

Summer 6-30-2018

The Differences of Online Review Textual Content: A Cross-Cultural Empirical Study

Liyi Zhang

School of Information Management, Wuhan University, China

Li Li

School of Information Management, Wuhan University, China

Follow this and additional works at: <http://aisel.aisnet.org/whiceb2018>

Recommended Citation

Zhang, Liyi and Li, Li, "The Differences of Online Review Textual Content: A Cross-Cultural Empirical Study" (2018). *WHICEB 2018 Proceedings*. 66.

<http://aisel.aisnet.org/whiceb2018/66>

This material is brought to you by the Wuhan International Conference on e-Business at AIS Electronic Library (AISeL). It has been accepted for inclusion in WHICEB 2018 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

The Differences of Online Review Textual Content:

A Cross-Cultural Empirical Study

Liyi Zhang^{1*}, *Li Li*²

¹School of Information Management, Wuhan University, China

²School of Information Management, Wuhan University, China

Abstract: This research do a cross-cultural study by examine the differences of online review textual content between China, America and Australian. Through the online review text segmentation, classifying words by coding schema, calculating word proportion of each category, the research analyzes the differences of online review textual content from the aspects of textual type, content preference and textual emotion. The research finds that, cultural differences have significant effect on the online review textual type, Chinese customers prefer to describe objective facts while American & Australian customers prefer to describe subjective feelings; For textual emotion, Chinese customers prefer to express negative emotions while American & Australian customers prefer to express positive emotions. But cultural differences show no significant effect on the online review content preference.

Keywords: culture, online review textual content, textual type, content preference, textual emotion

1. INTRODUCTION

Online reviews are the important information resource for both merchants and consumers in the Internet era. With the accelerating process of globalization and the rapid development of cross-border E-commerce, many online reviews from online platforms are not limited to a single country. And especially for the online travel platforms, because tourism and hotel industry are more accessible to customers from different countries, online travel reviews are from more diverse sources. Different countries have different cultural characteristics, thus, cultural factors have become a noticeable difference in online reviews.

Online platforms present both the ratings and the review texts of products or services, although ratings can provide the overall evaluation, the textual content of review texts can provide more details, such as descriptions of specific function of products and the consumers' emotional attitude, these detailed descriptions help people get a more thorough understanding of products or services when they read online reviews. Compared to the online rating, therefore, online review textual content can provide better decision support, it's particularly important to consider textual content when studying the impact of culture on online reviews. In this research, the online reviews of China, America and Australia are selected as the research object, a comparative analysis of differences of online review textual content are conducted, in order to explore the impact of culture on online review textual content. The results will help both merchants and consumers better understand the online reviews and guide them to make decisions.

2. LITERATURE REVIEW

Some researchers have already put forward specific cultural theories during the study of culture, one of the most influential theories is the cultural dimension theory proposed by Hofstede. The theory points out that cultural dimensions include power distance, uncertainty avoidance, individualism / collectivism, masculinity / femininity, long-term orientation / short-term orientation, self-indulgence / restraint^[1]. Besides, considering the accuracy and clarity of the transmission and reception of information in the process of communication, Hall has

* Corresponding author. Email: lyzhang@whu.edu.cn (Liyi Zhang).

put forward the high context / low context theory, the theory points out that the expression of information is direct and unambiguous in the culture preferring low context communication, whereas in the culture preferring high context communication the expression of is usually vague and implicit ^[2].

Hofstede's culture dimension theory and Hall's high context / low context theory have been used to study the impact of culture on online behaviors and information systems, such as the impact of culture on online search process ^[3], the impact of culture on e-WOM ^[4], and the impact of culture on online promotion effect ^[5]. Meanwhile, there are also studies just select different countries to compare the differences of online behavior. For example, Sakarya and Soyer compare the differences of consumption values (utilitarian / hedonic) and online shopping behavior (online purchase frequency, goods purchased, online site type used and payment method) between Turks and Britons ^[6].

