

*High Frequency Key Words in Tourism English in Newspapers:
A Corpus-Based Approach*

Chattraporn Junnak, Rajamangala University of Technology Isan, Thailand
Ladawan Juajamsai, Rajamangala University of Technology Isan, Thailand

The Asian Conference on Arts & Humanities 2017
Official Conference Proceedings

Abstract

Tourism English is a new approach in the field of English for Specific (ESP) and it has its own linguistics style. The aim of the current study is to analyze the most frequent and important vocabulary in the area of tourism. The data are extracted from a small corpus amounting of 246,601 words, compiled from newspaper online in tourism section of Thailand and the program “WordSmith Tool” was used to analyze Tourism English in this corpus. In the analysis, the vocabularies were analyzed and identified the key words that were most likely to occur in tourism term. The findings revealed that the most frequently used in the corpus are more nouns and the most five frequent tourism key words were “tourism”, “million”, “travel”, “government”, and “hotel”. It is hoped that this study could enhance the study of tourism English and support learning of tourism English.

Keywords: high frequency key words, tourism English, Corpus-Based Approach

iafor

The International Academic Forum
www.iafor.org

Introduction

Nowadays, tourism has become one of the largest and most rapidly growing sectors in the global economy during the 20th Century. Since, the role of tourism affects the economy in many ways such as personal and governmental income, and employment. Moreover, because of the globalization and information become an important role, the opportunities of communication has been increased all over the world that also has greatly encouraged the international tourism exchanges. For this reason, tourism English is widely used as a meaningful medium of international tourism communication in tourism industry.

In Thailand, tourism industry is one fundamental part of income for the country as it involved Thailand's economy by promoting foreign exchange earnings, creating new jobs, broadening the distributions of income, increasing rural development. In order to communicate with tourists around the world, language is used to represent people or things and to aid the creation of thoughts (Holmes, 1998). English is the world's most greatly used language (Quirk, Greenbaum, Leech, and Svartvik, 1985) and we use it as an important factor in sending and receiving the travel information. As Thailand is an exotic country with a variety of attractions and superb standard of hospitality and service thus Thailand is placed high on the list of the world's most favored tourism destinations (Tourism Authority of Thailand, 2004). As a result, the completed information about travel to places, tourist attractions and facilities of Thailand has been printed through the World Wide Web, newspapers, and magazines in order to promote the tourism industry. Travel articles in newspapers and magazines in these media are preferred by potential tourists as tools to find new travel destination (Siriwan, 2004). One of valuable new mediums for providing information to an international audience is online electronic version of newspaper (Rademan, 1998). As newspaper is one of the most interesting types of English reading for learners because it is used as source of up-to-date English language to extend their reading after they conclude their classroom study (Frederickson & Wedel, 1991).

To understand the communicative objectives of the travel news of the tourism, the vocabulary knowledge is very essential to the readers. However, to comprehend travel news is not an easy task for non-native speakers of English thus it is necessary that they require an excellent command of vocabulary in Tourism English and also how to read the text effectively. In addition, research into tourism English has been poor. This does not mean, however, that there has been no research in the tourism English area. There are some linguistic analyzes, often based on leaflet, blog, magazine, and website in tourism. Nevertheless, there is sufficient attention to tourism English from newspaper. Accordingly, analysis that involves the study of authentic text will be useful in giving a better comprehension of the nature of language used in this specific genre and could be employed to the design of English for tourism courses and materials for teaching.

Consequently, the analysis of the present study covers tourism English and the main reason for undertaking this research is to simplify tourism English reading comprehension by building up familiarity with tourism key words. The findings gained from this study may be developed further to be used as ESP material for tourism English.

Research Objectives

The objectives of the present study are as follows:

1. To create a special corpus in the field of tourism English.
2. To identify the tourism key words in the tourism English Corpus (TEC).

Methodology

The Source of Data for the Corpus

The Source of Data for the Business Corpus analysis is the online news in the Nation's Tourism sections during the period of one month from September 1st to September 30th, 2016. The total numbers of news articles were 90 files and the total number of tokens or running word were 278,514 words.

Research Procedure

The research procedures of this study consisted of eight important steps in order to create the Tourism Corpus. The method of data analysis can be described in Figure 1 that are related to how data were examined and how to make a selection of texts in order to carry on the top 100 most frequently occurring words list.

