THE LONG TAIL OF DESTINATION
IMAGE AND ONLINE MARKETING

Bing Pan, Ph.D.

Department of Hospitality and Tourism Management

School of Business and Economics

College of Charleston

66 George Street, Charleston, SC 29424 USA

Phone: 001-843-953-2025
Fax: 001-843-953-5697
Email: <bingpan@gmail.com>

Xiang (Robert) Li, Ph.D.

School of Hotel, Restaurant and Tourism Management

University of South Carolina

Columbia, South Carolina USA

Phone: 001-803-777-2764
Fax: 001-803-777-1224
Email: <robertli@mailbox.sc.edu>

Submitted exclusively to:

*Annals of Tourism Research*

DRAFT
THE LONG TAIL OF DESTINATION IMAGE AND ONLINE MARKETING

ABSTRACT

This study examines the linguistic structure of destination image using China as an example. The phrases the tourists use to describe China’s image follow the power-law distribution and exhibit the long tail pattern. The destination image is dominated by a few very popular phrases, but contains a large amount of phrases in small niches. Analysis on Google keyword search volumes shows that those phrases are likely to be the queries tourists use when searching destination information online. In addition, the tourists who use those niche phrases are more likely to travel to China. Thus, Destination Marketing Organizations should adopt the niche images as well as the commonly held images in their online marketing effort.

Keywords: Destination image; online marketing; search engine marketing; long tail; power-law distribution
INTRODUCTION

Tourism destination image (TDI), or the overall impression of one place (Bigne et al. 2001; Crompton 1979), is one of the most studied areas in the tourism literature (Stepchenkova and Morrison 2008). To date, empirical examination of TDI has generally employed either structured or unstructured approaches (Baloglu and Mangaloglu 2001; Echtner and Ritchie 1993; Selby and Morgan 1996). When analyzing and reporting results of TDI measurements, researchers generally focus on image items with the highest ratings in structured approaches, or the most frequently used words in unstructured approaches, as these represent the widespread and salient image held by tourists. Nevertheless, Destination Marketing Organizations (DMOs) often find such image results merely common sense (a typical comment mirrors the following, “I already know most customers think of us as so and so. Tell me something I don’t know!”). On the other hands, DMOs usually do not know what to do with the more obscure or “odd” images (for example, “Hmmm, why do I care that 15 customers are relating us with XXX?”). This paper attempts to investigate the linguistic structure of TDI and the connection between TDI phrases and online marketing, emphasizing the usefulness of those niche phrases.

Traditionally, most researchers have studied TDI from three perspectives: segmentation perspective, competitive analysis, and analysis of image components (Alcaniz et al. 2009; Gallarza et al. 2002). By and large, these studies serve to: (1) understand customer awareness and brand knowledge of a destination and evaluate a destination’s brand equity; (2) formulate product development and destination positioning strategies; and (3) assess the effectiveness of destination promotion efforts and predict tourists’ behavioral intention. In essence, most extant
TDI studies try to understand image as a brand component and utilize image profile as the basis for branding strategies (Cai 2002). However, search has become a dominant mode in tourists’ use of the Internet for trip planning (TIA, 2008). The importance of TDI partly lies in the fact that tourists search information and enter keywords in search engines based on their mental models, and mental models represent their understanding of the tourism information and their knowledge about the specific destinations (Pan and Fesenmaier, 2006). The phrases tourists use to describe their image of a destination may well be the keywords they use when searching that destination online. An information search perspective may bring new insights to TDI studies. Thus, understanding how TDI affects tourists’ information search provides opportunities for DMOs to apply their knowledge of TDI in search marketing efforts.

Therefore, the broad purpose of this study is twofold: to examine the linguistic structure of TDI; and to explore the importance of TDI phrases in online marketing. First, the utterance of any natural language generally follow power-law distribution (Pareto et al. 1971; Wikipedia 2009) which may exhibit long tail pattern (Anderson, 2004, 2006). This study first investigates if the distribution of destination phrases follows those patterns; second, the study correlates those image phrases expressed by tourists with search queries used in a search engine. Using an empirical study on American tourists’ image of China, this study demonstrates that TDIs held by tourists are rich and complex and image phrases generated through three open-ended questions follow the power-law distribution and exhibit the long tail characteristics. More importantly, empirical evidence demonstrate that many of those phrases are actually used by online users when looking for destination information via search engines; the aggregated volumes of niche phrases could surpass the few most popular phrases; potential tourists who use less common image phrases are more likely to travel to China.
THE LONG TAIL OF DESTINATION IMAGE AND SEARCH MARKETING

The notion of image has been widely used by marketing and behavioral science scholars to refer to people’s perception of a product, store, or corporate entity (Hampton et al. 1987; Jain and Etgar 1976; Spector 1961). Tourism researchers applied this idea to destination studies, and expanded the image definition to “include the perceptions or impressions a person has of a place” (McClinchey 1999, p. 9). Although tourism scholars have come up with numerous definitions of TDI (Li and Vogelsong 2006), most tend to agree that TDI is the overall impression of one place (Li et al. 2009). As Alcaniz and colleagues (2009, p. 716) put it, TDI “consists of all that the destination evokes in the individual; any idea, belief, feeling or attitude that tourists associate with the place.” Researchers have suggested that tourists’ overall TDI may be composed of two or three dimensions: TDI contains a cognitive and an affective component, with the former being the physical properties of a place, and the latter referring to the tourists’ feeling and evaluation of a destination. Some researchers (Gartner 1996; Pike and Ryan 2004) argue that there is a third dimension, conative image, which reflects the behavioral aspect (e.g., intention to visit) of one’s destination perception.

As indicated, the majority of TDI studies have used either structured or unstructured approaches (Baloglu and Mangaloglu 2001; Echtner and Ritchie 1993; Selby and Morgan 1996), but the former is much more widely employed than the latter (Alcaniz et al. 2009; Pike 2002). The structured approaches rate cognitive or affective image on a structured multi-attribute list via a semantic differential scale (Baloglu and Mangaloglu 2001). Typically, researchers have grouped these attributes into several dimensions by data reduction techniques such as factor analysis (Gallarza, et al., 2002). The battery of image attributes is generally pre-determined by the researchers based on literature review and the specific study context. Because most attribute
lists are developed idiosyncratically, the image items and dimensions tend to vary from study to study (Alcaniz et al. 2009). Attempts to generalize image dimensions and develop a standard image scale have been made (Echtner and Ritchie 1993). For instance, Gallarza, Saura, and Garcia (2002) review 25 empirical TDI studies that measured attribute-based images, and provide a list of the most commonly used image items. However, until now, there has been a lack of a universal image item list that is widely accepted and applicable across different destination types. Mossberg (1999) attributes the reasons to the use of factor analysis, free choice of naming each dimension, varying socio-cultural background of targeted image holders, and the way respondents were asked in image questions.

