and Vatican.

Regarding the scores assigned by locals in Spain, France and Italy to their most emblematic tourist spots, the results do not show great differences when comparing with values provided by foreigners. French value all their attractions below the average (like in Italy and Spain), with 3 values above the tenth and the smallest difference in Louvre. In Spain, locals value similarly their emblematic places, with the only difference of Prado that values substantially higher than most, except Russians.

Italians value their monuments almost identically to the average marked by English and Spanish, with variations from the minimum average. The Coliseum is scored almost a tenth more than the average and Pompeii is valued almost a tenth less. Curiously, they value

all Spanish monuments below average (except Sevilla Cathedral) and all French above average (except Versailles).

All these places are well-known worldwide and highly valued. Despite this, there is a percentage of travelers who assign them the worst possible score (1 out of 5). It is possible that some of those travelers have had a very bad experience, due to queues, confusions with tickets, dirty facilities, etc. But assigning the minimum score to emblematic places like these, seems quite irrational.

In this case, the percentages of visitors that assign that minimum score, coincide with that of the countries that better or worse value. German and French have the highest percentages, while Portuguese has the minimum. Interestingly, Japanese, despite valuing most of the places with lower scores, has the second lowest percentage of "1" assigned. This suggests a certain degree of rationality, honesty and respect on the part of Japanese.

Conclusion

We observe that, as was the case with previous studies focused mainly on hotels, there are differences in the way of valuing depending on the language used. When analyzing places with huge sample sizes and very high average ratings, we found that the differences in languages coincide with those of previous studies in most cases, but surprisingly they differ in others.

We highlight the one made with Chinese hotels (Liu et al., 2017) that uses the same languages as the present study. The results are similar: Portuguese and Russian speakers assign the best scores, while English, Spanish, Chinese and Italian speakers assign an average and similar score between them. But it is surprising that, in this case, German speakers assign the second highest score, when in our study they assign the lowest score. It is also striking that Japanese assign the lowest score, below even French.

As for the studies conducted with Portuguese hotels (Antonio et al., 2018; Pacheco, 2016), we observe how English-speaking visitors tend to assign somewhat higher scores than Spanish-speaking visitors, contrary to what happens in our study. On the other hand, by distinguishing opinions in Portuguese between opinions of residents in Portugal and Brazil (something we do not do in our study), it is clear that residents in Portugal tend to score similarly to the average (Spanish and English) and Brazilians who tend to score higher. This makes us consider that the high scores of Portuguese speakers in our study come mainly from Brazilians.

Regarding the study of attractions in London (Hale, 2016), we again observe general similarities with our study, but also certain differences. The main difference comes from the scores of French and German speakers, since they are not the ones who score with the lower ratings. French value slightly below the average, while Germans even value above that average.

We observe how our results coincide with what was suggested in the great study mentioned worldwide (Radojevic et al., 2018), when pointed that reviewers of countries with higher GDP tend to value the places they visit worse. In our case, the fact that Russians and Brazilians value with the highest scores, while Germans and French give the lowest scores seems to coincide with previous findings in this regard.

Regarding the hypothesis about whether the inhabitants of each country value their own most emblematic tourist places better or worse, we must conclude that we do not see such behavior, which seemed to occur in the case of hotels in China (Liu et al., 2017; Schuckert et al., 2015) and attractions in London (Hale, 2016). French, Spanish and Italians, value their national symbolic monuments very similarly to the way foreigners who visit them do.

The fact that Germans and French so often grant the minimum score to emblematic monuments suggest a lack of rationality. It is not only that they are very demanding and do not assign a 5 to emblematic places very often, their low marks are unexpected in countries with such a level of educational and cultural development. On the contrary, Japanese, although they are also demanding tourists who assign the maximum score less often, show more rationality when evaluating, avoiding assigning that minimum score.

When French people visit French monuments, run by French people and served in their language, they are less satisfied than foreigners. French seem to have such high expectations and requirements that even they are unable to meet when they manage their own monuments. Therefore, we cannot argue that the way of presenting and managing tourist attractions has a cultural component and that, when there are large cultural differences, visitors from distant countries may feel less satisfied.