At present, among the studies about the culture impact on online users and information systems, the studies about the culture impact on review behaviors are still scarce. Lai et al. from the perspective of individualism / collectivism study the online review textual content between China and America, and point out that American reviews are more likely to express their own opinions on products and contained more recommendations ^[7]. Fang et al. based on the Hofstede's cultural dimension theory, compare the online review differences between China and America in the average number of reviews, average number of votes per review, average word counts per review, percentage of 1-star reviews and so on ^[8]. Hong et al. study the impact of culture (individualism / collectivism) on the tendency of consumers to conform to prior opinion and the emotionality of the review text, find that ratings from collectivists are closer to the average prior rating and collectivists express less emotions in reviews ^[9]. Zhou, Xia and Zhang exploit developed opinion mining techniques to study the online review textual content between China and America and point out that Americans express their opinions more directly and pay more attention to product details ^[10].

3. RESEARCH METHOD

3.1 Data

With the prevalence of cross-border tourism and the maturity of online tourism platform, the impact of culture on online tourism reviews has become increasingly prominent. Bangkok is one of the most famous tourist destinations among many tourist cities. Well-developed tourism and moderate consumption price attract tourists from all over the world to favor Bangkok. This research selects AGODA which is the well-known online hotel booking website in the Asia Pacific region, and crawls all the reviews of the Bangkok Hotel by the end of June 2015. For each review, the data includes the user name, the travel type, the user nationality, the review time, the review theme, the review of advantages, the whole review content, and the rating.

3.2 Measurement

3.2.1 Culture

The cultural differences between China and America have always been the focus of many researchers, and they find that there are obvious cultural differences between the two countries, for example, China is a typical collectivist country, and attaches great importance to the status of rights and prefer high-context communication. However, America is just the opposite. Countries which have the similar cultural characteristics with America include Australia, Canada and so on.

Nowadays, China and America are the world's most populous country with advanced information technology, which has the large number of netizens and extensive influence. This research still chooses China and America to represent culture. Considering the demographic factors that the number of Chinese reviews will

be much larger than the number of America reviews, Australia which has the similar cultural characteristics with America also be selected as the research object to balance sample sizes.

3.2.2 Textual content

In the existing research on textual content and semantic information, Ghose and Ipeiritis point out that textual content can be divided into subjective information and objective information from a stylistic point of view, and they have different effects on product sales and review usefulness^[11]. Berger and Milkman focus on the subjective information, studied the impact of emotional tendencies and specific emotions (anger, sadness, etc.) on social transmission, find that positive emotions, as well as specific high-arousal emotions, are more likely to enhance transmission^[12]. Ludwig, Ruyter, and Friedman study the semantic content and linguistic style properties of online reviews, find that the increase of positive affective content in the review text and the similar linguistic style between product interest group will enhance the online retail sites' conversion rates^[13].

In the field of online tourism, by analyzing textual content of online reviews, researchers can explore the most frequently used words in online reviews, and further combine with the rating data to find topics which affect customer satisfaction most^{[14][15]}.

Based on the above research, this research analyses online review textual content from three aspects: textual type, content preference, textual emotion. Textual type includes: the description focusing on objective facts or the description focusing on subjective feelings. Content preference includes: the description preferring tangible objects or the description preferring intangible services. Textual emotion includes: the description with positive emotions or the description with negative emotions.

3.3 Research model

$$\text{Online review textual type} = \alpha_1 * C + \alpha_2 * R + \sum a_i * T_i + \sum b_j * H_j + \sum c_k * M_k + \varepsilon \quad (1)$$

$$\text{Online review content preference} = \beta_1 * C + \beta_2 * R + \sum a_i * T_i + \sum b_j * H_j + \sum c_k * M_k + \varepsilon \quad (2)$$

$$\text{Online review textual emotion} = \gamma_1 * C + \gamma_2 * R + \sum a_i * T_i + \sum b_j * H_j + \sum c_k * M_k + \varepsilon \quad (3)$$

C represents culture, it's a dummy variable, 1 indicates china, and 0 indicates America or Australia.