The structured approaches enjoy advantages such as flexibility, suitability for coding, and ease of analysis (Selby and Morgan 1996). However, TDIs are deemed as multiple, dynamic, and complex (Gallarza et al. 2002). Depending on the locale and scale, destinations in different places at different levels (country, state or province, or city) may demonstrate completely different characteristics. The predetermined destination attributes used in structured approaches may prevent respondents’ individualistic description of destinations (Baloglu and Mangaloglu 2001; Choi et al. 1999; Echtner and Ritchie 1993; Selby and Morgan 1996). The lack of subjective feedbacks could lead to difficulty in determining what exactly attracts tourists to their chosen destination (Coshall 2000; Embacher and Buttle 1989; Selby and Morgan 1996). Moreover, the typical structured approaches involve computing and reporting the mean values over all respondents. Underlying this practice is a somewhat questionable assumption that all respondents hold relatively homogeneous images. This has been deemed to be “inappropriate and can easily lead to false conclusions and subsequent mistakes in managerial decision-making” (Dolnicar and Grabler 2004, p.100). Hence, some researchers suggest that unstructured
approaches (e.g., open-ended questions) might provide richer data and reduce inherent bias and irrelevance of the structured approach (Stepchenkova and Morrison 2008). Nonetheless, when using unstructured methods, how to minimize potential semantic ambiguities and subjectivity in content analysis remains a methodological challenge (Stepchenkova and Morrison 2008). Fortunately, progress has been made in addressing this problem with the aid of software like CATPACII (Stepchenkova et al. 2009). Further, some researchers have attempted to combine both structured and unstructured methodologies to best capture different components of TDI (Baloglu and Mangaloglu 2001; Choi et al. 1999; Echtner and Ritchie 1993; Selby and Morgan 1996).

Notably, Echtner and Ritchie (1991; 1993) presented a particular conceptualization of TDI, which has been widely cited (Stepchenkova and Morrison 2008). Their three-dimensional framework of TDI incorporates three mutually overlapping continuums: (1) attribute-holistic continuum: TDI comprises both the perceptions of individual attributes of a destination and the holistic impression; (2) functional-psychological: each image component contains directly observable (functional) and abstract (psychological) characteristics; and (3) common-unique: TDI may range from common traits and features that can be used to compare all destinations to those unique to very few destinations. Note that the functional-psychological continuum is akin to the cognitive-affective typology of image in other TDI studies. Based on this conceptualization, Echtner and Ritchie (1991) proposed three open-ended questions to capture different components of TDI: (1) What images or characteristics come to mind when you think of XXX as a vacation destination?; (2) How would you describe the atmosphere or mood that you would expect to experience while visiting XXX?; and (3) Please list any distinctive or unique tourist attractions that you can think of in XXX.
To date, these three open-ended questions have been widely used in TDI studies. The first and second questions respectively capture the functional/holistic and psychological/holistic components of TDI, and were termed “stereotypical image” and “affective image” by Stepchenkova and Morrison (2008). The third item focuses on the unique aspect of TDI and was termed “uniqueness image” (Stepchenkova and Morrison 2008). As Alcaniz et al. (2009, p. 717) point out, although TDI studies have been frequently criticized for being atheoretical and lacking in conceptual frameworks, many researchers highlight Echtner and Ritchie (1993) as an exception, due to its conceptual and methodological rigor. The present study adopted Echtner and Ritchie (1991)’s three open-ended questions and Stepchenkova and Morrison (2008)’s terminology as the primary image measurement.

Power-Law Distribution of TDI Phrases and Long Tail Phenomenon  The first objective of this study is to explore the linguistic structure of TDI as self-reported by tourists. The answers to the three open-ended TDI questions are unique expressions of tourists’ understanding and impression of a destination. It is well-accepted that any natural language utterance follows the power-law distribution: e.g. in any corpus of natural language, the frequency of a word will be inversely proportional to its rank in frequency (Zipf 1949). Power-law distribution takes the form of $y = ax^k$ (Adamic and Huberman 2002). The distribution has been discovered in many unconstrained natural, social, or economic phenomena, such as sizes of cities (Gabaix 1999), magnitude and frequency of earthquakes (Abe and Suzuki 2004), Internet traffic and Web visits (Barabasi and Albert 1999), and the sizes of companies (Li 2002). However, answers to the three open-ended TDI questions represent are responses to specific survey questions and tend to be short phrases within a limited
topical scope. They are a special case of language utterance. It still needs to be verified whether or not those phrases will follow the power-law distribution.

The power-law distribution of products or market may present long tail characteristics and the 80-20 rule in businesses (Anderson 2006). The sales of nearly all businesses’ products follow the power-law distribution; that is, a small number of most popular products generate a large amount of revenue, while a large amount of niche products sell very little individually; the latter forms a “long tail” of niche products (Figure 1). This distribution could also exhibit the 80-20 rule whereby businesses make 80% of profit from 20% of top-selling products. Actually the 80-20 rule is a characteristic of a special case of the power-law distribution when the scale factor $k$ takes a value of around -1; when $k<-1$, 80-20 rule holds when the tail was partially cut off (Figure 1) (Pareto et al. 1971; Wikipedia 2009). Thus, the 80-20 rule and the long tail phenomenon are different manifestation of the power-law distribution. The former just stresses the importance of the “head” whereas the latter focuses on the significance of the “tail”. The 80-20 rule dominates in traditional businesses because they are restrained by the high storage and distribution costs; those businesses may intentionally cut off a large amount small niche products on the “tail” of the power-law graph due to unprofitability (Figure 1). For example, Wal-Mart only sells music CDs that have the potential for selling 100,000 to cover overhead. Thus the retailer ignores a large quantity of small-niche CDs due to unprofitability and limited shelf space (Anderson 2006). With the emergence of the Internet, online businesses, such as iTune or Amazon, could virtually distribute products at zero or very low cost, including those products intentionally ignored by traditional retailers. The aggregated volume of those niche products could amount to the majority of the revenue, even more than those popular products on the “head” of the power-law curve (Figure 1). Thus, the Internet actually helps businesses manifest the power-law distribution of their products fully online by taking
advantage of both the “head” and the “tail” of their product offering. This new phenomenon and business models was termed “the long tail phenomenon” (Anderson 2006).

