

## Corpus-based Creation of Tourism, Hotel, and Airline Business Word Lists

Piyapong Laosrirattanachai <sup>a,\*</sup>, Sugunya Ruangjaroon <sup>b</sup>

<sup>a</sup> piyapong.l@ku.th, Linguistics Program, Srinakharinwirot University, Thailand

<sup>b</sup> sugunya@g.swu.ac.th, Linguistics Program, Srinakharinwirot University, Thailand

\* Corresponding author, piyapong.l@ku.th

### APA Citation:

Laosrirattanachai, P. & Ruangjaroon, S. (2021). Corpus-based creation of tourism, hotel, and airline business word lists. *LEARN Journal: Language Education and Acquisition Research Network*, 14(1), 50-86.

Received  
23/06/2020

Received in  
revised form  
21/08/2020

Accepted  
10/11/2020

**Keywords**  
corpus  
linguistics,  
hospitality word  
list,  
tourism business  
word list,  
hotel business  
word list,  
airline business  
word list

### Abstract

A lack of technical vocabulary is a major problem for English for Specific Purposes (ESP) learners in a foreign setting. In this paper, we argue for using word lists to help learners expand their technical lexis repertoire. Therefore, we propose English word lists in three disciplines constructed from compiled corpora—the Tourism Business Word List (TBWL), the Hotel Business Word List (HBWL), and the Airline Business Word List (ABWL). The three word lists were derived from the vocabulary and technical terms appearing in the Tourism Business Corpus (TBC), the Hotel Business Corpus (HBC), and the Airline Business Corpus (ABC), which comprise language used in hospitality official websites, magazines, news, and work operation manuals. The corpora for ESP learners were carefully filtered through Filter Lexical Frequency, Filter Lexical Range, Filter Lexical Profiling, Filter Lexical Keyness, and via input and feedback from specialists and experts. Ultimately, the TBWL, HBWL,

	and ABWL were narrowed down to 378, 274, and 245 words, respectively, each of which was categorised into 13, 9, and 8 sub-word lists, respectively. The findings also revealed that the TBWL covered 7.76% of the TBC, the HBWL covered 7.67% of the HBC, and the ABWL covered 6.74% of the ABC.
--	--

## 1. INTRODUCTION

It has been extensively recognised that word lists and corpora are effective tools to help learners improve their vocabulary (Ma & Kelly, 2006; Nation & Waring, 1997; Read, 2000; Schmitt, 1997). In the language classroom, teachers use corpora as data-driven tools for teaching vocabulary (Smith, 2020) whilst using word lists to facilitate learners for intentional vocabulary learning outside the classroom. Some researchers have claimed that studying word lists was not much used by learners in learning vocabulary (Pookcharoen, 2016; Vo & Jaturapitakkul, 2016), which might be because it was too difficult for learners to use word lists on their own, particularly those learners with low proficiency (Fan, 2003). Even so, a number of scholars still suggested using word lists, and proposed technical word lists (Coxhead, 2000; Konstantakis, 2007; Lei & Liu, 2016; Todd, 2017; Wang, Liang & Ge, 2008; Ward, 2009; Yang, 2015) including Thai scholars (Chanasattru & Tangkiengsirisin, 2016; Itngam & Phoocharoensil, 2019; Tangpijaikul, 2014; Tongpoon-Patanasorn, 2017).

As one of the English skills, vocabulary is essential in communication. Those possessing knowledge of vocabulary but lacking the associated grammar would still understand the meaning of a sentence. Without vocabulary knowledge, it is arguably impossible to understand anything (Nosratinia et al., 2013; Wilkins, 1972). A language user needs vocabulary knowledge as a basic element to apply in listening, speaking, reading, writing, pronouncing, and putting words in the correct order in sentences (Kaya, 2014; Laufer & Nation, 1995). As a learner, encountering a few words with unknown meaning in every line makes understanding the text difficult and also results in the learner spending more time than usual in reading the text (Haynes & Baker, 1993; Laufer &

Sim, 1985; Nation, 1990, 2016). Vocabulary affects not only reading but also listening. Aitchison (2011) stated that English native speakers speak at six syllables per second on average and including pauses, gaps between speaking, and taking a breath, a speaker averages two hundred syllables per minute. This would be a big burden for learners entering the workforce where English is used as a medium for communication. Through our small survey, we found that learners in hospitality programs lacked vocabulary knowledge needed for language learning in the classroom, which obstructed their comprehension during listening to teachers or reading texts. Without adequate vocabulary knowledge, communication breakdown occurred when these learners entered the workforce.

Hospitality comprises tourism, hotel, and airline businesses and is one of the largest industries having an enormous impact on the economy in many countries all over the world. In 2019, hospitality businesses contributed approximately US\$8.9 trillion to the world's GDP (World Travel and Tourism Council, 2019). Each year, many travellers from different countries visit destinations abroad. English communication is one of the important elements in providing the best service for tourists. Service providers need to understand what is being requested and need to be able to provide appropriate information to tourists. Most learners found that understanding technical terms was one of their main problems during their studies (Evans & Green, 2007; Evans & Morrison, 2011; Ryan, 2012). As the basis, learners need to know the relevant technical terms used in hospitality businesses; not only in their textbooks (Bravo & Cervetti, 2009) but also other terms used in the real world (Hwang & Lin, 2010; Nation, 2001). Schmitt (1997) proposed vocabulary learning strategies are composed of 58 taxonomies, one of which was using word lists. Consequently, facilitating learning vocabulary in the field of hospitality was proposed by using specialised word lists designed based on a self-compiled corpus. To bridge the gap, we created hospitality word lists consisting of the Tourism Business Word List (TBWL), Hotel Business Word List (HBWL), and Airline Business Word List (ABWL) for learners to use, with the selection of the word design being carefully thought out. The word lists consist of specialist vocabulary lists and technical terms for learners to use as their reference tools.

This paper is organised as follows. In the Theoretical Background section, the background of word lists is introduced as well as the criteria commonly used in constructing a word list. The Methodology section discusses the procedures by which sources were chosen and how the corpora were compiled. Also, we propose mixed criteria abbreviated as the 6Fs, which are used to filter and create word lists. The Findings section presents the three word lists related to tourism, hotel, and airline businesses, respectively, in word family form. We conclude and discuss the findings and address the remaining issues in the Discussion section.

## 2. THEORETICAL BACKGROUND

### 2.1 Word List History and Commonly Used Criteria for Word List Construction

Originally, a word list referred to a list of high-frequency words generated by software programs (O'Keefle et al., 2007). However, nowadays, a word list refers to a list of words generated using complicated methods. Word lists are a good source for facilitating students to become autonomous learners (Todd, 2017).

Nation (2001) categorised word lists into 4 categories. The first category is a high-frequency word list which comprises basic words generally used in daily life. One of the most well-known high-frequency word lists is the General Service List (GSL), proposed by West (1953). The word list consists of 2,000 of the most used word families which cover about 80 per cent of each text (Nation & Waring, 1997). Some scholars pointed out some problems of West's GSL as it might be out of date regarding current English, as well as the word list itself being too big (Engels, 1968; Hwang, 1989; Nation & Hwang, 1995; Richards, 1974). In 2015, Brezina and Gablasova proposed a new GSL with the hope of solving these problems. However, West's GSL is still one of the most recognised word lists and is still used. The second category is the Academic Word List (AWL), which comprises words often used for academic purposes. Conventional AWLs have been proposed by Campion and Elley (1971), Praninskas (1972), Lynn (1973), Ghadessy (1979), and Xue and Nation (1984). The most well-known AWL was created by Coxhead (2000) comprising 570 words and covering about 10 per cent of each text. The third category is the Technical Word List (TWL) which

refers to words that appear with a high frequency and have a specific meaning in specific sources (Nation, 2001, 2016). This category has gained more attention from word list creators than the two previous categories since it serves as a good tool to help students in different fields to become familiar with words in their fields. Normally, a TWL covers about 5 per cent of the text; however, it may vary according to the field (Chung & Nation, 2003; Hyland & Tse, 2007). The last category is the low-frequency word list. This refers to a list of words not appearing in the three previous word lists. Such words appear at a very low frequency—approximately 1 or 2 times, in each text. This kind of word covers about 5 per cent of the text.