R represents the rating per review, we control it for different ratings reflect different degree of customer satisfaction and will affect the emotional tendencies of online reviews.

T_i represents the tourism type, we control it for tourists with different tourism type will have different expectations for the quality of the hotel and services.

H represents the different hotel, we control it to reduce the impact of different hotel facilities and service levels on customers' perception.

M represents the time effect, the use of month as a time variable because of the possible of unobservable impact and trend on online reviews at different time periods (for example, the decline in hotel quality).

3.4 Textual content measurement index

The difficulty of this research is how to measure textual type, content preference and textual emotion of online review textual content, and how to avoid deviations caused by grammatical differences between Chinese and English. Based on the existing research method^{[9][10]}, and combined with research purpose, this research solves the problem from the lexical level, using the content analysis method to construct a word coding schema about textual type, content preference and textual emotion. Then through the online review text segmentation, classifying words by coding schema, calculating word proportion of each category, the data of online review textual content can be quantified.

Online review textual type: online review textual type includes the description focusing on objective facts or the description focusing on subjective feelings. This research calculates the number of words that describe the

products and services accounts for the proportion of review length (total number of words), and the number of words that describe emotion accounts for the proportion of review length (total number of words) to quantify textual type.

Online review content preference: online review content preference includes the description prefer tangible objects or the description prefer intangible services. Tangible objects are related to physical facilities, while intangible services are related to all services and quality of services. This research calculates the ratio of words related to tangible objects and the ratio of words related to intangible services to quantify textual type.

Online review textual emotion: on emotion measurement, positive and negative words that express subjective feelings, as well as positive and negative words describing products or services both can reflect the emotional tendency of customers. This research calculates the ratio of positive words (negative words) to review length (total number of words) to quantify textual emotion.

3.5 Coding schema

The words related to research purpose need to be filtered manually and simultaneously through an iterative process. Based on the existing research and research purpose, this research constructs a coding schema to ensure the rationality and effectiveness of the filter work (Figure 1).

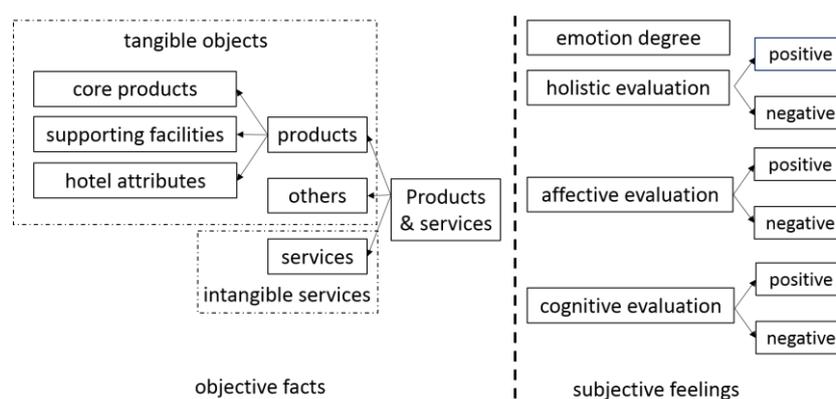


Figure 1. Coding schema

When describing products and services, core products refer to room, bed, bathroom and other core products of room. Supporting facilities refer to environment or facilities such as pool, breakfast, lobby. Hotel attributes include location, view, price, etc. And others refer to things that have nothing to do with the hotel, such as taxi, airport. Services refer to words such as service, check in, staff. When describing emotion, cognitive evaluation expresses the feeling of specific products and services, such as clean, delicious and friendly. Affective evaluation reflects the customer's mood, such as happy, like, enjoy and fun. Holistic evaluation includes good, perfect, excellent, etc. And the degree of emotion refers to the words such as very, too, extremely.