Since the conception of the theory, some researchers have raised questions regarding the model and argued that in reality the technological shift may further strengthen the status of “superstars” (the products that sell a lot) which may account for even larger amount of revenue (Elberse 2008; Elberse and Oberholzer-Gee 2006). Elberse and her colleague studied home video sales for all the titles from 2000 and 2005 and discovered the co-existence of the long tail and the “superstar” phenomenon. On one hand, the number of videos that can only sell a few copies per week doubled, indicating a longer tail; on the other hand, the number of videos accounting for the top 10% of sales volume dropped by half, indicating an increasingly dominating role of the “head” (Elberse and Oberholzer-Gee 2006). The authors explained the “superstar” phenomenon by the following reasons: (1) increased product offerings online lead to more visible “head” products, since consumer would prefer the best products rather than inferior ones; and (2) the “head” products are more profitable in the online economy because they can reach an even larger audience and the digital copies cost little to reproduce. In essence, this observation echoed the well-known “winner-take-all” argument (Frank et al. 1995). In rebuttal, Anderson (2008) argued that he and Elberse differ in the definition of heads: Elberse and her co-author took the top 10% or 20% of products as the “head” while Anderson used the storage capacity for traditional sale channels such as Walmart as the “head”. Using the latter definition, the aggregates of the “tail” products will amount to a larger volume than the “head” (Anderson 2008). Thus, depending on the definition of “head”, one can calculate the different volumes of the “tail” or the “head”, and thus stress the importance of either the 80-20 rule or the “long tail phenomenon”. In reality, both of them are intrinsic characteristics of the power-law distribution under different conditions.
Online Information Search and Search Marketing

The second objective of the study is to connect destination image with online marketing. With the dominating role of the Internet as an information source, search engines have gained more prominence and became one of the major tools that tourists use to navigate through the vast online information space (TIA, 2008). Thus, it is crucial for DMOs or tourism businesses to gain an online presence on search engines. Search Engine Optimization is unpaid marketing for the purpose of gaining top positions in search engines by tweaking the content of a Web site or conducting linking campaigns. In contrast, Search Engine Marketing is paid marketing in the form of bidding on advertising space on search engines using programs such as Google AdWords (Pan et al. 2007). Both efforts require targeting at specific keywords or phrases because different phrases or keywords represent different interests, goals, or tasks of users. Some queries have a large volume of searches, representing a large market, while others have a smaller volume, representing niche markets. However, the former will undoubtedly attract more business competitors than the latter, since more Web sites will be optimized toward popular queries and more businesses might bid on those through paid listing programs (Pan et al. 2007). Researchers have realized the importance of the niche products to small businesses and its implications for a new generation of tourists who are seeking individual experiences (Lew 2008; Xiang et al. 2009). In the TDI literature and practices, it seems that most past studies on TDI have focused more on popular and salient image items (the “head”) of a destination and somehow ignored small niche images (the “tail”). Yet the Internet may help connect a destination with tourists around the world who are interested in that destination’s niche products; each niche might be small individually, but combined they may amount to a large volume (Lew 2008). When manifested online, those niche tourists who use less common keywords and phrases to search for a long-haul destination may know more about
it and express very specific information needs; therefore they might be more likely to travel to the destination. Thus, the understanding of the niche search queries that potential tourists use is crucial in marketing a destination or business on search engines and on the Internet in general.

Studies have shown that tourists generate keywords for searching a specific destination based on their understanding and experience of the destination, their computer and search expertise, and their understanding of available tourist information online (Pan and Fesenmaier, 2006). All those elements constitute tourists’ mental models when planning a trip. TDI is the total impression of a place (Bigne et al. 2001; Crompton 1979), which involves the knowledge and understanding of, and experience with the destination. Thus, TDI is part of tourists’ mental model when they search online. Hence, the authors postulate that the phrases tourists use to describe a specific TDI are likely to be the queries that tourists type in search engines when searching for information. However, due to the different nature and goal of the communication processes, i.e., one being answering a survey question and the other being the phrases entered for information search, whether TDI phrases and search keywords are correlated remains to be substantiated.

In summary, the purpose of this study was to examine the distribution of TDI phrases, and to explore the connection of TDI studies and extant research on search marketing. More specifically, this study attempts to achieve these goals by focusing on the following questions: (1) Do the phrases tourists use to describe TDI follow the power-law distribution? (2) Are TDI phrases used as keywords in information search for a destination? and, (3) Are those tourists who use the niche phrases more likely to travel to the destination? To answer these questions, the following section describes the details of the research methods.
Study Methods

This study adopts two different data sources. The first one is an online panel survey and the second involves Google search volumes for certain keywords as reported by Google AdWords Keyword Tool (Google 2009a). The online survey was a large-scale marketing research project on American tourists’ perceptions of mainland China as a vacation destination. China was used for this study for two reasons: (1) China is becoming increasingly active in tourism promotion and many international affairs, which leads to reasonable destination awareness among respondents; (2) As one of the world’s top destination country, China provides a wide array of attractions which allows sufficient variation in image held by potential respondents (Li, Pan, Zhang, and Smith, 2009). The survey targeted the American leisure travel population, operationalized as American adults, age 18 or older, who have taken at least one leisure vacation in the past 12 months. To reach a representative sample of this population, an American online survey company sent the survey invitations to its national and demographically representative panel. The outgoing survey was demographically balanced to represent the American travel population as reported by the Travel Industry Association (TIA) (TIA, 2007). Key measures for population balancing include region, age, household income, and gender (TIA 2007; U.S. Census Bureau 2005).

The online survey was conducted during May and June 2008, and the final analysis includes 3,263 valid and complete responses. Echtner and Ritchie (1993)’s three open-ended questions are used to measure respondents’ perception of China. Each respondent was allowed to provide up to three, but at least one answer to each question, and they were not allowed to skip any of the three questions. The survey also asks about respondents’ likelihood of taking a leisure vacation to mainland China in the next five years. Other questions of the survey ask about
respondents’ international and Asia-related travel experiences, impression of different Asian
countries, and demographic information.

In order to examine the linkage of TDI phrases with search engine queries, the second
data set was the actual search volumes on Google for uniqueness image phrases, since online
searches for destinations are mostly functional (Pan et al. 2007). Google AdWords keyword Tool
(Google 2009a; Joshi and Motwani 2006) intends to help Google advertisers find appropriate
keywords to bid on, so it gives the search volumes of search keywords for the current month and
also the average for the past 12 months, based on actual user queries in English and from the
United States (Google 2009a). In this study, the average volumes of those search phrases from
Google in the past 12 months (July, 2008 to July, 2009) were retrieved from the Keyword Tool
on August 4, 2009.