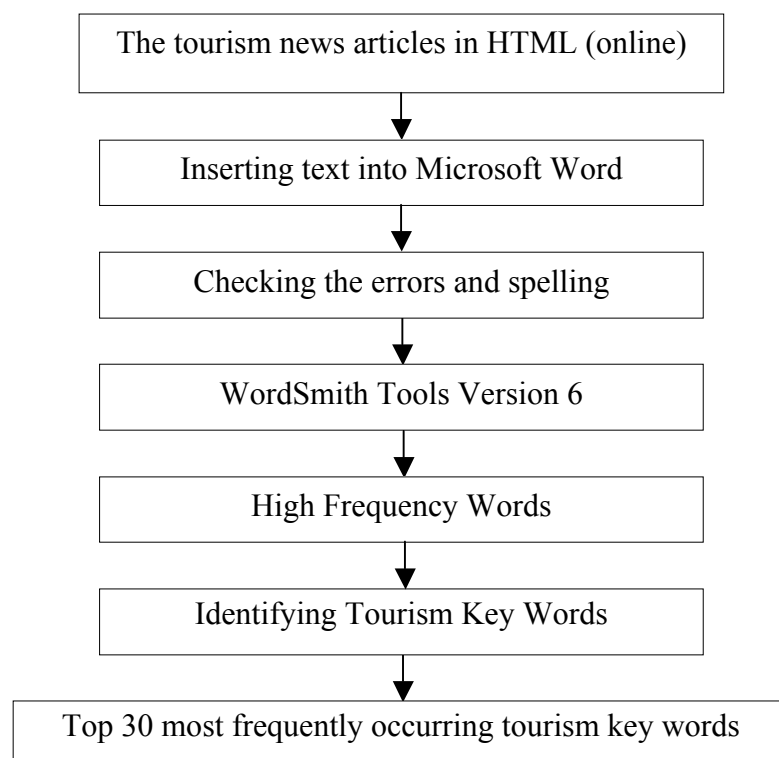


Figure 1: Research Procedures

As shown in Figure 1, there are six important steps to create the Tourism English Corpus of the study as follows:

Step 1: The news articles in HTML (online)

In order to collect the data, the tourism news articles in the online version were selected as important data for the study. The data were collected for a period of one month from September 1st to September 30th, 2016.

Step 2: Inserting text into Microsoft Word

After receiving the data in HTML format, the online texts from the tourism news articles were inserted into “Microsoft Word” as a word document (*.doc) and the total of news articles were 90 files.

Step 3: Checking the errors and spelling

At this step, the checking of the mistakes and the spellings of the words inspection were corrected before selecting texts in the analyzing program.

Step 4: WordSmith Tools Version 6

After checking the mistakes of the words inspection, all news articles were inserted into the “WordSmith Tool Version 6” for analyzing the most frequent words in the Tourism English Corpus (TEC).

Step 5: High Frequency words

All files of tourism news articles were inserted into the “WordSmith Tool Version 6” for analyzing then the program presented the words occurring frequently in the text files.

Step 6: Identifying Tourism Key Words

After receiving the most frequently words of the Tourism Corpus, the frequency word list was checked by using the Longman Business English Dictionary manually to identify the technical vocabulary. Then, the top 30 most frequently occurring tourism key words were obtained.

Results

As mentioned earlier, this study aimed to investigate high frequency tourism key words. Therefore, tourism news articles in an online version during the period of one month were selected as an important input for the present study. The information of this corpus is shown in table 1 showed.

Table 1 Statistical Details of the Tourism English Corpus (TEC)

	No. of Articles	No. of Tokens	Word types
Tourism	90	246,601	7,820

According to Table 1, the text statistics of the whole tourism English corpus showed that the tokens or running words were 246,601 words in 90 news articles and 7,820 word types.

Table 2 The Top 30 Most Frequently Occurring Tourism Key Words of the Tourism English Corpus (TEC)

Rank	Word	F	POS	Rank	Word	F	POS
1	tourism	186	n.	16	domestic	46	adj.
2	million	89	n.	17	industry	46	n.
3	travel	76	n./v.	18	local	45	adj.
4	government	75	n.	19	place	43	n./v.
5	hotel	75	n.	20	area	42	n.
6	business	71	n.	21	private	42	adj.
7	ASEAN	66	n.	22	sector	41	n.
8	economic	61	adj.	23	development	40	n.
9	market	57	n.	24	visa	39	n.
10	price	55	n.	25	high	38	n./adj./adv.
11	growth	53	n.	26	president	38	n.
12	city	51	n.	27	money	37	n.
13	centre	49	n.	28	resort	36	n.
14	time	47	n.	29	plan	35	n./v.
15	billion	46	n.	30	area	34	n.
Total							

