Lexical Input Used for English for Tourism Coursebooks at Tertiary Level: A Corpus-Based Approach

Anchalee Veerachaisantikul[1], Wara Chansin[2], Kamontip Nuamkoksoong[3]

[1] *Rajamangala University of Technology Isan, Thailand. e-mail: anchalee.ve@rmuti.ac.th*
[2] *Naresuan University, Phitsanulok. e-mail: warac@nu.ac.th*
[3] *Rajamangala University of Technology Isan, Thailand. e-mail: kamontip.nu@rmuti.ac.th*

ARTICLE INFO	ABSTRACT
Keywords: corpus-based approach, Coursebooks, English for tourism, lexical input, tertiary level **DOI:** http://dx.doi.org/10.210 93/ijeltal.v10i2.2129	*Effective English for Specific Purposes (ESP) training depends on coursebooks for relevant and genuine vocabulary. Corpus-based analysis is a reliable empirical technique for lexical evaluation in ESP coursebook assessment, as shown in this research. Thus, this quantitative corpus-based research focused on (1) the most common general English lexical items and (2) domain-specific vocabulary in tertiary English for Tourism coursebooks. The study was a quantitative corpus-based one that included 3,423,973 words from all primary reading materials, dialogues, and exercises in the 2020–2022 English for Tourism coursebooks. Using corpus linguistics methods, WordSmith Tools discovered and evaluated lexical frequency lists and domain-specific keywords. This study found that, although generic high-frequency terms are important, tourism-related language dominates the materials' specialized vocabulary. Generally, high-frequency lexical terms, essential for language development, were detected. The most common domain-specific vocabulary, mostly nouns, was tourism terminology, including "guest," "hotel," "room," "travel," and "tourism." This study provided actual lexical data to help curriculum developers and teachers plan, choose, and adjust tourist education resources. This aligns with real-world language needs and improves ESP instruction.*

1. Introduction

With the increasing global connectedness, English has become the preferred language for communication in various industries, particularly the dynamic domains of travel and hospitality. There is an increasing need for professionals who can communicate effectively in English, particularly in positions related to tourism, as the global tourist industry continues to experience steady growth and makes a substantial contribution to economic expansion.

Because of this tendency, the relevance of English for Specific Purposes (ESP) in higher education has been critically emphasized. In particular, an emphasis has been given to English for Occupational Purposes (EOP), which includes English for Tourism (Basturkmen, 2010; Dudley-Evans & St. John, 1998; Hsu, 2009). ESP stands for English for Specific Purposes. It cannot be denied that selecting and promoting appropriate vocabulary is one of the most important aspects of teaching English to non-native speakers. This is due to the fact that lexical knowledge is required for communicative competence, particularly in professional contexts where linguistic accuracy and appropriateness are of utmost importance (Basturkmen, 2010; Nation, 2006; Coxhead, 2021).

Although commercial textbooks are widely used in English for Specific Purposes (ESP) schools, scholars have persistently questioned the degree of linguistic correctness with which these materials match the actual language that is used in true industrial contexts. Many textbooks simplify content to make it more accessible to teachers, which often leads to a failure in adequately representing the complex linguistic requirements that tourism professionals face in various contextual settings. It has been shown in many studies that coursebooks often simplify the information, and as a result, they fail to convey the actual language complexities that are frequently experienced by tourist professionals (Dudley-Evans and St. John, 1998; Flowerdew, 2013). Moreover, Afzalia (2014) and Blanka (2015) have highlighted the significant role that rich vocabulary knowledge plays in ESP, especially within professional domains like tourism, where the exact use of specialized language has a direct influence on the quality of service and, ultimately, the happiness of customers (López-Rocha, 2020).