Data Cleaning and Data Analysis. The three types of image phrases went through the
data cleaning process, in which clear typos were corrected and phrases with exact meaning were
merged (such as merging “the Great Wall” and “Great Wall”). There are three stages in the data
analysis: (1) the researchers used text analysis software CATPACII to calculate the frequencies
of different phrases used in the responses (Woelfel and Stoyanoff, 1993). The frequencies of
different phrases were ranked from most to least used ones; the top, middle, and bottom phrases
were identified. In addition, the authors look at the percentage of respondents who used at least
one of the top phrases. Also the distribution of phrases was modeled on the power-law
distribution using graphing and linear regression on the logarithmic transformation of phrase
frequencies and their ranks; (2) the frequencies of those phrases appearing in the survey were
correlated with corresponding search volumes on Google; (3) the distinctness (e.g. uniqueness)
of the phrases used by the respondents was correlated with their travel intention. The goal of this
The first step was to demonstrate the importance of niche keywords in search engine marketing, so as to establish the importance of niche phrases for TDI.

Results

This section details the results from the three stages of analyses. Table 1 lists the profile of the 3,263 respondents. More than 70% of respondents had traveled outside the United States before, and nearly two-thirds (64.1%) of respondents have taken at least one leisure vacation outside of the United States, which indicates they are more active tourists than the general American adult population. This comes as no surprise as respondents to this survey represent the currently active American leisure travel population.

_____________ Insert Table 1 Here _______________

The Power-Law Distribution of China’s Destination Image  The frequencies of the image phrases are tallied and the distribution of those phrases is modeled in the first stage of analysis. In total, there are 1,641 distinct and 8,082 total phrases describing the stereotypical images of China, while there are 1,001 distinct and 7,851 total phrases for affective image and 463 and 5,775 for uniqueness image (Table 2). In Figure 2, the x axis is the rank of the phrases (i.e., the most popular phrase ranked 1, the second 2, etc.) and the y axis represents their corresponding frequencies. All TDI phrases are dominated by a few phrases on the top, but contain a large amount of small niches. More specifically, for stereotypical image, “the Great Wall” (740 times), “Crowded” (396), and “Food” (367), are the top ones, followed by “Culture” (196), “Communism” (163), and “People” (141). However, there are 1,166 phrases that only appear once, and 149 phrases appear twice. Affective image is dominated by phrases such as
“Exciting” (664), “Happy” (592), and “Busy” (446). For uniqueness image phrases, “The Great Wall” and “Beijing” each appear more than 1,000 times, followed by “Hong Kong” and “Shanghai” both of which appear more than 300 times. Notably, some niche image phrases are erroneous or unrelated, such as “HooHingMing City” (the correct spelling is “Ho Chi Minh City”, a Vietnamese city) and “Nintendo” (the name of a Japanese company); but most are interesting niche concepts/phrases, such as “Gansu”, “Yin Yang” and “Monasteries”. Table 2 lists the top 10 most frequent, 10 least frequent, and 10 median frequent (by rank) phrases.

Graphing and regression are performed to predict phrase frequencies from their ranks with power-law distribution. The result indicates that the three types of image phrases fit the distribution nicely. Figure 2 includes three plots of logarithmic transformations of the phrases’ rank and frequency. The straight-line shape of the three graphs is a clear indication of the power-law distribution (Adamic and Huberman 2002). The modeling of the distribution of the three image phrases showed highly significant results: Stereotypical image, $y = 1213.389 x^{-1.016}$ (R Square = 0.96, p<0.001); affective image, $y = 1280.269 x^{-1.042}$ (R Square = 0.96, p<0.001); and Uniqueness image, $y = 2197.860 x^{-1.225}$ (R Square = 0.99, p<0.001). The k values for stereotypical and affective image phrases are very close to -1, indicating that those two distributions naturally follow the 80-20 rule.

Table 3 lists the top phrases by frequency and their coverage on the total volumes of phrases and the number of the respondents. For example, for the fifth row in the first table, the top 328 phrases by frequency constitute 20% of the total number of unique phrases; they
contributed to 80.6% of the total volume of all phrases; 97.4% of respondents used at least one of those 328 phrases. Interestingly, for stereotypical and affective image phrases, the top 20% of unique phrases contained a little more than 80% of total phrase volume, indicating the 80-20 rule. For the phrases about uniqueness image, the top 20% of phrases accounted for more than 90% of volume, indicating an even more concentrated “head” in the distribution (Table 3). In addition, 80% of respondents used one of top 20 phrases of stereotypical image and affective image phrases, while more than 80% of respondents used one of the top two phrases in uniqueness phrases (“the Great Wall” and “Beijing”). If one defines the top 20% phrases as the “head”, the numbers of phrases vary: 328 phrases for stereotypical image, 200 phrases for affective image, but only two phrases for uniqueness image. Thus, there exists the long tail of the image phrases, but the aggregated volumes of 80% of distinct tail phrases are not large enough to surpass the top 20% most popular distinct phrases. However, one can hardly promote more than top the 200 concepts associated with a destination. Traditionally, marketers promote a destination with far fewer concepts. On the other hand, if one takes the top 10 phrases as the “head”, the total volume of the “head” phrases only covers around 30%, 38%, and 73% of all phrase volumes of stereotypical, affective, and uniqueness image, respectively. Thus, in the first two cases, the “tail” is significant enough to surpass the volume of the “head.” In the distribution of uniqueness image phrases, the “head” is a lot fatter and the “tail” is a lot slender than other image types. This indicates high concentration on the popular attractions, especially “the Great Wall” and “Beijing” (Table 2).

__________________ Insert Table 3 Here____________________

_Online Google Searches and Uniqueness Image Items_ In the second stage, the frequencies of the uniqueness image phrases was correlated with search volumes on Google as
reported by Google AdWords Keyword Tool (Google 2009a; Joshi and Motwani 2006). First, the survey itself provided the context information about China: for example, respondents typed in “historic sites” in the survey form but implicitly they meant “historic sites China”. Thus, if “China” or “Chinese” were not already in the phrase, the authors added the keyword “China” to the phrases. Those revised uniqueness phrases are entered in the Tool to extract the average monthly search volume. Google Keyword Tool reported search volumes of 286 of the 463 uniqueness image phrases (with the lowest frequency being 16). The tool reports “Not Enough Data” on the rest 179 phrases. However, that does not mean that those phrases did not have any search traffic. Google explains: “If a keyword is uncommon or very specific, there may not be a substantial amount of search information about it” (Google 2009b, bold added). Those phrases are given a value of 0 as their search volumes in the correlation analysis between the frequencies and their corresponding search volumes. Table 4 listed the top, middle, and bottom most used phrases in the survey and their corresponding Google search volumes. The correlation analysis of the frequencies of those unique image phrases with their corresponding search volumes shows a Pearson’s correlation value of 0.652 with a p value less than 1.8 x 10^{-57} (n=463). Furthermore, for those 179 phrases that Google Keyword Tool reported “Not Enough Data”, 164 of them appeared only once among all survey results (Table 4). The results clearly indicate that those very niche image phrases are also likely to be used very infrequently in Google searches. In conclusion, most TDI phrases about China were used in searches on Google, and TDI descriptions and search queries are indeed correlated.