Previous studies mentioned the influence of geographic and psychic distance on online hotel ratings (Phillips et al., 2019), argued that the quality requirements of each culture could determine the scores (Schuckert et al., 2015) or suggested that the use of a known language favor the scores (Mariani et al., 2019). But this study seems to show that cultural distances or differences are not determining, observing GDP as the only explanatory cause. There are cultures that tend to rate tourism services better or worse, no matter if tourism spots are in their own country or abroad, places with religious component, museums or palaces.

Managers of all types of tourism services should take these results into account, being aware that there are certain nationalities that tend to score better than others. If Russians score better, it is not because the services are better suited to their specific needs, they just tend to score everything better. In the opposite case of low scores, it is not always related with adapt their services properly, some nationalities simply tend to assign lower scores. A hotel, a museum or a destination probably do not have any problem with Germans or French because

they are assigned lower scores than the rest, simply their culture implies that on a 1-5 scale they tend to assign lower scores.

These preliminary results, of a purely descriptive nature, are simply a first step in the measurement of differences by languages in the valuation of tourist spots. The next steps should be aimed at obtaining larger databases and introducing additional variables to design more elaborate statistical models, which will allow us to reach more precise conclusions. It would be interesting to expand the study with more countries and attractions. It would also be interesting to include places that do not obtain ratings as high as those in this study, to verify whether these trends are maintained in the language scores.

References

Antonio, N., de Almeida, A., Nunes, L., Batista, F., & Ribeiro, R. (2018). Hotel online
    reviews: Different languages, different opinions. *Information Technology & Tourism*,
    18(1), 157–185. https://doi.org/10.1007/s40558-018-0107-x

Cantallops, A. S., & Salvi, F. (2014). New consumer behavior: A review of research on
    eWOM and hotels. *International Journal of Hospitality Management*, *36*, 41–51.

Chen, R. X. Y., Cheung, C., & Law, R. (2012). A review of the literature on culture in hotel
    management research: What is the future? *International Journal of Hospitality
    Management*, *31*(1), 52–65. https://doi.org/10.1016/j.ijhm.2011.06.010

Hale, S. A. (2016). User reviews and language: How language influences ratings.
    *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in
    Computing Systems*, 1208–1214. ACM.

Holmqvist, J., & Grönroos, C. (2012). How Does Language Matter for Services? Challenges
    and Propositions for Service Research. *Journal of Service Research*, *15*(4), 430–442.
    https://doi.org/10.1177/1094670512441997

Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights:

Investigating language-specific drivers of hotel satisfaction with 412,784 user-

generated reviews. *Tourism Management*, *59*, 554–563.

https://doi.org/10.1016/j.tourman.2016.08.012

Mariani, M. M., Borghi, M., & Kazakov, S. (2019). The role of language in the online

evaluation of hospitality service encounters: An empirical study. *International*

Journal of Hospitality Management, *78*, 50–58.

https://doi.org/10.1016/j.ijhm.2018.11.012

Pacheco, L. (2016). An analysis of online reviews by language groups: The case of hotels in

Porto, Portugal. *European Journal of Tourism Research*, *14*, 66–74.

Phillips, P., Antonio, N., de Almeida, A., & Nunes, L. (2019). The Influence of Geographic

and Psychic Distance on Online Hotel Ratings. *Journal of Travel Research*,

0047287519858400. https://doi.org/10.1177/0047287519858400

Radojevic, T., Stanisic, N., Stanic, N., & Davidson, R. (2018). The effects of traveling for

business on customer satisfaction with hotel services. *Tourism Management*, *67*, 326–

341. https://doi.org/10.1016/j.tourman.2018.02.007

Schuckert, M., Liu, X., & Law, R. (2015). A segmentation of online reviews by language

groups: How English and non-English speakers rate hotels differently. *International*

Journal of Hospitality Management, *48*, 143–149.

https://doi.org/10.1016/j.ijhm.2014.12.007

Tripadvisor. (2019). Media Center. Retrieved from MediaRoom website:

https://tripadvisor.mediaroom.com/US-about-us