Following the occurrence of these issues, future corpus-based research has changed its focus to evaluating the lexical content of ESP materials and the language that is associated with tourism. As an example, comprehensive research conducted by Wang and Wang (2017) highlights the growing use of corpus-based techniques in the process of enhancing the accuracy and relevance of textbooks in the field of language instruction. Studies have also focused on creating specialized word lists from genuine tourism corpora, such as Tourism Business Word Lists (Laosrirattanachai & Ruangjaroon, 2021) and lexical collocation analyses in tourism research publications (Trinant & Kijpoonphol, 2021). Furthermore, preliminary attempts have been undertaken to assess the vocabulary and graphics in certain English for Tourism textbooks (e.g., Al-Busaidi, 2022, examining earlier versions such as English for International Tourism, 2003, and Marketing for Hospitality and Tourism, 2021). Other research focuses on the needs of tourist workers, consistently identifying communication issues, such as limited vocabulary, as a key obstacle (Namtapi, 2022), proving the need for suitable teaching resources. Furthermore, Wagner (2023) analyzes curriculum and materials design in ESP for tourism, emphasizing the challenge of selecting relevant content. These studies, although informative, often focus on specific linguistic features, regional contexts, or materials published before the most recent curricular breakthroughs.

Despite this considerable body of work, a key empirical gap persists: comprehensive, detailed, and current corpus-based assessments of lexical input in recently released (2020-2022) English for Tourism coursebooks designed for tertiary-level education are noticeably absent. Even while past research has shown the benefits of corpus linguistics, not all have provided a complete, up-to-date lexical profile of the resources now in use. Therefore, no previous study has extensively examined the domain-specific vocabulary and generic English

lexical words in a contemporary corpus of English for Tourism coursebooks to assess how well they correspond to language demands in the real world. Because of this gap, teachers and curriculum authors find it challenging to make fully data-informed decisions on the most appropriate and authentic lexical material for their students.

In this context, advancements in corpus linguistics have enabled the creation of effective and equitable tools for studying language. Corpus-based methodologies are becoming more and more common in ESP research, especially when it comes to studying how language is used at school and work. These tools allow you to see many details about word frequency, collocations, and usage trends. We may compare the language in textbooks to both major general corpora like the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA), as well as smaller, more specialized corpora (McEnery & Hardie, 2012). Previous research found that corpora are an excellent way to assess how well training materials fit with people's real language use. For example, Crosthwaite and Cheung (2019) investigated the possibility of corpus technologies to compare textbook material to real utterances made by individuals in disciplinary settings. They found that the terminology and patterns of usage were quite different, which might make it harder to prepare students for real-life encounters. Huynh & Nguyen (2020) also undertook similar research that looked at the differences between the language used in ESP textbooks and the language used in actual jobs. The survey discovered that most of the course materials do not include the words and phrases that people need for different jobs. These studies demonstrate that employing corpus-driven evaluations may assist in finding missing words and make English for Specific Purposes (ESP) more useful.

The current study tries to fill the gap to what we already know about the language employed in English literature for tourism textbooks. Even though corpus linguistics is being used more and more for lexical analysis in a number of fields, there is still a need for more focused research on the specific terms that can be found in English for Tourism textbooks that are currently being used in schools for tourism. Our work used a corpus-based technique to close this gap and build on the well-acknowledged advantages of corpus linguistics. The primary objective was to identify commonly used tourism-specific vocabulary and explore its prevalence and occurrence in general-purpose corpora from real-world visitor settings. The findings aim to provide textbook authors, curriculum developers, and instructors with the necessary information and support to develop materials that are effectively aligned with the language requirements of students pursuing professions in tourism. The final objective of this program is to bridge the current divide between academic learning and the linguistic requirements of real-world professionals by enhancing vocabulary education and better equipping students to meet the specialized language requirements of the tourism industry. The research questions is: "What are the most frequently occurring words in English for tourism coursebooks?"

2. Literature Review

2.1 English for Tourism

The globalization of communication has made English the dominant language in many industries, including travel and hospitality. Crystal (1997) supports this concept, stating that 337 million individuals spoke English as their primary language and 235 million as a second

language in 1995. This is why it has become the global language of business and technology, making it easier for salespeople and technical professionals to communicate. In addition, English for Specific Purposes (ESP) has seen remarkable growth because of the growing need for English in various contexts, including academic and professional ones. According to Basturkmen (2010), two of the subfields that are seeing expansion are English for Occupational Purposes (EOP) and English for Tourism Purposes (EFT), which are gaining popularity.