Laufer (1989) stated that when reading any text, one should know the vocabulary for about 95 per cent of the text. Therefore, it could be inferred that to communicate understandably, one should possess a vocabulary size of about 95 per cent of the text. When combining GSL, AWL, and TWL, the total would be about 95 per cent. Thus, the technical word lists with respect to the hospitality word lists in this study (covering about 5 per cent) should be part of the knowledge base of learners in hospitality fields.

## 2.2 Word List Creation

According to the literature review, many criteria are used to create a word list as show in Table 1.

**TABLE 1**

*Size of Corpus and Criteria Used to Create Word Lists*

Word list	Corpus size (tokens)	Criteria				
		Frequency	Range	Lexical Profiling	Expert's view	Keyness
General Service List (West, 1953)	5,000,000	✓				
Academic Word List (Coxhead, 2000)	3,500,000	✓	✓	✓		
Business Word List (Konstantakis, 2007)	600,000	✓	✓	✓		
Medical Academic Word List (Wang, Liang & Ge, 2008)	1,093,011	✓	✓	✓	✓	
Basic Engineering English	250,000	✓	✓	✓		

Word list	Corpus size (tokens)	Criteria				
		Frequency	Range	Lexical Profiling	Expert's view	Keyness
Word List (Ward, 2009)						
Academic Vocabulary in Agriculture Research Articles (Martinez, Beck & Panza, 2009)	826,416	✓	✓	✓		
Academic Vocabulary in Chemistry Research Articles (Valipouri & Nassaji, 2013)	4,000,000	✓	✓			
Academic Word List for Applied Linguistics Research Articles (Khani & Tazik, 2013)	1,553,450	✓	✓	✓		
Technical Keywords for Business (Tangpijaikul, 2014)	890,000			✓		✓
Vocabulary of Agriculture Semi-Population Article (Muñoz, 2015)	455,366	✓		✓		
Nursing Academic Word List (Yang, 2015)	1,006,934	✓	✓	✓		
New Medical Academic Word List (Lei & Liu, 2016)	6,200,000	✓	✓			
Opaque Engineering Word List (Todd, 2017)	1,150,000		✓	✓		✓
Science Academic Word List (It-ngam & Phoocharoensil, 2019)	5,500,000	✓	✓	✓	✓	

In summary, there are 5 major criteria used to create a word list that vary between scholars, with frequency, range, and lexical profiling being the most common. In linguistics, frequency is one of the important concepts in studying language use (Lindquist, 2009). The first 100 high-frequency words cover about half of the text (Zipf, 1935). Therefore, paying attention only to the frequency for the creation of a word list is inadequate (Schmitt & Schmitt, 2014). However, while the frequency is not the most appropriate way to create a word list, it cannot be ignored.

Range is also an important criterion to consider. Coxhead (2000) stated that considering the frequency alone for a very long text leads to bias. To solve the problem, Coxhead proposed range as one of the criteria. Lexical profiling is about classifying words into groups; however, a word should not appear in more than one group. This criterion removes

any words that the word list creator believes are irrelevant. For example, Coxhead (2000) ignored all words appearing in the GSL to create her AWL. However, lexical profiling has been criticised because the criterion itself reduces the chance of some words appearing in the word list since some words might have other meanings related to the specific fields (Billuroglu & Neufeld, 2007; Cabre, 1999; Gardner & Davies, 2014; Paquot, 2007; Pearson, 1998; Valipouri & Nassaji, 2013). However, Itngam and Phoocharoensil (2019) advocated using a lexical profiling method to create a specialised word list and claimed that students should be aware of words in the GSL and AWL prior to learning specialised words.

Expert inputs and feedback are also important and should not be omitted. Expert opinion regarding the inclusion of a specific word in a word list should be obtained from people active in the field the word list is created for. Their experiences of using words on a daily basis is a valuable input that can help a word list creator to make good decisions regarding which words to include in a list (Chung & Nation, 2004).

Keyness is the result of keyword analysis. Keyword analysis, used broadly among corpus linguists, refers to identifying keywords appearing in a corpus (Gabrielatos & Machi, 2012). A keyword is an unusually high-frequency word appearing in the target corpora compared to the reference corpora considered, instead of just using the frequency alone (Rayson & Garside, 2000; Scott, 2008). Normally, the reference corpora are huge, such as the British National Corpus (BNC) (Johnson & Esslin, 2006; Scott, 2001). Generally, keyword analysis is calculated using log-likelihood or chi-square techniques (Anthony & Gladkov, 2007).

There is no definitive proof to indicate which criterion is the best because each has its advantages and disadvantages. Hyland and Tse (2007) stated correspondingly that word lists in different fields should be created in different ways. Thus, we proposed to use all the criteria based on applying mixed criteria in a step-by-step approach. The mixed criteria (called the Six Filters or 6Fs) are presented and discussed in the following section.

### 3. METHODOLOGY

#### 3.1 Compiling Corpora

To create an efficient and fair word list, a corpus should be chosen and compiled carefully. Since we aimed to gather authentic language use, we surveyed graduates working in three areas using Google Form prior to creating the specific word lists. The QR code generated from the Google Form was shared to a private Facebook group with more than 4,500 members. The members of this private Facebook group are lecturers, students, and graduates of one of the programs concerning hospitality businesses in Thailand. The requirements of the respondents were that they must have worked in a tourism, hotel, or airline business for more than one year. The respondents were asked to provide information including general information, for example, hospitality field, years of experience, and position. The most essential information was that apart from verbal communication with tourists, which English language source should people who are going to enter the hospitality workforce be familiar with. The Google form was kept open for responses for 2 weeks. As a result, there were 446 respondents with 367 respondents mentioning the ability to understand English in websites related to a tourism, hotel, or airline business was essential since tourists always ask questions about the information in websites. Some other sources were also suggested including magazines, news, and work operation manuals. As a result of this needs analysis, we started from their recommended language sources. Consequently, the corpora were compiled by gathering data from 2 main sources in each hospitality business, namely websites and the corpus depending on availability of access to the information. The corpora information is shown in Table 2.

**TABLE 2**

*Sources and Tokens of Tourism Business, Hotel Business and Airline Business Corpora*

Hospitality Business Corpus	Sources	Number of Sources	Tokens
Tourism Business Corpus	- Official tourism websites	152	31,701,430

Hospitality Business Corpus	Sources	Number of Sources	Tokens
(TBC)	- Tourism magazines		
Hotel Business Corpus (HBC)	- Official hotel websites - Hotel business news	124	4,835,926
Airline Business Corpus (ABC)	- Official airline websites - Airline work operation manuals	120	15,542,604

The Tourism Business Corpus comprises 31,701,430 running words from 152 different sources as follows: 1) 100 official tourism websites of the first 100 countries with the largest number of travellers ranked by the United Nations World Tourism Organization (2017), and 2) 52 tourism magazines published from 2017 to 2018.

The Hotel Business Corpus was compiled by collecting data from 100 official hotel websites ranked as the best 100 hotels in 2017 by Gifford (2018) from [www.travelandleisure.com](http://www.travelandleisure.com), and hotel news from 2017 to 2018. As a result, the Hotel Business Corpus contained 4,835,926 running words.

The Airline Business Corpus was compiled by gathering data from 100 official airline business websites rated as the best 100 airlines in 2017 by Skytrax (2018) and from airline work operation manuals. As a result, the corpus had 15,542,604 running words.