Table 2 provides the list of the top 30 most frequently occurring words of the TEC. The total of the frequency in this table were calculated from the 1st rank up to 150th rank. There were 1,659 occurrences of the whole corpus. It is to be expected that high frequency words in the corpus were nouns and adjectives. In this list, the top five high frequency words were “tourism” appeared 186 times, “million” 89 times, “travel” 76 times, “government” 75 times, and “hotel” 75 times. Moreover, the above data indicated that the parts of speech of most key words in tourism were noun and the rest are adjectives. Yet, some words can be used both as noun and as verb (3rd rank), and as noun, adjective and as adverb (25th rank).

Table 3 Comparison of 10 Most Frequently Occurring Words of the Tourism English Corpus and the General Service List (GSL)

		General Service List (GSL)		Tourism English Corpus (TEC)	
Rank	Word	F		Word	F
1	say	2,793		tourism	186
2	make	2,378		million	89
3	man	2,132		travel	76
4	time	1,940		government	75
5	state	1,773		hotel	75
6	new	1,698		business	71
7	year	1,673		ASEAN	66
8	take	1,586		economic	61
9	come	1,577		market	57
10	now	1,531		price	55

According to Table 3, the purpose of the comparison between tourism key words of the Tourism English Corpus (TEC) and the General Service List (GSL) is to show that tourism key words are technical vocabulary while the GSL are common vocabulary. From these two lists, it can be seen that no one matched word appeared. From the data, the most frequent key word occurring in the TEC was “tourism” with 186 occurrences while “say” was the 1st rank in the General Service List (GSL) appearing 2,793 times (West, 1953).

Conclusion

This study attempted to study a corpus of tourism news articles from the tourism section. We gathered 90 news articles from the tourism news articles during September 1st to September 30th, 2016. We collected the online news articles of the in Tourism section and saved into “Microsoft Word”. Then, “WordSmith Tools Version 6” was used to identify the most frequent words and tourism key words of Tourism English Corpus (TEC). The number of tokens or running words of the TEC were 278,514 words.

The objectives of this study to analyze the vocabulary occurred in the Tourism English Corpus. The findings of the present study could be useful information in order to consider for framing the ESP course improvements and materials in the future. Furthermore, the findings pointed out that the most five frequent tourism key words were “tourism”, “million”, “travel”, “government”, and “hotel”. The findings indicated that the tourism key words of tourism appeared less frequently than the function words. However, the tourism key words are still important for learners who want to achieve in reading tourism texts.

To sum up, this study demonstrated the most frequent tourism key words. It can be said that the findings of the study could be a helpful guidance and a considerable point for teachers or course designers when they produce materials for teaching Tourism English.

Limitation of the Study

The study had a number of limitations as follows:

1. The study focused on a relatively small sampling, consisting of 90 news articles from newspaper online. Therefore, the findings may not be applicable to other kinds of tourism information such as the hotel services or rental services. Additionally, the results might not be representative of tourism English of other sources.
2. The study focused on a limited number of linguistic features that are only content words. Thus, here are other kinds of linguistic features, such as tense, voice, and clause modal verbs, imperatives, personal pronouns, and adjectival pre-modifiers which are useful for learners of the genre to improve their reading, to use each linguistic feature accurately, and to create the text effectively.

References

Frederickson, T. & Wedal, P. (1991). *English by newspaper: Reading Comprehension techniques*. Bangkok: Post Books.

Holmes, S.W. (1998). *Our Addiction to Attributive Adjective. Etc.* Vol. 55.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comparative Grammar of the English Language*. New York: Longman Group.

Rademann, T. (1998). Using Electronic Newspaper in Modern English-Language. *Press Corpora: Benefits and Pitfalls. ICAME Journal*. Vol. 22, 49-72.

Siriwan, J. (2004). Message from Governor. The official Website for TAT Governor. Retrieved September 20, 2016 from <http://tatgovernor.com>

Tourism Authority of Thailand. (2004). Thailand Moving Forward with Vision. TAT News. Retrieved September 20, 2016 from http://www.tatnes.org/tat_news/detail.asp?id=1165.

West, M. (1953). *A General Service List of English Words*. London. Longman, Green.