4. DATA PREPROCESSING & EMOTION CLASSIFICATION

In order to ensure data balance that each hotel has the same amount of data in China, America and Australia, this research finally chooses 7 hotels among all the hotels of Bangkok, and selects all Chinese, American, Australian reviews, then eliminates empty reviews and short reviews (such as: good!), finally gets 1073 reviews. There are 549 Chinese reviews, 254 American reviews and 270 Australian reviews respectively. After data acquisition, Chinese and English word segmentation software are utilized to extract words of all reviews and count word frequency, 3440 Chinese words and 3077 English words are got.

Any word that related to research purposes should be included, and some exceptions should be considered, such as stop words; and generic nouns such as “size”, “people”, “effort”, “fault”, etc. due to the lack of specificity, these words need to be filtered out.

Coding is primarily conducted by two researchers respectively to classify words according to the coding schema. When two researchers are in disagreement of any word classification, the third researchers joins and discusses together. Misspelling and English singular-plural pair should be correct during coding.

Coding begins from the high frequency word, and stops at the point when words appeared generic and irrelevant to the coding schema, resulting in 193 Chinese words and 195 English words. Then, coded words are marked in the original online reviews, words and word frequency involved in each review can be obtained. Through calculating word proportion of each category in coding schema, textual content can be quantified..

To judge the positive or negative emotional tendency of emotion words, researchers bring words into the context of the original online reviews to judge it artificially.

5. DATA ANALYSIS

The research calculates word proportion of each measurement index to quantify textual content, and table 1 shows descriptive analysis of measurement indexes. The proportion of words describing objective facts to review length (objective facts / review length), the proportion of words describing objective facts to total number of words (objective facts / total number of words), the proportion of words describing subjective feelings to review length (subjective feelings / review length), the proportion of words describing subjective feelings to total number of words (subjective feelings / total number of words) measure the online review textual type. And, “tangible objects / review length”, “tangible objects / total number of words”, “intangible services / review length”, “intangible services / total number of words” measure the online review content preference. And, “positive emotion / review length”, “positive emotion / total number of words”, “negative emotion / review length”, “negative emotion / total number of words” measure the online review textual emotion.

Table 1. Descriptive analysis

	China					America & Australian				
	N	Min.	Max.	Avg.	Std. error	N	Min.	Max.	Avg.	Std. error
objective facts / review length	549	0.00%	57.14%	19.62%	9.17%	524	0.00%	42.86%	13.69%	6.77%
objective facts / total number of words	549	0.00%	100.00%	58.39%	17.31%	524	0.00%	100.00%	51.15%	16.69%
subjective feelings / review length	549	0.00%	55.56%	14.79%	9.51%	524	0.00%	57.14%	13.52%	8.73%
subjective feelings / total number of words	549	0.00%	100.00%	41.43%	17.22%	524	0.00%	100.00%	47.37%	16.98%
tangible objects / review length	549	0.00%	54.55%	16.53%	8.61%	524	0.00%	42.86%	11.71%	6.54%
tangible objects / total number of words	549	0.00%	100.00%	49.43%	18.34%	524	0.00%	100.00%	43.91%	17.99%
intangible services / review length	549	0.00%	57.14%	3.09%	4.87%	524	0.00%	23.08%	1.98%	2.96%

5.1 Culture effect on textual type

When studying the culture effect by textual type measurement indexes, the result indicates that culture has a significant effect on textual types. Besides, the coefficient of “objective facts / review length” and “objective facts / total number of words” is positive, indicating that compared with American and Australian culture, Chinese culture is more descriptive of objective facts. And, the coefficient of “subjective feelings/ total number of words” is negative, indicating that Chinese culture has less description of emotion. But, the culture effect on “subjective feelings / review length” is not much significant, the reason may be that, coding stops at the point when words appear generic and irrelevant to the coding schema, but due to the diversity of emotion words, many low frequency emotion words are eliminated, so the analysis result has been disturbed.