All the texts were first copied and pasted into Notepad and then saved as text files. Afterwards, all the files were processed using software programs and the 6Fs were used subsequently to create the word lists in three disciplines.

### 3.2 Data Processing

To create the three word lists, the researchers proposed criteria arranged systematically called the Six Filters (6Fs) as follows.

Filter Lexical Frequency – This was the first filter used to create the word lists. Coxhead’s frequency criterion was applied in this study (Coxhead, 2000). In her study, Coxhead compiled a corpus of 3,500,000 tokens, and any word that appeared at least 100 times was considered as passing the frequency criterion. To create the word lists in this study, the following equations were used to set the cut-off point of the Filter Lexical Frequency.

Tourism Business Word List (TBWL)	Hotel Business Word List (HBWL)	Airline Business Word List (ABWL)
$100 \times \frac{31,701,430}{3,500,000}$	$100 \times \frac{4,835,926}{3,500,000}$	$100 \times \frac{15,542,604}{3,500,000}$
= 906	= 138	= 444

In summary, words that appeared in the TBC at least 906 times would pass the Filter Lexical Frequency criterion, while words in the HBC and ABC would pass the Filter Lexical Frequency criterion if they appeared at least 138 and 444 times, respectively. The AntWordProfiler program (Anthony, 2018) was used to determine the frequencies in this study.

Filter Lexical Range – This was the second filter and again we used the AntWordProfiler program to extract words for inclusion in the word lists. The range criterion of Coxhead (2000) was applied. Coxhead compiled her corpus by gathering language data from 28 sources. She claimed that words that appeared in at least 15 resources passed the range criterion test. In the current study, words that appeared in at least 50 per cent of the total sources passed this criterion. Therefore, words that appeared at least 76, 62, and 60 times in the TBC, HBC, and ABC, respectively, passed the criterion and tended to be included in the word lists.

Filter Lexical Profiling – The main concept of lexical profiling is that a word should be put in only a single word list. This eliminates irrelevant words from the created word lists. The researchers used the GSL (West, 1953), AWL (Coxhead, 2000), Function Word List (FWL), Abbreviation List (AL), and the Proper Name List (PNL) as the target word lists. The FWL, AL, and PNL were created by Nation (2018) and can be downloaded at <https://www.wgtn.ac.nz/lals/resources>. Any words that appeared in the mentioned word lists would be ignored. The AntWordProfiler program was used in this filter.

Filter Lexical Keyness – This was used to consider unusually high-frequency words appearing in the TBC, HBC, and ABC compared to the British National Corpus (BNC) used as the reference corpora in this study based on the log-likelihood applied in the Key-BNC program (Graham, 2018). The cut-off point used in the Filter Lexical Keyness was applied

from Todd (2017). In his study, Todd considered the 500 words with the highest log-likelihood value from his corpus of 1,150,000 tokens. When calculating the size of the corpus in the current study, the first 13,783, 2,103, and 6,758 words with the highest Log-likelihood values in the TBC, HBC, and ABC, respectively, were considered as passing the cut-off point. However, these numbers of words would be decreased again later since the words passing the Filter Lexical Keyness must also pass the frequency and range criteria set in the Filter Lexical Frequency and Range. Words not appearing in the GSL and AWL were ignored.

Filter Expert Consultation – The inputs and feedback from the experts and specialists in the field were gathered to ensure that the word lists were well designed and authentically used in the industry because they use ESP both receptively and productively on a daily basis. Their inputs helped decide which words were appropriate to be included in the word lists (Chung & Nation, 2004; Martinez et al., 2009). In the current study, the criterion of Chung and Nation (2004) was used to create the word-list check list with 4 rating scales. The first scale referred to words with a meaning irrelevant to the tourism, hotel, and airline business fields. The second scale referred to words with a meaning of little relevance to the tourism, hotel, and airline business fields. The third scale referred to words with a meaning very relevant to the tourism, hotel, and airline business fields, including polysemous words which have one meaning in a general context and a specific meaning relevant to the tourism, hotel, and airline business when appearing in a hospitality context. The fourth scale referred to words with a meaning specific to the tourism, hotel, and airline business that were not included in other fields. The list of words in the tourism field passing the previous 4 filters was distributed to 5 experts who had been working in tourism-related businesses for more than 5 years. The lists of words in the hotel and airline business were treated the same way with 5 experts from the hotel field and 5 experts from the airline business field. The researchers considered the mode value of each word. Finally, words with a mode value of 3 or 4 were added to the word lists.

Filter Lexical Difficulty – Since a long list of words might cause recognition difficulties for users, dividing such a list into shorter sub-word lists is one way to solve this problem. In the current study, the researchers used the VocabProfile program (Cobb, 2018) to divide the three main word lists into sub-word lists based on the difficulty of the

words. The program operates by categorizing words into 25 Base Lists of vocabulary. Each Base List comprises 1,000 words which are separated based on the commonly used level, where the more commonly a word is used, the earlier it appears in the Base List.

#### 4. FINDINGS

Before processing the data, the tokens were changed to the type form, with the number of types in each corpus being: 302,128 types in TBC, 65,737 types in HBC, and 134,862 types in ABC. These were then analysed by using the 6Fs (Table 3).

**TABLE 3**

*Number of Words in the 3 Word Lists Using the 6Fs*

Research Procedure	TBWL		HBWL		ABWL	
	Satisfying the filter itself	Satisfying itself and previous filter(s)	Satisfying the filter itself	Satisfying itself and previous filter(s)	Satisfying the filter itself	Satisfying itself and previous filter(s)
**Tokens to Types	31,701,430	302,128	4,835,926	65,737	15,542,604	134,862
Filter Lexical Frequency	2,465	2,465	3,548	3,548	2,381	2,381
Filter Lexical Range	2,109	1,785	1,243	1,216	2,047	1,714
Filter Lexical Profiling	273	273	178	178	176	176
Filter Lexical Keyness	446	719	346	524	682	858
**Types to Word Families	719	672	524	403	858	606
Filter Expert Consultation	378	378	274	274	245	245
Filter Lexical Difficulty	378 words separated into 13 sub-word lists		274 words separated into 9 sub-word lists		245 words separated into 8 sub-word lists	

#### 4.1 Filter Lexical Frequency

To construct the TBWL, HBWL, and ABWL, we started by applying the first filter, the Filter Lexical Frequency. The Filter Lexical Frequency criteria varied and were determined by the size of the corpus. When analysing the frequency criteria using the AntWordProfiler, 2,465 words in TBC, 3,548 words in HBC, and 2,381 words in ABC passed the criteria. The first 30 highest-frequency words in the three corpora were mostly function words, e.g. *the, and, of, not, of, to, is, for, with, this, that, with,*

and *from*. Some examples of content words in the first 30 highest-frequency words were *visit, year, hotel, guest, flight, travel, and baggage*.

## 4.2 Filter Lexical Range

The data were passed to the Filter Lexical Range to create further corpora. The numbers of words passing both the Filter Lexical Frequency and the Filter Lexical Range were 1,785 (TBC), 1,216 (HBC) and 1,714 (ABC), respectively. The three corpora were next filtered by considering the Filter Lexical Range. In the TBC, the number of the words passing both the Filter Lexical Frequency and the Filter Lexical Range was 1,785 and for the HBC and ABC was 1,216 and 1,714, respectively. Some sample words that passed the two criteria were: *the, a, able, about, above, adapt, beer, convenience, harbour, lagoon, appoint, adjacent, approach, delay, apart, and drop*. When comparing this group of words with the words filtering by the Lexical Frequency, it was found that the proportion of content words substantially increased. It also reflected words that were more related to the tourism, hotel, and airline businesses.