___________________ Insert Table 4 Here___________________
Correlation of Distinctness of Phrases and Travel Intention  

This analysis explores the correlation between those respondents using the niche phrases and their intention to travel to China. The distinctiveness of phrases was defined as the reciprocal of the probability of a phrase to appear in a corpus of language utterance, which is computed as the reciprocal of the proportion of its frequency to the number of total phrases in a text corpus (Jurafsky et al. 2000). For example, in a body of text with 1,000 phrases, a phrase appeared 100 times in the text will have a distinctiveness value of $1/(100/1000) = 10$.

The distinctiveness value for each phrase was calculated and the average distinctiveness values of all three phrases of each of the three image questions were assigned to each respondent. First, the univariate Analysis of Variance model in SPSS was used to determine the significant demographic and travel experience variables that may influence respondents’ intention to visit China. The results showed that five factors (i.e., ethnicity, gender, age, number of world areas traveled, and whether this respondent visited China before) significantly affected on one’s travel intention. These five factors were entered in the regression model as control variables. A following multi-regression analysis was then used to find out the significance of average uniqueness values for each respondent to travel intention incorporating the effects of the control variables (Table 5). The results indicated that distinctiveness values of uniqueness image had a significant impact on respondents’ intention to travel ($p < 0.001$). The respondents who used more distinctive uniqueness image phrases were more likely to have greater intention to travel to China ($\beta = 0.70$). This finding makes intuitive sense: given China as an exotic destination for most American citizens, those respondents mentioning unique attraction names tend to know more about China and are more likely to travel to China in the near future. Thus, the niche phrases of TDI might be more effective than those popular and commonly used phrases when
marketing online. Put differently, the niche market as represented by those tourists who used niche phrases in information searches, are more likely to convert.

CONCLUSION

This study reports that the destination image items American travelers used to describe China do follow the power-law distribution: a few phrases or attractions are well known to many of the respondents; hundreds of niche phrases are used very few times individually, but collectively they account for a large volume. The distribution of stereotypical and affective image phrases follow both the 80-20 rule and the Long Tail pattern, if one defines the “head” as the top 10 phrases in the latter case. The distribution of uniqueness image phrases was more concentrated on the “head” because the top 10 phrases covered more than 73 percent of total volume.

The primary theoretical contributions of the present study lie in the linguistic approach to TDI research and the linking of linguistic structure of TDI with search marketing. Unlike most previous studies, which approached TDIs as either pictorial representations (Day et al. 2002; MacKay and Fesenmaier 1997) or symbolic manifestation of a destination (Echtner and Ritchie 1993; White 2005), the present study focused on the linguistic portrayal of TDI. More specifically, most extant TDI studies tend to focus on the most commonly held images of a destination (Dann 1996; Pike 2002). This study shows that the TDI phrases are actually very rich and diverse, and those “tail” image items deserve more research efforts. Notably, there is not always an apparent cutoff point which divides popular image phrases from niche ones; depending on the marketing purpose, the choice of top attributes is rather arbitrary. Thus, all image phrases demand research attentions. In addition, this study links the understanding of TDI
to research on search engine queries for destination information. Findings from the study show that many uniqueness image phrases generated through TDI studies are the same phrases that tourists type in search engines for a destination. The destination image - search keywords link is of critical importance to both TDI studies and search engine marketing. The authors hope this new conceptual perspective may lead to a new stream of research.

The general managerial contribution lies in the validation of the importance of niche products and niche market in the Internet age. Different “head” and “tail” sections of image phrases might be suitable for different marketing channels. The most popular (e.g., top 20%) phrases are very important and represent the majority of tourists; however, it is unlikely that all those image attributes could be promoted effectively. To avoid diluting a brand’s identity or sending confusing brand messages, the positioning literature traditionally suggests destinations to focus on several key themes in marketing efforts in mass media channels (Calantone et al. 1989; Chen and Uysal 2002; Ries and Trout 1986). This classic strategy accomplishes effectiveness by essentially compromising niche markets to more mainstream market. The present study argues that such compromise might no longer be necessary in today’s environment. For TDI researchers, this calls for more attention to be given to those uncommon even obscure destination images. This study also found that holders of the “tail” images are not only more knowledgeable about a destination, but also more likely to visit the destination. For today’s DMOs, it is becoming necessary and cost-effective on the Internet to capture the “niche” image held by only a few tourists. A new segmentation approach might be employed based on the distinctiveness of phrases the tourists type in. Indeed, the latest technology allows a website to track precisely the keywords the tourists type in search engines before they reach the home page; a highly customizable website could deliver different content and marketing messages according
to the specific keywords used. The one-on-one marketing is not only feasible but also highly achievable (Peppers et al. 1999). Thus, one can take full advantages of the aggregated niche market by targeting all of them. Furthermore, destination Web sites should not only tell customers what they already know, they should provide novel and exciting information beyond customers’ initial mental models. Adding more niche information may create positive and enticing surprises and higher satisfaction in trip planning (Pan and Fesenmaier 2006). In addition, providing more niche attractions and unique characteristics can also help alleviate the congestions in popular attractions and implicitly direct tourists to less visited areas. In conclusion, taking advantage of the long tail nature of TDI could allow destination marketers to conduct cost-effective online marketing campaigns if segmented and marketed correctly; it can also possibly be beneficial to the sustainable development of a destination on balancing tourist flows. Targeting niche market online represents the paradigm change for online marketing and electronic commerce afforded by the Internet (Hoffman and Novak 1997).

The specific managerial contribution is on offering strategic implications for destination marketing of China. United States is currently China’s No. 1 long-haul origin market (China National Tourism Administration 2010). When asked to name any distinctive tourist attractions (i.e., their “uniqueness image”) in China, respondents of this study generated far less responses. Further, their responses were dominated by top two phrases “the Great Wall” and “Beijing.” The two terms contributed to almost half of the phrase volume; and about 85 percent of respondents used at least one of these two. This rather monotonous image may be due to unfamiliarity of most American respondents to attractions in China, or a lack of information diversity in China’s marketing communications with the American market. Thus, China’s destination marketers may need to revisit some of their strategies. When marketing China to U.S. tourists on traditional
mass media channels, one can focus on the top few image items such as the iconic image (Litvin and Mouri 2009) due to the limited capacity and higher cost; but when marketing on the Internet, focusing on the long tails of phrases could be as valuable as or even more cost-effective than the popular ones. For example, popular destinations or attractions such as “the Great Wall” and “Beijing” always have high levels of competition among all the marketers of China, making them more expensive to target (Pan, Litvin, and O'Donnell, 2008). Targeting the niche phrases could help expand the product offerings to less popular destinations or attractions (Lew 2008).