English for Tourism (EFT) is a program that assists students in developing the language skills necessary for success in the tourist industry within ESP, as suggested by Basturkmen (2010) and Dann (1996). Communication skills are improved via the use of EFT programs in locations such as hotels, travel agencies, customer service, and tour guides. In their respective works, Gilmore (2007) and Harwood (2005) highlighted the importance of the learner-centered approach that ESP and EFT take in teaching a second language. Furthermore, according to Gilmore (2007), ESP is a needs-based program that adapts both its material and its methodology to meet the requirements of the students. In addition, EFT, like other specialist languages (Bühler, 2011), has specific roles, structures, and tenses; thus, tourism language teaching must be tailored to its needs. In addition, the distinctiveness of tourist vocabulary is a standard collection of symbols and codes used to communicate business information (Dann, 1996). This specialty language must be used accurately and properly in various circumstances, including hotel check-ins, facility and schedule information, directions, and tourist information exchange. Therefore, to communicate effectively with tourists, one must have a comprehensive awareness of the language components that are special to tourism (Bühler, 2011).

2.2 The Importance of Lexical Input in ESP

Nation and Newton (1997) expressed that vocabulary is a key item of foreign language communication. It also directly influences English language students' comprehension and expression. When students need more than simply a broad grasp of the English language, vocabulary is an important part of English for Specific Purposes (ESP). Furthermore, ESP proficiency is knowledge of technical terminology, common collocations, and language relevant to one's employment or study (McEnery & Wilson, 1996). ESP students must alter their lexical intake to meet complex communication needs, which differ from standard English. Students must focus on language to communicate effectively and responsibly in their profession or studies.

2.3 Word Frequency

Given the foundational role of vocabulary, educators bear a responsibility to provide relevant lexical information and elucidate its practical application (Blue & Harun, 2000). A cornerstone of effective vocabulary instruction is the strategic sequencing of lexical input. It is pedagogically imperative to ensure that students possess a solid understanding of high-frequency words before proceeding to less frequent vocabulary, as mastery of the former offers substantial advantages for overall language comprehension and production. Moreover, the efficacy of vocabulary acquisition and learner attitudes can be significantly enhanced through the thoughtful design of vocabulary and other skill-development exercises. This necessitates that educators meticulously consider the appropriate language level for their students, aligning lexical challenges with existing proficiency. Linguistic

research on vocabulary consistently demonstrates that acquiring knowledge of the most prevalent and frequently utilized words yields invaluable insights for students. Consequently, robust word frequency data profoundly influences syllabus and material design, underscoring that the most common words in the English language are often the most beneficial for both students and course developers.

2.4 High-Frequency Words

High-frequency words are essential for language mastering, indicating a class of keywords whose absence makes traditional usage unfeasible. These words deserve special attention in language teaching because they often appear in both spoken and written English. Moreover, word frequency in reading texts can be divided into four categories: high-frequency words, academic or sub-technical words, technical words, and low-frequency words (Mackey & Mountford, 1978). The teaching processes for high-frequency words differ greatly from those for low-frequency words; as a result, both teachers and students must attend training that focuses on these terms and their acquisition strategies. Besides pre-teaching, teachers can help students learn high-frequency terminology by introducing new words into textbook-based assignments and examining their multiple meanings. Also, all English students must study high-frequency phrases.

2.5 Challenges of Authentic Language Representation in ESP Textbooks

When it comes to providing organized language input, ESP classes continue to depend primarily on textbooks. Many concerns have been raised by academics regarding the degree to which these materials correspond to language usage in the actual world (Dudley-Evans & St. John, 1998). Scholars such as Flowerdew (2013) have categorically claimed that a significant number of English as a Second Language (ESL) textbooks are either too generic or out of date, and as a result, they do not possess the linguistic authenticity that is necessary for effective professional communication. Due to the apparent obsolescence of the language taught in textbooks and the language used in real life, several academics have been motivated to use corpus-based approaches to rigorously assess and enhance the quality of the content that is taught in classes.

2.6 Vocabulary Coverage and Comprehension

As students prepare to join the professional world, the language in their textbooks must represent actual, industry-relevant terminology. Previous research, as proposed by Kennedy (1998), showed that specialist corpora, such as the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), or tourism-specific corpora, are useful reference points for assessing the sufficiency of textbook material. Lexical coverage could improve reading comprehension and communication beyond exposure. To grasp a book, students must understand 95-98% of the words. If tourism textbooks fail to address this lexical barrier, students may struggle to use their language abilities in professional situations (McEnery & Hardie, 2012). Thus, a thorough lexical profile evaluation of these resources is required to ensure linguistic correctness and improve language teaching performance.