## 4.3 Filter Lexical Profiling

Words that had passed both the Filter Lexical Frequency and the Filter Lexical Range were then analysed by their profile using the AntWordProfiler program. The results are shown in Table 4.

**TABLE 4**

*Number of Words Appearing in Different Profiles*

Profile	TBWL		HBWL		ABWL	
	Number	%	Number	%	Number	%
1 <sup>st</sup> 1,000 GSL	731	40.95	609	50.08	850	49.59
2 <sup>nd</sup> 1,000 GSL	317	17.76	153	12.58	217	12.66
AWL	234	13.11	117	9.62	248	14.47
FWL	182	10.20	148	12.17	174	10.15
AL	5	0.28	4	0.33	7	0.41
PNL	43	2.41	7	0.58	42	2.45
The rest	273	15.29	178	14.64	176	10.27

Profile	TBWL		HBWL		ABWL	
	Number	%	Number	%	Number	%
Total	1,785	100	1,216	100	1,714	100

The Filter Lexical Profiling removed words that appeared in the GSL, AWL, FWL, AL, and PNL. As a result, 273 words, 178 words, and 176 words were left to be considered for the TBWL, HBWL, and ABWL, respectively. The words removed by this Filter were mostly salient to domain-specific contexts. Some examples are *archaeological*, *breathhtaking*, *cathedral*, *excursion*, *accommodation*, *acre*, *amenities*, *deluxe*, *heritage*, *airbus*, *altitude*, *cancellation*, *charter*, and *clearance*.

#### 4.4 Filter Lexical Keyness

Since the Filter Lexical Profiling might remove some polysemous words located in the GSL and AWL, we considered words located in the GSL and AWL that had a possible meaning relevant to the tourism, hotel and, airline businesses by using the Filter Lexical Keyness and considering the log-likelihood value. The Key-BNC program was used to calculate this statistic, resulting in 1,405 words, 801 words, and 1,452 words in the TBWL, HBWL, and ABWL, respectively, passing the Keyness criteria. However, they were again treated by extracting only those passing the Filter Lexical Frequency and the Filter Lexical Range. This resulted in 446 words in the TBWL, 346 words in the HBWL, and 682 words in the ABWL being extracted and passing the Filter Lexical Keyness. Combining the words from the Filter Lexical Profiling and the Filter Lexical Keyness resulted in 719 words in the TBWL, 524 words in the HBWL, and 858 words in the ABWL.

#### 4.5 Filter Expert Consultation

The input and feedback from the experts and specialists in the relevant field were gathered to ensure that the word lists were well designed and authentically and genuinely used in the industry because these experts and specialists use ESP both receptively and productively on a daily basis. Their contribution helped to decide which words were appropriate to be included in the word lists. Before reaching the last filter, the opinions of professionals using an authentic language was needed. The researchers

applied Filter Expert Consultation to gather data. However, a long list of words might exhaust the professionals leading to inaccurate responses. Thus, the words in the list were changed from type form into word family form, which decreased the number of words in each word list, so that 719 words became 672 words in the TBWL, 524 words became 403 words in the HBWL, and 858 words became 606 words in the ABWL. The TBWL was then distributed to 5 professionals in the tourism field. Likewise, the HBWL was distributed to 5 professionals in the hotel field, and the ABWL was distributed to 5 professionals in the airline business field. The results from the professionals' input were that 378 words in the TBWL, 274 words in the HBWL, and 245 words in the ABWL were used and relevant to the tourism field, hotel field, and airline business fields, respectively. These words were the final set and were used to facilitate teaching and learning the English vocabulary used in the tourism, hotel, and airline businesses.

#### **4.6 Filter Lexical Difficulty**

According to Laosrirattanachai and Ruangjaroon (2020), word lists with a close relationship, such as in the tourism, hotel, and airline businesses, have a high possibility of sharing some words. Consequently, we compared the three word lists and found that there were 36 common words appearing in the TBWL, HBWL, and ABWL. These 36 words were put in the first sub-word list of each of the 3 fields. Then, the remaining words in each word list were allocated into sub-word lists using the VocabProfiler program, with 30 words assigned to each sub-word list except for the last sub-word list that contained the remaining words. Ultimately, the TBWL, the list with the largest size compared to the other two, consisted of 13 sub-word lists, while the HBWL comprised 9 sub-word lists, and the ABWL, as the smallest, comprised 8 sub-word lists.

#### **4.7 Coverage of Word Lists**

Then, we used the AntWordProfiler to find the word list coverage for each corpus. The results showed that the TBWL covered 7.76 per cent of the TBC, the HBWL covered 7.69 per cent of the HBC, and the ABWL covered 6.76 per cent of the ABC. The proportions of coverage were adequate to enhance learners' vocabulary knowledge, as Chung and

Nation (2003) and Hyland and Tse (2007) stated that a technical word list covers about 5 per cent of the text, though this could vary with different texts.

## 5. DISCUSSION

One of the main advantages of the created word lists is that these lists prepare learners for their career path because the corpora used to construct the word lists have been mainly compiled from authentic sources, such as websites, and not from textbooks or research articles as is common with other constructed word lists. When comparing the TBWL, HBWL, and ABWL in the study with other related word lists, for example, Business Word List (Konstantakis, 2007), Academic Word List of Business English (Tongpoon-Patanasorn, 2017), and Technical Keywords for Business (Tangpijaikul, 2014), the TBWL, HBWL, and ABWL are more appropriate within the hospitality business context. Their coverage is shown in Table 5.

**TABLE 5**

*Coverage of the Six Word Lists*

Word list/Corpus	TBC	HBC	ABC
TBWL/HBWL/ABWL	7.76	7.69	6.76
Business Word List	5.17	5.60	5.74
Academic Word List of Business English	0.52	0.83	0.55
Technical Keywords for Business	0.27	0.59	0.37

The TBWL, HBWL, and ABWL cover larger proportions of the TBC, HBC, and ABC, respectively than the other three word lists. Nation and Waring (1997) considered a created technical word list should cover approximately 5 per cent of the text. The coverages of the created word lists in the current study were the TBWL covered about 7.76 per cent of the TBC, the HBWL covered about 7.69 per cent of the HBC, and the ABWL covered about 6.76 per cent of the ABC. These values indicated that the three word lists would be worth learning for both academic and career opportunity purposes. Even though Konstantakis' Business Word List covers the TBC, HBC, and ABC at 5.17, 5.60, and 5.74 per cent,

*LEARN Journal: Vol. 14, No. 1 (2021)* *page 65*

respectively, the TBWL, HBWL, and ABWL are still more appropriate and useful in a hospitality business context considering their levels of coverage.

The 6Fs were used to extract relevant words for inclusion in the word lists. Each filter had its advantages and disadvantages, but by using them in combination, the 6Fs provided a systematic approach. The Filter Lexical Frequency, the most extensively used to create a word list (Lindquist, 2009), can extract an enormous number of words. However, irrelevant words (such as function words or generally used words) covering approximately 50 per cent of the text (Zipf, 1935), were also extracted. The different sizes of different language sources could lead to bias. The Filter Lexical Range could reduce such bias by considering the data distribution in the different sources. However, even though the Filter Lexical Range could minimise the bias, there were still some irrelevant words in the extracted words. The Filter Lexical Profiling could remove a number of irrelevant words, especially function words (Cobb & Horst, 2001). Among the 6Fs, the Lexical profiling might be criticised by many researchers since it eliminated all words appearing in the GSL, AWL, and other word lists used as referent word lists. The method itself allowed some flexibility to remove some words considering only their meaning with regard to hospitality contexts or some polysemous words that appeared in the required word list (Billuroglu & Neufeld, 2007; Gardner & Davies, 2014; Paquot, 2007; Schmitt & Schmitt, 2014; Valipouri & Nassaji, 2013). However, erased words having a possible hospitality meaning should be reincluded in the list. This is addressed by the Filter Lexical Keyness that draws back any words that appear frequently and significantly compared to a large target corpus like the British National Corpus (Laosrirattanachai & Ruangjaroon, 2020). It should be noted that to make the filters consistent, the drawn back words using the Filter Lexical Keyness must also pass the frequency and range criteria. After obtaining words from the four earlier filter programs, the fifth filter (Filter Expert Consultation) utilises human decision making based on professional experience to decide which words should be in the final lists. Researchers have confirmed that expert experience was one of the most valuable sources of opinion on whether words should be

contained in the lists (Chung & Nation, 2004; Martinez et al., 2009). Many scholars have argued that a large-sized word list makes it difficult for practical application as it could overwhelm learners (Brezina & Gablasova, 2015; Engels, 1968; Hwang, 1989; Nation & Hwang, 1995; Richards, 1974). Consequently, the lists in the current study were divided into shorter sub-lists based on their difficulty before being presented to learners with the principle that easier sub-lists should be learnt prior to more difficult ones.