This study is limited in that it focuses on China as a tourist destination. The novelty of the place and the U.S. respondents’ lack of familiarity might potentially influence the concentrated “head” of uniqueness image phrase distribution. Future studies could focus on exploring the distribution graph of other destination TDI in different geographic levels, such as states, provinces or cities. Exploring the correlation of actual tourist volumes with their major TDI image phrases will be a further step in demonstrating the importance of a long tail to the lifeline of a destination. Furthermore, how some respondents developed their “tail” images in the first place remains a fascinating question. From an information search perspective, these potential customers probably know more about a product because they either have higher level of involvement (Havitz and Dimanche 1999) or are at a more advanced stage of the purchase funnel than others (Barry 1987). Or, some respondents may simply have a more sophisticated “baseline image” (Li et al. 2009) because of personal interests, education, or cultural or ethnic background. Follow-up research on this group of customer may shed new light on the understanding of information search and image formation.

In conclusion, although popular or commonly held images are good for marketing a destination on mass media channels, destination marketers should capture, mine, and understand
their TDI in many niches for the purpose of online marketing. This will also help DMO marketers to expand their potential customer base, take on one-to-one marketing efforts, and also protect their most popular and crowded destinations and attractions. With the development of more refined Internet technology in which personal data of each customer or potential customer could be captured, analyzed, and distributed, this more refined approach to understand TDI is becoming more crucial for online marketing in general and search engine marketing in particular.
REFERENCES

Abe, S., and N. Suzuki
Adamic, L., and B. Huberman
2002 Zipf’s law and the Internet. Glottometrics 3:143-150.
Alcaniz, E.B., I.S. Garcia, and S.S. Blas
Anderson, C.
Anderson, C.
2008 Debating the Long Tail. In Conversation Starter.
Baloglu, S., and M. Mangaloglu
Barabasi, A., and R. Albert
Barry, T.
Bigne, J.E., M.I. Sanchez, and J. Sanchez
Cai, L.
Calantone, R., C. Di Benedetto, A. Hakam, and D. Bojanic
Chen, J., and M. Uysal
China National Tourism Administration
2010 Inbound tourist arrivals from January to December, 2009.
Choi, W.M., A. Chan, and J. Wu
Coshall, J.T.
Crompton, J.L.
Dann, G.
Day, J., S. Skidmore, and T. Koller

Dolnicar, S., and K. Grabler
Echtner, C., and B. Ritchie
Echtner, C.M., and J.R. Ritchie
Elberse, A.
Elberse, A., and F. Oberholzer-Gee
Embacher, J., and F. Buttle
Frank, R., P. Cook, and S. Rosen
Gabaix, X.
Gallarza, M.G., I.G. Saura, and H.C. Garcia
Gartner, W.C.
Google
Google
2009b Why do I see 'No data' or 'Not enough data' in my Keyword Tool results?
Hampton, R.D., B.S. Guy, and J.M. Sinkula
Hoffman, D., and T. Novak
Jain, A.K., and M. Etgar
Joshi, A., and R. Motwani
Jurafsky, D., J. Martin, and A. Kehler
2000 Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition: MIT Press.
Lew, A.

Li, W.

Li, X., B. Pan, L.G. Zhang, and W.W. Smith

Li, X., and H. Vogelsong

Litvin, S., and N. Mouri

MacKay, K., and D.R. Fesenmaier

McClinchey, K.A.
1999 Rural images, tourism and sustainability: Perceptions of rural accommodation operators and their visitors in Waterloo-Wellington Region, Ontario Wilfrid Laurier University.

Mossberg, L.L.

Pan, B., and D.R. Fesenmaier

Pan, B., S.W. Litvin, and T.E. O'Donnell

Pareto, V., A. Schwier, and A. Page

Peppers, D., M. Rogers, and B. Dorf

Pike, S.

Pike, S., and C. Ryan

Ries, A., and J. Trout

Selby, M., and N.J. Morgan

Spector, A.J.
Stepchenkova, S., A.P. Kirilenko, and A.M. Morrison
Stepchenkova, S., and A.M. Morrison
TIA
U.S. Census Bureau
2005 United State Census Bureau Web Site.
White, C.
2005 Destination image: to see or not to see? Part II. Management 17:191-196.
Wikipedia
2009 Pareto principle.
Xiang, Z., B. Pan, and D. Fesenmaier
Zipf, G.
Table 1. Demographic Profile of Survey Respondents

<table>
<thead>
<tr>
<th>Age (n=3261)</th>
<th>Ethnicity (n=3211)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-24 years</td>
<td>White 83.2%</td>
</tr>
<tr>
<td>25-34 years</td>
<td>Black 5.5%</td>
</tr>
<tr>
<td>35-44 years</td>
<td>Latino/Hispanic 3.5%</td>
</tr>
<tr>
<td>45-54 years</td>
<td>Asian/Pacific Islander 6.0%</td>
</tr>
<tr>
<td>55-64 years</td>
<td>Other 1.7%</td>
</tr>
<tr>
<td>65 and over</td>
<td></td>
</tr>
<tr>
<td>Average Age</td>
<td>Household Income (n=2880)</td>
</tr>
<tr>
<td></td>
<td>&lt;$25,000 13.1%</td>
</tr>
<tr>
<td></td>
<td>$25,000-$49,999 31.5%</td>
</tr>
<tr>
<td></td>
<td>$50,000-$74,999 23.8%</td>
</tr>
<tr>
<td></td>
<td>$75,000-$99,999 15.2%</td>
</tr>
<tr>
<td></td>
<td>$100,000 + 16.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender (n=3263)</th>
<th>Household Asset (n=2474)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male 48.8%</td>
<td>Less than $49,999 38.6%</td>
</tr>
<tr>
<td>Female 51.2%</td>
<td>$50,000-$99,999 14.2%</td>
</tr>
<tr>
<td></td>
<td>$100,000-$199,999 17.5%</td>
</tr>
<tr>
<td></td>
<td>$200,000-$299,999 8.6%</td>
</tr>
<tr>
<td></td>
<td>$300,000-$399,999 5.5%</td>
</tr>
<tr>
<td></td>
<td>$400,000+ 15.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education (n=3241)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Attended grade school 0.1%</td>
<td></td>
</tr>
<tr>
<td>Graduated grade school 1.5%</td>
<td>Less than $49,999 38.6%</td>
</tr>
<tr>
<td>Attended high school 1.2%</td>
<td>$50,000-$99,999 14.2%</td>
</tr>
<tr>
<td>Graduated high school 15.1%</td>
<td>$100,000-$199,999 17.5%</td>
</tr>
<tr>
<td>Attended college 32.1%</td>
<td>$200,000-$299,999 8.6%</td>
</tr>
<tr>
<td>Graduated college 34.8%</td>
<td>$300,000-$399,999 5.5%</td>
</tr>
<tr>
<td>College post-graduate 15.1%</td>
<td>$400,000+ 15.6%</td>
</tr>
</tbody>
</table>
Table 2. TDI Phrases for China