2.7 Corpus-Based Approach to Textbook Analysis

A corpus is a collection of naturally occurring linguistic material collected from either written texts or transcriptions of recorded speech that serves as a foundation for linguistic analysis and description (Nation & Newton, 1997; Kennedy, 1998; Biber, Conrad, & Reppen, 1998).

Linguistic data is frequently stored on microfiches, printed volumes, or other computer-readable media.

Furthermore, corpus-based studies play a vital role in helping students develop their vocabulary, which is promptly converted into useful classroom input material, allowing instructors and course planners to obtain information from highly reliable sources. The presence of corpus-based word lists, for example, ensures course creators that the vocabulary used is appropriate and valuable to students. Furthermore, it supplies detailed pragmatic information on language use.

In conclusion, corpus-based methods are recommended for assessing and improving lexical input in English for Specific Purposes (ESP) materials. When it comes to tourism, communication that is both specialized and authentic is essential. Based on these guidelines, this study tries to overcome the discrepancy between true tourist communication and educational language by revealing tertiary English for tourists' lexical needs.

3. Research Methodology

3.1 The Source of Data for the Corpus

This research aimed to construct a corpus of English lexical input from tertiary-level course materials to assess the lexical profile of contemporary English for Tourism coursebooks. This collection constituted the foundation for further lexical analysis and evaluation, directly addressing the study's research goals.

The corpus data were collected from a meticulously selected array of tertiary-level English for Tourism textbooks released between 2020 and 2022. To ensure the findings were globally representative and pertinent, only coursebooks from reputable publishers specializing in English for Specific Purposes (ESP) and English Language Teaching (ELT) materials were used. These materials had to be explicitly dedicated to English for Tourism, covering key thematic areas such as hospitality, travel operations, cultural tourism, and customer service. Furthermore, they were specifically chosen for their design for university-level students enrolled in specialized English for Tourism programs or related undergraduate degrees that include a significant English for Tourism component. Their employment as essential teaching resources in educational institutions throughout the 2020-2022 academic year demonstrated their pedagogical value. To provide a complete picture of lexical input, the corpus contained material from all main reading passages, debates, and activities in each coursebook.

3.2 Tools of Analysis

This study employed a systematic corpus-based approach to analyze the lexical input of English for Tourism coursebooks, utilizing a combination of specialized software and established linguistic resources to identify and categorize vocabulary.

3.2.1 Selecting Coursebooks

The COVID-19 pandemic, which impacted universities between 2020 and 2022, made it very difficult to select and obtain coursebooks for the English for Tourism subject. Most institutions were closed or had limited access at that time. In this very chaotic situation, the Thai Qualifications Framework for Higher Education (TQF) was a vital guide. In situations where it was not always feasible to check out educational resources on campus, it provided

precise standards for judging the quality of such items and how they should be used. There were two methods that were used in order to ascertain which textbooks were being utilized. Formal inquiries were dispatched to university departments providing tourism or associated courses where direct communication or public postings were impractical. Concurrently, coursebook information was obtained from publicly accessible sources that were often immediately aligned with the course criteria stipulated by TQF. This careful technique guaranteed the systematic collection of all primary reading materials, debates, and activities. The study's corpus was subsequently established by a thorough analysis of these utilizing TQF-informed criteria.

3.2.2 Concordance Software

This study utilized WordSmith Tools (version 8) to perform the primary lexical analyses. The Wordlist module was employed to generate raw frequency lists of all lexical items present in the compiled English for Tourism coursebook corpus, providing initial counts of general high-frequency and potential domain-specific words. Subsequently, the KeyWords module was instrumental in identifying statistically significant tourism-specific vocabulary by comparing the frequency of words in the coursebook corpus against a large general reference corpus. The Concord module was used for contextual verification of identified terms, allowing for a qualitative examination of their usage patterns within the coursebook texts.

3.2.3 The Oxford Advanced Learner's Dictionary (OALD)

The Oxford Advanced Learner's Dictionary (OALD, 10th edition) served as a key reference tool during the manual verification and classification of identified lexical items. It was specifically consulted to confirm parts of speech and primary general meanings of high-frequency and key words, assisting in accurate categorization, particularly when disambiguation was required based on context.