The created word lists might be useful for learners in the tourism, hotel, and airline business fields. The lists could expand learners' vocabulary knowledge and prepare them for the hospitality workforce. Teachers would be able to give students guidance and urge them to use the lists wisely since learning vocabulary in classes might be inadequate due to limitations of time and textbooks for used as learning material, as the latter are normally made up of easier words than are often applied in the authentic language used in workplaces. Learners would soon find that words appearing in the word lists were created from authentic language commonly used in workplaces, some of which might otherwise only rarely be seen in more traditional classroom texts.

To provide students with word lists for autonomous learning, teachers should first consider a learner's English competency. Since the words contained in the word lists varied from beginner to advanced level, choosing sub-word lists to suit a student's needs and level is important. The difficulty levels of words contained in the sub-word lists vary, but with regard to hospitality, other sub word-lists might be difficult and might never be encountered in learning at the university level. Because of this, teachers can choose, provide, and guide appropriate sub-word lists to suit students' needs. Consequently, teachers should ensure their students learn word lists one-by-one and do not skip or step back too far so that the students can make connections between sub-word lists. It would be much better for learners to be evaluated for their pre-existing knowledge before learning an appropriate level of a word list (Schmitt, 2010).

Even though using a word list was one of the vocabulary-learning strategies mentioned by Schmitt (1997), there are differing views on word list implementation. Some studies have claimed that the use of a word list could provide a powerful benefit and hence should be integrated in a curriculum (Ma & Kelly, 2006; MacArthur & Littlemore, 2008; Nation & Waring, 1997, Read, 2000; Schmitt, 1997; Smith, 2020; Vongpumivitch et al., 2009). However, others argue against the use of word lists in a classroom, as a word list application was perceived as one of the least useful strategies in learning vocabulary and might not be accepted in teaching widely as its language was too difficult for most learners. Instead it has been advocated that vocabulary should be learned from context instead of vocabulary-focused learning material (Boulton, 2008, 2009; Fan, 2003; Hulstijn, 2001; Lamy et al., 2012; Yoon & Hirvela, 2004). However, we view a word list as having complementarity with both advantages and disadvantages, depending on whether teachers apply different materials in different settings appropriately. We have advocated the application of corpora-driven and word-lists learning in enhancing learners' vocabulary knowledge. Consequently, we have suggested applying the created word lists in teaching and learning vocabulary both in class and autonomously as a supplement.

For example, we recommend teachers assign students to construct their vocabulary portfolio on their own as a do-it-yourself (DIY) task as proposed by Smith (2020). Students could design their DIY word lists by selecting words from the three word lists based on their awareness and needs and use it as part of their meaningful learning and training to think in English under the guidance of instructors.

## TABLE 6

### *Sample of DIY Word List Output in the Three Disciplines*

Word	Sentence extracted from corpus	Making learners' own vocabulary portfolio
Words retrieved from TBWL		
arch	The Gateway of India is an arch-	Passing through the arch,

Word	Sentence extracted from corpus	Making learners' own vocabulary portfolio
	monument built in the early twentieth century in the city of Mumbai, in the Indian state of Maharashtra	you enter a palace.
cathedral	You will notice how local craftsmen adorned the inside of the <b>cathedral</b> with intricate decorations and stone sculptures depicting saints, angels and sacred symbols.	A guide conducts tours of the <b>cathedral</b> every morning at 9.00.
Words retrieved from HBWL		
banquet	The most efficient way to produce food from a cost of sales and labour point of view is to have a main kitchen, <b>banquet</b> floor kitchen and satellite kitchens for the outlets.	The food was delivered to the <b>banquet</b> by a catering service.
toiletries	The rooms also include the signature Hyatt Grand bed, luxury bath <b>toiletries</b> and hi-tech amenities, including a 42" flat screen TV, to guarantee the optimum comfort for guests.	There were shampoo and other <b>toiletries</b> in my bag.
Words retrieved from ABWL		
aisle	The child became so unruly during the flight that he ran down the <b>aisle</b> .	The flight attendant came down the <b>aisle</b> serving drinks.
starboard	The forward entry door is located directly behind the first officer's seat on the <b>starboard</b> side of the aircraft flight deck	When flight crew talk about <b>starboard</b> , they are referring to the right side of the plane.

## REMAINING ISSUES

Since we argued for ESP word lists that can help language learning beyond basic English communication for learners of hospitality, we proposed using word lists and corpora in a practical manner. To come to terms with pedagogic practices, we will customise some beneficial and effective learning materials, such as interactive online lessons and quizzes, as an innovation developed from the three word lists. Their benefit to vocabulary knowledge development needs to be evaluated. Future study should examine the effects or the outcomes of using word

lists and assess whether they help learners enhance their vocabulary knowledge. Learners might find learning from corpora more beneficial when such material is constructed more meaningfully and so is of more direct use in their studies and future career. They might feel more engaged and satisfied because we as instructors provide them with options in learning ubiquitously.

## LIMITATIONS OF THE STUDY

A limitation of the study was accessing online sources. The study aimed to gather the data from online sources including websites and other sources in the form of PDF files. PDF files of tourism magazines were available, but not for the hotel and airline magazines. The available PDF files of airline magazines are mostly about selling tax-free products and text on the airplanes and engines which are not relevant to the service sector. The second source of HBC was collected from online news, which mostly provides information about hotels and resorts, for example, rooms, facilities, and services. Tourism and airline news mostly provide information about economic impacts, which are rarely relevant to the service sector. The work operation manuals were chosen as the second source for compiling the ABC on purpose because information received from graduates working in the airline business indicated that airline personnel are required to read work operation manuals containing information about the rules, regulations, and security policy with regard to working at an airport. They found that understanding the manuals was virtually impossible without prior vocabulary knowledge concerning airline work operation manuals.

## ABOUT THE AUTHORS

***Piyapong Laosrirattanachai***: a graduate student of the Linguistics program at the Faculty of Humanities, Srinakharinwirot University, Bangkok, Thailand. He is now an assistant professor of the English for Service Industry program at Kasetsart University. His research interests include vocabulary teaching and learning, and corpus linguistics.

**Sugunya Ruangjaroon:** an Associate Professor of Linguistics at Srinakharinwirot University, Bangkok, Thailand. She serves as the Director for the Graduate Studies, Faculty of Humanities. Her research interests are syntax, phonetics and phonology, and language teaching.