<table>
<thead>
<tr>
<th>Stereotypical Image Phrases</th>
<th>Top 10</th>
<th>Freq.</th>
<th>Median 10</th>
<th>Freq. last 10</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Wall</td>
<td>740</td>
<td>Antiquity</td>
<td>3</td>
<td>Worried about Safety</td>
<td>1</td>
</tr>
<tr>
<td>Worried about The</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crowded</td>
<td>396</td>
<td>Asian Culture</td>
<td>3</td>
<td>Cleanliness of the Food</td>
<td>1</td>
</tr>
<tr>
<td>Food</td>
<td>367</td>
<td>Asian Food</td>
<td>3</td>
<td>Writings</td>
<td>1</td>
</tr>
<tr>
<td>Culture</td>
<td>196</td>
<td>Asian People</td>
<td>3</td>
<td>Xenophobic</td>
<td>1</td>
</tr>
<tr>
<td>Communism</td>
<td>163</td>
<td>Atmosphere</td>
<td>3</td>
<td>Xian</td>
<td>1</td>
</tr>
<tr>
<td>People</td>
<td>141</td>
<td>Awesome</td>
<td>3</td>
<td>Yao Ming</td>
<td>1</td>
</tr>
<tr>
<td>Lots of People</td>
<td>134</td>
<td>Bad Food</td>
<td>3</td>
<td>Year Of The Dog</td>
<td>1</td>
</tr>
<tr>
<td>Language Barriers</td>
<td>115</td>
<td>Beautiful Women</td>
<td>3</td>
<td>Yin Yang</td>
<td>1</td>
</tr>
<tr>
<td>Busy</td>
<td>98</td>
<td>Beijing Olympics</td>
<td>3</td>
<td>Yuck</td>
<td>1</td>
</tr>
<tr>
<td>Historical Sites</td>
<td>98</td>
<td>Big Buildings</td>
<td>3</td>
<td>Yummy</td>
<td>1</td>
</tr>
</tbody>
</table>

Total Number of Phrases: 8,082, Unique Phrases: 1,641

<table>
<thead>
<tr>
<th>Affective Image Phrases</th>
<th>Top 10</th>
<th>Freq.</th>
<th>Median 10</th>
<th>Freq. last 10</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exciting</td>
<td>664</td>
<td>Smiling</td>
<td>4</td>
<td>Wishes Filled</td>
<td>1</td>
</tr>
<tr>
<td>Happy</td>
<td>592</td>
<td>Stern</td>
<td>4</td>
<td>Withdrawn</td>
<td>1</td>
</tr>
<tr>
<td>Busy</td>
<td>446</td>
<td>Stifled</td>
<td>4</td>
<td>Wondering</td>
<td>1</td>
</tr>
<tr>
<td>Interesting</td>
<td>272</td>
<td>Structured</td>
<td>4</td>
<td>Workers</td>
<td>1</td>
</tr>
<tr>
<td>Crowded</td>
<td>237</td>
<td>Stuffy</td>
<td>4</td>
<td>Work-Oriented</td>
<td>1</td>
</tr>
<tr>
<td>Scary</td>
<td>182</td>
<td>Technical</td>
<td>4</td>
<td>Worldly</td>
<td>1</td>
</tr>
<tr>
<td>Funny</td>
<td>144</td>
<td>Teeming</td>
<td>4</td>
<td>Worshipful</td>
<td>1</td>
</tr>
<tr>
<td>Different</td>
<td>142</td>
<td>Threatening</td>
<td>4</td>
<td>Wow</td>
<td>1</td>
</tr>
<tr>
<td>Boring</td>
<td>135</td>
<td>Uncertainty</td>
<td>4</td>
<td>Young</td>
<td>1</td>
</tr>
<tr>
<td>Friendly</td>
<td>133</td>
<td>Unhealthy</td>
<td>4</td>
<td>Yummy Food</td>
<td>1</td>
</tr>
</tbody>
</table>

Total Number of Phrases: 7,851, Unique Phrases: 1,001

<table>
<thead>
<tr>
<th>Uniqueness Image Phrases</th>
<th>Top 10</th>
<th>Freq.</th>
<th>Median 10</th>
<th>Freq. last 10</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Wall</td>
<td>1663</td>
<td>Forests</td>
<td>4</td>
<td>Xiamen</td>
<td>1</td>
</tr>
<tr>
<td>Beijing</td>
<td>1111</td>
<td>Gobi Desert</td>
<td>4</td>
<td>Xiao</td>
<td>1</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>384</td>
<td>Hangzhou</td>
<td>4</td>
<td>Xiao Xing</td>
<td>1</td>
</tr>
<tr>
<td>Shanghai</td>
<td>357</td>
<td>Harbin</td>
<td>4</td>
<td>Xuhng</td>
<td>1</td>
</tr>
<tr>
<td>Forbidden City</td>
<td>208</td>
<td>Mandarin</td>
<td>4</td>
<td>Yangtze River Gorge</td>
<td>1</td>
</tr>
<tr>
<td>Tiananmen Square</td>
<td>175</td>
<td>Nanking</td>
<td>4</td>
<td>Yearly Festivals</td>
<td>1</td>
</tr>
<tr>
<td>Temples</td>
<td>128</td>
<td>New Year Festival</td>
<td>4</td>
<td>Yunnan</td>
<td>1</td>
</tr>
<tr>
<td>Olympics</td>
<td>73</td>
<td>Parades</td>
<td>4</td>
<td>Zanny</td>
<td>1</td>
</tr>
<tr>
<td>Villages</td>
<td>59</td>
<td>Shenzhen</td>
<td>4</td>
<td>Zion</td>
<td>1</td>
</tr>
<tr>
<td>Three Gorges Dam</td>
<td>58</td>
<td>Tea Gardens</td>
<td>4</td>
<td>Zoo</td>
<td>1</td>
</tr>
</tbody>
</table>

Total Number of Phrases: 5,775, Unique Phrases: 463
Table 3. Top TDI Phrases of China and Their Coverage