3.2.3. The Oxford Dictionary of Travel and Tourism

The Oxford Dictionary of Travel and Tourism was critically utilized for validating and contextualizing the identified domain-specific tourism vocabulary. This specialized dictionary provided essential definitions and insights into operational terms, acronyms, and industry-specific terminology found within the coursebook corpus, ensuring the accurate classification and interpretation of specialized terms.

3.2.4 Collins COBUILD Bank of English (BoE)

The Collins COBUILD Bank of English (BoE), a sub-corpus of 650 million words extracted from the larger 4.5 billion-word COBUILD corpus, served as the primary general reference corpus for the KeyWords analysis within WordSmith Tools. Its diverse range of written and spoken materials allowed for the statistical identification of domain-specific lexical items in the English for Tourism coursebook corpus by highlighting words that were unusually frequent in the specialized corpus compared to general English use.

3.2.5 General Service List (GSL)

The General Service List (GSL), specifically the revised edition by Bauman and Culligan, was employed as a benchmark to distinguish high-frequency general English lexical words from more specialized or less common vocabulary within the coursebook corpus. Words found

within the GSL were categorized as foundational, high-frequency items, enabling a clear differentiation in the lexical analysis.

3.3 Research Procedures

This study employed a two-stage methodology, encompassing data collection and lexical analysis. These two main phases involved a systematic series of procedures designed to investigate the English thoroughly for tourism coursebooks.

Stage 1: Data Collection

The first part of this research included a systematic selection of English for Tourism coursebooks produced between 2020 and 2021. This research started by selecting appropriate undergraduate tourism programs at several Thai institutions. Initially, attempts were made to contact colleges directly via formal letters to ask about their recommended textbooks and course materials. However, this direct method was rendered untenable because of the COVID-19 epidemic that swept throughout the whole world and also caused colleges to close their doors. The researchers altered their approach by carefully examining online articles from the Thai Qualifications Framework for Higher Education (TQF) to circumvent the issue that they were currently studying. Several public and private colleges in Thailand's northern, central, northeastern, and southern regions provided these publicly available materials. They offered comprehensive course materials, including lists of assigned or suggested literature. This thorough review revealed that course books from three major publishers—Oxford University Press, Longman, and Thammasat University Press—were the most frequently prescribed or referenced materials in the identified programs. These meticulously selected sources subsequently formed the foundation for the tourism corpus developed in this study.

Stage 2: Lexical Analysis and Corpus Preparation

The corpus of English for Tourism coursebooks for the 2020-2021 academic year was rigorously prepared for lexical analysis, building upon the original data collection. A total of 15 unique coursebooks were deliberately chosen to provide the core dataset for this investigation.

 The content of these coursebooks was initially converted to Microsoft Word (docx) format to ensure compatibility and prevent technical complications with the selected corpus analysis tools. It was necessary to conduct comprehensive editing and rectification of typographical errors and linguistic irregularities in the Word documents during the conversion process to enhance the quality of the data. Subsequently, these refined papers were transformed into plain text (.txt) files, which served as the input for the lexical analytical process. The lexical analysis used a corpus-based approach in lexical semantics to determine the unique words and language patterns that are important to the tourism industry (e.g., Stubbs, 2001; Baker, 2006). This strategy must focus on how language is actually used, as evidenced in the corpus, rather than relying on definitions that aren't tied to a specific context.

We used a customized English for Tourism dictionary to ensure that lexical recognition and interpretation were correct and useful in the tourist environment. The Oxford Dictionary of Travel and Tourism is the most trusted source for this vocabulary. It covers all the terms used in the tourist sector. Subsequently, the analysis was conducted using WordSmith Tools. Concordances and word frequency lists were produced from the plain text files that were

prepared using this tool. The Wordlist Tool function in WordSmith Tools was utilized to organize the vocabulary systematically from the coursebooks in both alphabetical and frequency-based order. This approach provided a fundamental understanding of the lexical landscape of the corpus. This initial phase facilitated the identification of high-frequency items and served as a foundation for additional targeted analyses, including keyword and collocation analysis.