## REFERENCES

- Aitchison, J. (2011). *The articulate mammal: An introduction to psycholinguistics*. Routledge.
- Anthony, L. (2018, June, 22). *AntWordProfiler*. [Software]. <http://www.laurenceanthony.net/software>.
- Anthony, M., & Gladkov, K. (2007). Rhetorical appeals in fundraising. In U. C. Biber & A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 121-151). John Benjamin.
- Billuroglu, A., & Neufeld, S. (2007). BNL 2709: *The essence of English* (4th ed.). Rustem Kitabevi.
- Boulton, A. (2008). Bringing corpora to the masses: Free and easy tools for language learning. In N. Kübler (Ed.), *Corpora, language, teaching, and resources: From Theory to practice*. Peter lang.
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37-54.
- Bravo, M. A., & Cervetti, A. (2009). Teaching vocabulary through text and experience in content areas. In M. F. Graves (Ed.), *Essential readings on vocabulary instructions* (pp. 141-152). International Reading Association.
- Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, 36(1), 1-22.
- Cabre, T. (1999). *Terminology, theory, methods and applications*. John Benjamins.
- Campion, M., & Elley, W. (1971). *An academic vocabulary list*. New Zealand Council for Educational Research.
- Chanasattru, S., & Tangkiengsirisin. (2016). Developing of a high frequency word list in social sciences. *Journal of English Studies*, 11(2016), 41-87.
- Chung, T., & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language*, 15(2), 103-116.

- Chung, T., & Nation, P. (2004). Identifying technical vocabulary. *System*, 32(2), 251-263.
- Cobb, T. (2018, June, 22). *Web Vocabprofile*. [Online program]. <http://www.lex tutor.ca/vp/>.
- Cobb, T., & Horst, M. (2001). Reading academic English: Carrying learners across the lexical threshold. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 315-329). Cambridge University Press.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Engels, L. (1968). The fallacy of word counts. *International Review of Applied Linguistics*, 6(1-4), 213-231.
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6(1), 3-17.
- Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first year experience in Hong Kong. *English for Specific Purposes*, 30(3), 198-208.
- Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, 87(2), 223-241.
- Gabrielatos, C., & Machi, A. (2012). Keyness: Appropriate metrics and practical issues. Paper presented at *the CADS International Conference*, Bologna, Italy.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305-327.
- Ghadessy, M. (1979). Frequency counts, words lists, and materials preparation: A new approach. *English Teaching Forum*, 17(1), 24-27.
- Gifford, J. (2018, April, 1). *The top 100 hotels in the world*. <http://www.travelandleisure.com/worlds-best/hotels-top-100-overall>.
- Graham, D. (2018, June, 22). *Key-BNC*. [Software]. <http://crs2.kmutt.ac.th/Key-BNC/>.
- Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English text. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary* (pp. 130-152). Ablex.

- Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal, and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge University Press.
- Hwang, K. (1989). *Reading newspapers for the improvement of vocabulary and reading skills* [Master Thesis], Victoria University of Wellington, New Zealand].
- Hwang, Y., & Lin, S. (2010). A study of medical students' linguistic needs in Taiwan. *Asian ESP Journal*, 6(1), 35-58.
- Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly*, 41(2), 235-253.
- It-ngam, T., & Phoocharoensil, S. (2019). The development of science academic word list. *Indonesian Journal of Applied Linguistics*, 8(3), 657-667.
- Johnson, S., & Esslin, A. (2006). Language in the news: Some reflections on keyword analysis using WordSmith Tools and the BNC. *Leeds Working Papers in Linguistics and Phonetics*, 11.
- Kaya, J. (2014). *Vocabulary learning strategies: A study of Congolese English language learners* [Master's thesis, Southern Illinois University Carbondale].
- Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for Applied Linguistics Research Articles. *RELC Journal*, 44(2), 209-232.
- Konstantakis, N. (2007). Creating a business word list for teaching business English. *Estudios De Linguística Inglesa Aplicada (ELIA)*, 7, 79-102.
- Lamy, N., & Klarskov Mortensen, H. J. (2012). Using concordance programs in the modern foreign languages classroom. Module 2.4. In G. Davies (Ed.), *Information and communications Technology for language teachers* (ICT4LT). Thames Valley University.
- Laosrirattanachai, P. & Ruangjaroon, S. (2020). The word lists of hospitality service review construction. *Journal of Studies in the English Language*, 15(1), 107-158.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Laurén & M. Nordmann (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Multilingual Matters.

- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-332.
- Laufer, B., & Sim, D. (1985). Taking the easy way out: Non-use and misuse of contextual clues in EFL reading comprehension. *English Teaching Forum*, 23(2), 7-10.
- Lei, L., & Liu, D. (2016). A new medical academic word list: A Corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, 22(2016), 42-53.
- Lindquist, H. (2009). *Corpus linguistics and the description of English*. Edinburgh University Press.
- Lynn, R. (1973). Preparing word lists: A suggested method. *RELC Journal*, 4(1), 25-32.
- Ma, Q., & Kelly, P. (2006). Computer-assisted vocabulary learning: Design and evaluation. *Computer-Assisted Language Learning*, 19(1), 15-45.
- MacArthur, F., & Littlemore, J. (2008). A discovery approach to figurative language learning with the use of corpora. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistics approaches to teaching vocabulary and phraseology* (pp. 159-188). Mouton de Gruyter.
- Martinez, I., Beck, S., & Panza, C. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, 28(3), 183-198.
- Muñoz, V. (2015). The Vocabulary of agriculture semi-popularization articles in English: A corpus-based study. *English for Specific Purposes*, 45(3), 26-44.
- Nation, P. (1990). *Teaching and learning vocabulary*. Heinle and Heinle.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, P. (2016). *Making and using word lists for language learning and testing*. Philadelphia John Benjamins Publishing Company.
- Nation, P. (2018, April, 10). *Resources*.  
<https://www.wgtn.ac.nz/lals/resources>.
- Nation, P., & Hwang, K. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, 23(1), 35-41.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary, description, acquisition and pedagogy* (pp. 6-19). Cambridge University Press.

- Nosratinia, M., Divani, N. S., & Zaker, A. (2013). Language learners' internal factors and practical applications: A case of vocabulary learning strategies. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, 17(1), 100-115.
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Paquot, M. (2007). Towards a productively-oriented academic wordlist. In e. a. J. Walinski (Ed.), *PALC Proceedings* (pp. 127-140). Peter Lang.
- Pearson, J. (1998). *Terms in context*. John Benjamins.
- Pookcharoen, S. (2016). Thai EFL University Teachers' Beliefs and Practices about Vocabulary Learning Strategies. *Language Education and Acquisition Research Network (LEARN) Journal*, 9(2), 155-172.
- Praninskas, J. (1972). *American university word list*. Longman.
- Rayson, P., & Garside, R. (2000). *Comparing corpora using frequency profiling proceedings of the workshop on comparing corpora* (pp. 1-6). Hong Kong: Hong Kong: Association for Computational Linguistics.
- Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.
- Richards, J. (1974). Word lists: Problems and prospect. *RELC Journal*, 5(2), 69-84.
- Ryan, G. (2012). *Technical vocabulary acquisition through texts. A corpus and a case study in theology classroom*. Lambert Academic Publishing.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 199-227). Cambridge University Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research Manual*. Palgrave Macmillan.
- Schmitt, N., & Schmitt, D. (2014). A Reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484-503.
- Scott, M. (2001). Mapping key words to problem and solution. In M. Scott & G. Thompson (Eds.), *Patterns of text. In honour of Michael Hoey* (pp. 109-128). John Benjamins.