<table>
<thead>
<tr>
<th>Phrases for Stereotypical Image</th>
<th>Top N Phrases</th>
<th>Top % of Phrases</th>
<th>% of Volume of Phrases</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1%</td>
<td>9.2%</td>
<td>26.2%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1%</td>
<td>14.1%</td>
<td>38.6%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.3%</td>
<td>23.0%</td>
<td>55.8%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.6%</td>
<td>30.3%</td>
<td>65.4%</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>1.5%</td>
<td>44.7%</td>
<td>80.2%</td>
<td></td>
</tr>
<tr>
<td>328</td>
<td>20.0%</td>
<td>81.9%</td>
<td>97.4%</td>
<td></td>
</tr>
<tr>
<td>1641</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phrases for Affective Image</th>
<th>Top N Phrases</th>
<th>Top % of Phrases</th>
<th>% of Volume of Phrases</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1%</td>
<td>8.5%</td>
<td>23.0%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.2%</td>
<td>16.0%</td>
<td>36.7%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.5%</td>
<td>28.2%</td>
<td>57.1%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.0%</td>
<td>37.5%</td>
<td>68.9%</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>2.2%</td>
<td>48.4%</td>
<td>80.0%</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>20.0%</td>
<td>84.4%</td>
<td>97.5%</td>
<td></td>
</tr>
<tr>
<td>1001</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phrases for Uniqueness Image</th>
<th>Top N Phrases</th>
<th>Top % of Phrases</th>
<th>% of Volume of Phrases</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2%</td>
<td>28.8%</td>
<td>65.7%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.4%</td>
<td>48.0%</td>
<td>84.6%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.6%</td>
<td>54.7%</td>
<td>88.5%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.1%</td>
<td>64.5%</td>
<td>90.1%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2.2%</td>
<td>73.0%</td>
<td>93.0%</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>20.0%</td>
<td>92.4%</td>
<td>98.5%</td>
<td></td>
</tr>
<tr>
<td>463</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4. Search Volumes for Top, Middle, and Bottom Phrases

<table>
<thead>
<tr>
<th>Rank</th>
<th>Keywords</th>
<th>Google Search Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>great wall china</td>
<td>550000</td>
</tr>
<tr>
<td>2</td>
<td>beijing china</td>
<td>550000</td>
</tr>
<tr>
<td>3</td>
<td>hong kong china</td>
<td>201000</td>
</tr>
<tr>
<td>4</td>
<td>shanghai china</td>
<td>368000</td>
</tr>
<tr>
<td>5</td>
<td>forbidden city china</td>
<td>12100</td>
</tr>
<tr>
<td></td>
<td>tiananmen square</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>china</td>
<td>6600</td>
</tr>
<tr>
<td>7</td>
<td>temples china</td>
<td>2900</td>
</tr>
<tr>
<td>8</td>
<td>olympics china</td>
<td>165000</td>
</tr>
<tr>
<td>9</td>
<td>villages china</td>
<td>1600</td>
</tr>
<tr>
<td></td>
<td>three gorges dam</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>china</td>
<td>6600</td>
</tr>
<tr>
<td>11</td>
<td>yangtze river china</td>
<td>90500</td>
</tr>
<tr>
<td>12</td>
<td>tibet china</td>
<td>74000</td>
</tr>
<tr>
<td></td>
<td>buddhist temples</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>china</td>
<td>390</td>
</tr>
<tr>
<td></td>
<td>xian terracotta</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>warriors china</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>pandas china</td>
<td>6600</td>
</tr>
<tr>
<td>16</td>
<td>food china</td>
<td>90500</td>
</tr>
<tr>
<td>17</td>
<td>shops china</td>
<td>9900</td>
</tr>
<tr>
<td>18</td>
<td>mountains china</td>
<td>14800</td>
</tr>
<tr>
<td>19</td>
<td>museums china</td>
<td>1300</td>
</tr>
<tr>
<td>20</td>
<td>sichuan china</td>
<td>22200</td>
</tr>
<tr>
<td>454</td>
<td>xiamen china</td>
<td>27100</td>
</tr>
<tr>
<td>455</td>
<td>xiao china</td>
<td>2400</td>
</tr>
<tr>
<td>456</td>
<td>xiao Xing china</td>
<td>0</td>
</tr>
<tr>
<td>457</td>
<td>xuhng china</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>yangtze river gorge</td>
<td></td>
</tr>
<tr>
<td>458</td>
<td>china</td>
<td>0</td>
</tr>
<tr>
<td>459</td>
<td>yearly festivals china</td>
<td>0</td>
</tr>
<tr>
<td>460</td>
<td>yunnan china</td>
<td>22200</td>
</tr>
<tr>
<td>461</td>
<td>zanny china</td>
<td>0</td>
</tr>
<tr>
<td>462</td>
<td>zion china</td>
<td>110</td>
</tr>
<tr>
<td>463</td>
<td>zoo china</td>
<td>3600</td>
</tr>
</tbody>
</table>
Table 5. The Analysis of Variance of Distinctness of Phrases to Travel Intention

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Estimated Parameters</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>981.34</td>
<td>1</td>
<td>981.34</td>
<td>1094.13</td>
<td>3.16</td>
<td>0.00**</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>99.12</td>
<td>4</td>
<td>24.78</td>
<td>27.63</td>
<td>(Fixed Factor)</td>
<td>0.00**</td>
</tr>
<tr>
<td>Age</td>
<td>61.47</td>
<td>1</td>
<td>61.47</td>
<td>68.54</td>
<td>0.01</td>
<td>0.00**</td>
</tr>
<tr>
<td>Number of World Areas Visited</td>
<td>54.29</td>
<td>1</td>
<td>54.29</td>
<td>60.53</td>
<td>-0.11</td>
<td>0.00**</td>
</tr>
<tr>
<td>Visited China</td>
<td>43.51</td>
<td>1</td>
<td>43.51</td>
<td>48.51</td>
<td>-0.73</td>
<td>0.00**</td>
</tr>
<tr>
<td>Gender</td>
<td>35.37</td>
<td>1</td>
<td>35.37</td>
<td>39.44</td>
<td>0.25</td>
<td>0.00**</td>
</tr>
<tr>
<td>Distinctiveness of Stereotypical Image Phrases</td>
<td>0.77</td>
<td>1</td>
<td>0.77</td>
<td>0.86</td>
<td>0.08</td>
<td>0.35</td>
</tr>
<tr>
<td>Distinctiveness of Affective Image Phrases</td>
<td>0.79</td>
<td>1</td>
<td>0.79</td>
<td>0.88</td>
<td>0.11</td>
<td>0.35</td>
</tr>
<tr>
<td>Distinctiveness of Uniqueness Image Phrases</td>
<td>29.51</td>
<td>1</td>
<td>29.51</td>
<td>32.91</td>
<td>0.70</td>
<td>0.00**</td>
</tr>
<tr>
<td>Error</td>
<td>2062.01</td>
<td>2299</td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>40138.00</td>
<td>231</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>2470.80</td>
<td>231</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** significant at 0.01 level.
Figure 1. An Illustration of Power-Law Distribution \( y = ax^k \), when \( a = 1 \) and \( k = -1 \)
Note: The graphs are truncated at one tenth of all unique phrases

Figure 2. The Distribution of Phrase Frequency with Ranks and Their Log-Log Transformation