Table 2. Culture effect on textual type

	objective facts /review length	objective facts/total number of words	subjective feelings /review length	subjective feelings /total number of words
constant	24.051**	85.835***	3.972	13.989
culture	6.113***	7.812***	1.042 ⁺	-6.473***
rating	-0.111	-3.039***	1.497***	3.071***
travel type	YES	YES	YES	YES
hotel	YES	YES	YES	YES
sample size	1073	1073	1073	1073
R square	0.206	0.188	0.145	0.17

Note: +<0.1, *p<0.05, **p<0.01, ***p<0.001

5.2 Culture effect on content preference

When using “intangible services / review length” to measure the content preference, the culture effect on content preference is significant, the positive coefficient indicates that compared with American and Australian culture, Chinese culture describes more intangible services in online reviews. But, when using “intangible services / total number of words” to measure the content preference, the culture effect is not significant. Besides, when using “tangible objects / review length” or “tangible objects / total number of words” to measure the content preference, although culture effect is significant, the positive coefficient indicates Chinese culture describes more tangible objects in online reviews. Thus, Chinese culture describes more of both tangible objects and intangible services. Cultural differences show no significant effect on the online review.

Table 3. Culture effect on content preference

	tangible objects / review length	tangible objects /total number of words	intangible services /review length	intangible services /total number of words
constant	23.042**	78.555***	1.008	7.281
culture	4.998***	6.39***	1.114***	1.422 ⁺
rating	0.03	-2.165***	-0.141	-0.874**
travel type	YES	YES	YES	YES
hotel	YES	YES	YES	YES
sample size	1073	1073	1073	1073
R square	0.176	0.141	0.131	0.11

Note: +<0.1, *p<0.05, **p<0.01, ***p<0.001

5.3 Culture effect on textual emotion

When using “positive emotion / review length” or “positive emotion / total number of words” to measure the positive textual emotion, the culture effect on positive textual emotion is significant, and the positive coefficient indicates that compared with American and Australian culture, Chinese culture expresses less positive emotion. When using “negative emotion / review length”, the culture effect on negative textual emotion is significant, and the negative coefficient indicates that Chinese culture expresses more negative emotions. Thus, for textual emotion, Chinese customers prefer to express negative emotions while American & Australian customers prefer to express positive emotions.

Table 4. Culture effect on textual emotion

	positive emotion / review length	positive emotion / total number of words	negative emotion / review length	negative emotion / total number of words
constant	-4.376	-14.673	11.862***	36.875***
culture	-1.697**	-11.189***	0.431**	0.417
rating	1.669***	4.703***	-0.515***	-2.205***
travel type	YES	YES	YES	YES
hotel	YES	YES	YES	YES
sample size	1073	1073	1073	1073
R square	0.175	0.291	0.228	0.237

Note: +<0.1, *p<0.05, **p<0.01, ***p<0.001

6. CONCLUSION & DISCUSSION

This research examines differences of online review textual content between different culture and select China, America and Australian as research object. The research finds that, cultural differences have significant effect on the online review textual type, Chinese customers prefer to describe objective facts while American & Australian customers prefer to describe subjective feelings; For textual emotion, Chinese customers prefer to express negative emotions while American & Australian customers prefer to express positive emotions. But cultural differences show no significant effect on the online review content preference.

Different from the existing research, this research obtains data from the highly diverse online review text, and especially studies the different characteristics of textual content. The method that calculating word proportion of each measurement index to quantify textual content overcomes the impact of grammatical differences to some extent.