- Scott, M. (2008). *Oxford WordSmith Tools 5.0*. Liverpool: Lexical Analysis Software.
- Skytrax. (2018, April, 1). *The top 100 airlines of 2017*. <http://www.airlinequality.com/info/top-100-airlines-2017/>.
- Smith, S. (2020). DIY corpora for Accounting & Finance vocabulary learning. *English for Specific Purposes*, 57(2020), 1-12.
- Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP classroom. *RELC Journal*, 45(1), 51-65.
- Todd, R. W. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes*, 45(1), 31-39.
- Tongpoon-Patanasorn, A. (2017). Constructing an academic word list of business English: A corpus-based approach. *Humanities and Social Sciences Journal*, 34(2), 1-31.
- United Nations World Tourism Organization. (2018, April, 7). *UNWTO tourism highlights*. <http://mkt.unwto.org/publication/unwto-tourism-highlights>.
- Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *English for Specific Purposes*, 12(4), 248-263.
- Vo, T. D., & Jaturapitakkul, N. (2016). The use of vocabulary learning strategies by Thai EFL learners studying Vietnamese as a third language. *Language Education and Acquisition Research Network (LEARN) Journal*, 9(2), 105-121.
- Vongpumivitch, V., Huang, J., & Chang, Y. (2009). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes*, 28(1), 33-41.
- Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical academic word list. *English for Specific Purposes*, 27(4), 442-458.
- Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, 28(3), 170-182.
- West, M. (1953). *A general service list of English words*. Longman.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.
- World Travel and Tourism Council. (2019, December, 20). *Economic impact reports*. <https://wtcc.org/Research/Economic-Impact>.

- Xue, G., & Nation, P. (1984). A university word list. *Language learning and communication*, 3, 215-229.
- Yang, M. (2015). A nursing academic word list. *English for Specific Purposes*, 37(1), 27-38.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2. *Journal of Second Language Writing*, 13(4), 257-283.
- Zipf, G. (1935). *The psycho-biology of language*. MIT Press.

## APPENDIX 1

The 378 words of the Tourism Business Word List (TBWL) separated into 13 sub-word lists.

### The 1st sub-TBWL

- |                   |             |             |               |
|-------------------|-------------|-------------|---------------|
| 1. accommodate    | 2. airport  | 3. arrive   | 4. atmosphere |
| 5. available      | 6. bay      | 7. book     | 8. capital    |
| 9. convenience    | 10. cuisine | 11. culture | 12. depart    |
| 13. executive     | 14. express | 15. guest   | 16. holiday   |
| 17. international | 18. Ideal   | 19. journey | 20. leisure   |
| 21. lounge        | 22. luxury  | 23. offer   | 24. private   |
| 25. region        | 26. relax   | 27. request | 28. reserve   |
| 29. serve         | 30. service | 31. tour    | 32. transport |
| 33. travel        | 34. trip    | 35. visit   | 36. welcome   |

### The 2nd sub-TBWL

- |                |              |             |             |
|----------------|--------------|-------------|-------------|
| 37. amaze      | 38. animal   | 39. art     | 40. autumn  |
| 41. bar        | 42. bath     | 43. beach   | 44. camp    |
| 45. church     | 46. city     | 47. climb   | 48. country |
| 49. delicious  | 50. discover | 51. double  | 52. east    |
| 53. experience | 54. fair     | 55. farm    | 56. fly     |
| 57. forest     | 58. found    | 59. garden  | 60. green   |
| 61. hall       | 62. hill     | 63. history | 64. hour    |
| 65. huge       | 66. internet |             |             |

---

---

### The 3rd sub-TBWL

67. island	68. lake	69. local	70. mileage
71. mountain	72. nature	73. outdoor	74. park
75. place	76. plan	77. plenty	78. reach
79. rent	80. rich	81. river	82. rock
83. safe	84. sail	85. sea	86. secure
87. shop	88. spacious	89. spot	90. spring
91. square	92. station	93. stay	94. stone
95. store	96. street		

### The 4th sub-TBWL

97. town	98. track	99. train	100. view
101. weather	102. wide	103. wild	104. woods
105. access	106. admire	107. advance	108. adventure
109. attend	110. attract	111. beer	112. brick
113. bridge	114. cable	115. castle	116. century
117. classic	118. cliff	119. coast	120. contact
121. decorate	122. delight	123. desert	124. distance
125. district	126. dive		

### The 5th sub-TBWL

127. environment	128. event	129. expense	130. famous
131. fantastic	132. fascinate	133. feature	134. flag
135. folk	136. gate	137. gather	138. giant
139. gift	140. golf	141. gorgeous	142. guide
143. hire	144. hotel	145. impress	146. improve
147. incredible	148. invite	149. modern	150. mount
151. mud	152. official	153. organize	154. path
155. period	156. pine		

### The 6th sub-TBWL

157. planet	158. policy	159. pool	160. popular
161. port	162. rare	163. royal	164. restaurant
165. recommend	166. saint	167. scenic	168. schedule

169. season	170. shelter	171. shore	172. site
173. surroundings	174. ski	175. solid	176. tent
177. theatre	178. ticket	179. tip	180. tower
181. tradition	182. traffic	183. typical	184. valley
185. various	186. vehicle		

### The 7th sub-TBWL

187. village	188. wander	189. abroad	190. agriculture
191. archaeological	192. airline	193. ancient	194. annual
195. architect	196. border	197. budget	198. capacity
199. carve	200. celebrate	201. ceremony	202. concert
203. conservation	204. cruise	205. craft	206. contemporary
207. currency	208. dedicate	209. display	210. diverse
211. domestic	212. era	213. enhance	214. emergency
215. exhibit	216. explore		

### The 8th sub-TBWL

217. extraordinary	218. facilitate	219. fee	220. festival
221. gallery	222. goods	223. harbor	224. heritage
225. highlight	226. holy	227. host	228. ingredient
229. inhabitant	230. insight	231. inspire	232. interior
233. landscape	234. legend	235. lodge	236. marine
237. numerous	238. museum	239. ocean	240. overlook
241. pace	242. palace	243. palm	244. participate
245. passion	246. peak		

### The 9th sub-TBWL

247. platform	248. preserve	249. prior	250. province
251. remote	252. republic	253. resort	254. resource
255. retreat	256. route	257. rural	258. sculpture
259. stun	260. summit	261. symbol	262. territory
263. theme	264. trail	265. treasure	266. ultimate
267. unique	268. urban	269. vast	270. arch
271. array	272. authentic	273. avenue	274. bathe
275. bronze	276. café		

### The 10th sub-TBWL

277. calendar      278. canal      279. cathedral 280. cave  
281. dynamic      282. escort      283. exotic      284. ferry  
285. fort      286. habitat      287. harvest      288. infrastructure  
289. indigenous      290. inn      291. marble      292. magnificent  
293. monument      294. mineral      295. overnight 296. pearl  
297. recreation      298. refresh      299. romance 300. sacred  
301. spectacular      302. soak      303. statue      304. steep  
305. stroll      306. superb

### The 11th sub-TBWL

307. temple      308. terrace      309. tropical      310. venue  
311. villa      312. volcanic      313. altitude      314. courtyard  
315. feast      316. globe      317. hike      318. iconic  
319. inland      320. jewel      321. landmark 322. memorable  
323. paradise      324. passport      325. peninsula 326. renovate  
327. renown      328. sanctuary 329. spa      330. surf  
331. terrain      332. transit      333. vacation      334. visa  
335. wilderness      336. dune

### The 12th sub-TBWL

337. ecosystem      338. elegance      339. excursion      340. flora  
341. hospitality      342. hub      343. lush      344. plateau  
345. refund      346. trek      347. vibrant      348. boutique  
349. canyon      350. culinary      351. fauna      352. lagoon  
353. majestic      354. motel      355. pristine      356. picturesque  
357. panorama      358. gourmet      359. itinerary      360. kayak  
361. scuba      362. backpack      363. campsite      364. breathtaking  
365. coastline      366. countryside

### The 13th sub-TBWL

367. downtown      368. limestone      369. nightlife      370. underground  
371. sightseeing      372. sunset      373. oneway      374. underwater  
375. waterfall      376. wellness      377. wildlife      378. wheelchair

**APPENDIX 2**

The 274 words of the Hotel Business Word List (HBWL) separated into 9 sub-word lists.