The results show some guiding significance. On the one hand, Chinese customers focus on describing objective facts and can provide more online reviews on specific hotel products, whereas American and Australian customers are more concerned about personal feelings. Thus, for hotel managers, more attention should be paid to Chinese online reviews when taking steps to improve hotel facilities and management, the details will help to understand deficiency. Meanwhile, hotel staff should be required to concern more about the emotional state of American and Australian customers in the service contact and improve their emotional experience. For hotel customers, when referring to the online reviews, more information about specific conditions of hotel products will be revealed from Chinese customers' online reviews, and their reviews are more objective. On the other hand, because Chinese customers are more inclined to express negative emotions, hotel managers need to be more careful in dealing with Chinese's feedback and complaints, managers should respond promptly and properly so as to avoid further complaining in online reviews. Therefore, the impact of

cultural factors should be taken into consideration when referring to online reviews while making business marketing strategy or consumer consumption decision. This research selects China, America, Australia as research objects, moreover, countries similar to Chinese culture also include: Japan, Russia, Thailand, etc. Most of them are in Asia; and countries similar to American and Australia culture include: Canada, New Zealand, and other European countries, the results can be further applied to these countries.

There are also some limitations in this research. First, considering the sample balance of each hotel, only 7 hotels of 1073 online reviews are selected, under the background of extensive big data research and application, if the sample size can be further expanded, the results will be more accurate. Second, although most culture comparative researches have selected China and America as the research object, if more representative countries can be chosen, the results will be more objective. Finally, if natural language processing method can be used to identify emotions, the error will be reduced more effectively.

ACKNOWLEDGEMENT

This work is supported by the MOE Project of Key Research Institute of Humanities and Social Science of China under Grant 14JJD870002.

REFERENCES

- [1] Hofstede G. (1980). *Culture's Consequence*. Beverly Hills: SAGE Publications Inc, 25.
- [2] Hall E T. (1976), *Beyond Culture*. Garden City, New York: Anchor Press, 34.
- [3] Vuylsteke A, Wen Z, Baesens B, et al. (2010). Consumers' search for information on the internet: how and why china differs from western europe. *Journal of Interactive Marketing*, 24(4): 309-331.
- [4] Luo C, Wu J, Shi Y, et al. (2014). The effects of individualism–collectivism cultural orientation on eWOM information. *International Journal of Information Management*, 34(4): 446-456.
- [5] Phang C W, Tan C H, Sutanto J, et al. (2014). Leveraging O2O commerce for product promotion: an empirical investigation in mainland china. *IEEE Transactions on Engineering Management*, 61(4): 623-632.
- [6] Sakarya S, Soyer N. (2013). Cultural differences in online shopping behavior: Turkey and the United Kingdom. *International Journal of Electronic Commerce Studies*, 4 (2): 213-237.
- [7] Lai J, He P, Chou H M, et al. (2013). Impact of national culture on online consumer review behavior. *Global Journal of Business Research*, 7(1): 109-115
- [8] Fang H, Zhang J, Bao Y, et al. (2013). Towards effective online review systems in the Chinese context: A cross-cultural empirical study. *Electronic Commerce Research & Applications*, 12(3): 208-220.
- [9] Hong Y, Huang N, Burtch G, et al. (2016). Culture, conformity and emotional suppression in online reviews. *Journal of the Association for Information Systems*, Forthcoming(11).
- [10] Zhou Q, Xia R, Zhang C. (2016). Online shopping behavior study based on multi-granularity opinion mining: China versus America. *Cognitive Computation*, 8(4): 587-602.
- [11] Ghose A, Ipeirotis P G. (2011). Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Transactions on Knowledge & Data Engineering*, 23(10):1498-1512.
- [12] Berger J, Milkman K L. (2012). What makes online content viral? *Journal of Marketing Research*, 49 (2): 192-205.
- [13] Ludwig S, Ruyter K D, Friedman M, et al. (2013). More than words: the influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, 77(1): 87-103.
- [14] Berezina K, Bilgihan A, Cobanoglu C, et al. (2015). Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews. *Journal of Hospitality Marketing & Management*, 25(1): 1-24.
- [15] Xiang Z, Schwartz Z, Gerdes J H, et al. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44(4): 120–130.