**The 1st sub-HBWL**

1. accommodate	2. airport	3. arrive	4. atmosphere
5. available	6. bay	7. book	8. capital
9. convenience	10. cuisine	11. culture	12. depart
13. executive	14. express	15. guest	16. holiday
17. international	18. Ideal	19. journey	20. leisure
21. lounge	22. luxury	23. offer	24. private
25. region	26. relax	27. request	28. reserve
29. serve	30. service	31. tour	32. transport
33. travel	34. trip	35. visit	36. welcome

**The 2nd sub-HBWL**

37. amaze	38. art	39. bar	40. base
41. bath	42. beach	43. beauty	44. bed
45. breakfast	46. bring	47. business	48. centre
49. check	50. club	51. coffee	52. collect
53. comfort	54. cook	55. couple	56. delicious
57. dinner	58. discover	59. done	60. double
61. drink	62. excite	63. experience	64. floor
65. garden	66. heart		

**The 3rd sub-HBWL**

67. history	68. indoor	69. inform	70. island
71. lake	72. local	73. lunch	74. mountain
75. nature	76. outdoor	77. park	78. rate
79. rich	80. river	81. rock	82. room
83. sign	84. spacious	85. special	86. square
87. stay	88. table	89. treat	90. view
91. wedding	92. access	93. attract	94. benefit
95. casual	96. ceiling		

---

---

### The 4th sub-HBWL

97. cheese	98. classic	99. coast	100. complain
101. contact	102. create	103. deck	104. decorate
105. design	106. desk	107. dish	108. due
109. entertain	110. event	111. extend	112. favour
113. feature	114. golf	115. guide	116. hotel
117. impress	118. improve	119. include	120. incredible
121. invite	122. item	123. locate	124. modern
125. partner	126. pool		

### The 5th sub-HBWL

127. property	128. provide	129. rare	130. register
131. restaurant	132. scenic	133. shower	134. site
135. standard	136. surround	137. tradition	138. twin
139. valley	140. valuables	141. various	142. approximate
143. architect	144. award	145. blend	146. celebrate
147. corporate	148. custom	149. craft	150. contemporary
151. dedicate	152. distinct	153. diverse	154. expand
155. explore	156. extraordinary		

### The 6th sub-HBWL

157. facility	158. fee	159. gallery	160. heritage
161. ingredient	162. host	163. innovate	164. inquire
165. inspire	166. interior	167. intimate	168. landscape
169. legend	170. menu	171. ocean	172. occupancy
173. overlook	174. passion	175. reception	176. reside
177. resort	178. retreat	179. stun	180. sophisticate
181. trail	182. ultimate	183. unique	184. vast
185. adjacent	186. array		

### The 7th sub-HBWL

187. authentic	188. boast	189. chef	190. champagne
191. destination	192. exotic	193. fare	194. furnish
195. magnificent	196. laundry	197. marble	198. premier

199. premium	200. refine	201. refresh	202. romance
203. signature	204. sparkling	205. stroll	206. Spectacular
207. suite	208. terrace	209. venue	210. villa
211. balcony	212. butcher	213. cocktail	214. complimentary
215. exquisite	216. gym		

### The 8th sub-HBWL

217. iconic	218. massage	219. pastry	220. Memorable
221. renown	222. sanctuary	223. shuttle	224. soothe
225. spa	226. tranquil	227. tub	228. vacation
229. amenity	230. banquet	231. elegance	232. excursion
233. hospitality	234. lush	235. vacancy	236. vibrant
237. beverage	238. boutique	239. culinary	240. majestic
241. panorama	242. oasis	243. pristine	244. Picturesque
245. butler	246. gourmet		

### The 9th sub-HBWL

247. sumptuous	248. yoga	249. unwind	250. utensil
251. cutlery	252. deluxe	253. valet	254. toiletries
255. brunch	256. concierge	257. bellhop	258. appetizer
259. bathroom	260. ballroom	261. babysit	262. bartender
263. breathtaking	264. busser	265. doorman	266. fireplace
267. housekeeper	268. getaway	269. lifestyle	270. medium
271. sunset	272. walk-in	273. wellness	274. Wildlife

## APPENDIX 3

The 245 words of the Airline Business Word List (ABWL) separated into 8 sub-word lists.

### The 1st sub-ABWL

1. accommodate	2. airport	3. arrive	4. atmosphere
5. available	6. bay	7. book	8. capital
9. convenience	10. cuisine	11. culture	12. depart
13. executive	14. express	15. guest	16. holiday

17. international	18. Ideal	19. journey	20. leisure
21. lounge	22. luxury	23. offer	24. private
25. region	26. relax	27. request	28. reserve
29. serve	30. service	31. tour	32. transport
33. travel	34. trip	35. visit	36. welcome

### The 2nd sub-ABWL

37. allow	38. bag	39. baggage	40. base
41. board	42. business	43. call	44. card
45. carrier	46. carry	47. charge	48. class
49. clearance	50. comfort	51. country	52. danger
53. engine	54. first	55. fly	56. flyer
57. hour	58. inform	59. land	60. load
61. middle	62. mile	63. pass	64. prepare
65. rate	66. report		

### The 3rd sub-ABWL

67. responsible	68. return	69. rule	70. safe
71. seat	72. secure	73. sign	74. space
75. special	76. store	77. tax	78. weather
79. window	80. adult	81. advance	82. advice
83. agent	84. alcohol	85. assist	86. attendant
87. battery	88. bound	89. captain	90. cart
91. chief	92. claim	93. commit	94. deck
95. deliver	96. direct		

### The 4th sub-ABWL

97. distance	98. duty	99. economy	100. gate
101. identify	102. Individual	103. instruct	104. instrument
105. item	106. length	107. log	108. loss
109. mask	110. medical	111. minor	112. operate
113. organize	114. pat	115. plane	116. pocket
117. port	118. provide	119. refuse	120. recommend
121. remain	122. require	123. respect	124. row
125. schedule	126. seal		

### The 5th sub-ABWL

127. senior	128. spare	129. ticket	130. tower
131. traffic	132. transfer	133. tray	134. accompany
135. aircraft	136. airline	137. annual	138. approve
139. assign	140. capacity	141. charter	142. climate
143. companion	144. code	145. consent	146. compensate
147. corporate	148. consult	149. crew	150. cruise
151. currency	152. custom	153. delay	154. device
155. disabled	156. domestic		

### The 6th sub-ABWL

157. electronic	158. infant	159. excess	160. facilitate
161. emergency	162. jet	163. liquid	164. numerous
165. passenger	166. permit	167. pilot	168. proceed
169. prohibit	170. republic	171. restrict	172. route
173. specify	174. update	175. weigh	176. zone
177. cabin	178. cargo	179. carriage	180. comply
181. destination	182. duration	183. escort	184. evacuate
185. exit	186. fare		

### The 7th sub-ABWL

187. fleet	188. haul	189. immigrate	190. overhead
191. premium	192. strap	193. tag	194. terminal
195. upgrade	196. airways	197. aisle	198. altitude
199. automate	200. aviation	201. brace	202. compartment
203. congestion	204. fasten	205. disarm	206. Complimentary
207. notify	208. passport	209. ramp	210. runway
211. transit	212. turbulent	213. vent	214. visa
215. cockpit	216. hub		

### The 8th sub-ABWL

217. luggage	218. portable	219. refund	220. galley
221. aft	222. itinerary	223. lavatory	224. recline

225. disembark	226. deplane	227. purser	228. decompress
229. airbus	230. airside	231. armrest	232. copilot
233. headset	234. inflight	235. jumpseat	236. landside
237. layover	238. legroom	239. onboard	240. pregnant
241. seatback	242. seatbelt	243. takeoff	244. taxiway
245. wheelchair			