4. Results

This section contains 2020-2021 English for Tourism coursebook corpus lexical analysis results. This analysis sought to answer the research question, "What are the most frequently occurring words in English for tourism coursebooks?" by identifying and characterizing the most common lexical elements in these instructional materials.

4.1. Lexical Profile of the English for Tourism Coursebook Corpus (TC)

The initial frequency analysis provided a foundational understanding of the most common lexical items present in the English for Tourism coursebooks. Table 1 presents the top 100 most frequently occurring words in the Tourism Corpus (TC).

Table 1: The Top 100 Most Frequently Occurring Words of the Tourism Corpus (TC)

Rank	Word	F	%	Rank	Word	F	%
1	THE	9,819	5.42	51	SERVICE	389	0.21
2	A	9,196	5.07	52	HOLIDAY	386	0.21
3	BE	4,857	2.68	53	WHERE	365	0.20
4	TO	4,722	2.34	54	YES	363	0.20
5	AND	4,442	2.45	55	TWO	362	0.20
6	YOU	3,650	2.01	56	EACH	360	0.20
7	OF	3,517	1.94	57	INFORMATION	347	0.19
8	IN	3,354	1.85	58	WHEN	346	0.19
9	FOR	1,747	0.96	59	MORE	343	0.19
10	YOUR	1,411	0.78	60	LISTEN	342	0.19
11	I	1,244	0.69	61	ALL	341	0.19
12	ON	1,156	0.64	62	GOOD	331	0.18
13	AT	1,129	0.62	63	THINK	325	0.18
14	CAN	1,118	0.62	64	OUT	325	0.18
15	DO	1,102	0.61	65	THEM	320	0.18
16	THAT	1,091	0.60	66	GET	320	0.18
17	WITH	1,062	0.59	67	TAKE	317	0.17
18	WHAT	1,060	0.58	68	EMPLOYEE	313	0.17
19	OR	963	0.53	69	SO	310	0.17
20	IT	960	0.53	70	SOME	285	0.16
21	HAVE	925	0.51	71	MAKE	280	0.15
22	GUEST	882	0.48	72	ANY	279	0.15
23	HOTEL	872	0.48	73	TOUR	276	0.15
24	FROM	773	0.43	74	TIME	270	0.15
25	ABOUT	734	0.40	75	THEN	268	0.15
26	WE	703	0.39	76	MOST	264	0.15
27	THIS	672	0.37	77	COUNTRY	264	0.15
28	ROOM	635	0.35	78	JOB	263	0.15
29	IF	605	0.33	79	NEW	253	0.14
30	WHICH	590	0.33	80	FLIGHT	252	0.14
31	THESE	554	0.31	81	TOURIST	251	0.14
32	AN	540	0.30	82	THEIR	250	0.14

Rank	Word	F	%	Rank	Word	F	%
33	WORK	538	0.30	83	ASK	247	0.14
34	AS	533	0.29	84	PLEASE	245	0.14
35	HOW	522	0.29	85	MY	241	0.13
36	TRAVEL	518	0.29	86	WANT	237	0.13
37	TOURISM	514	0.28	87	STUDENT	235	0.13
38	BY	492	0.27	88	RIGHT	235	0.13
39	ROOM	486	0.27	89	SEE	234	0.13
40	THEY	475	0.26	90	HERE	234	0.13
41	LIKE	465	0.26	91	FIND	232	0.13
42	ONE	451	0.25	92	DIFFERENT	228	0.13
43	THERE	450	0.25	93	PRACTICE	226	0.12
44	WOULD	427	0.24	94	JUST	221	0.12
45	GO	419	0.23	95	SPEAKING	218	0.12
46	PEOPLE	415	0.23	96	PARIS	217	0.12
47	QUESTIONS	411	0.23	97	ME	215	0.12
48	LOOK	400	0.22	98	NOW	213	0.12
49	USE	395	0.22	99	THAN	212	0.12
50	BUT	387	0.21	100	UP	209	0.12

Table 1 shows that the most common terms are mostly grammatical elements, such as 'THE' (F=9,819, 5.42%), 'A' (F=9,196, 5.07%), 'BE' (F=4,857, 2.68%), and 'TO' (F=4,722, 2.34%). These high-frequency function terms are common in the English language corpus. However, some topic terms explicitly relevant to the tourism area appear prominently in the top 100, showing their pedagogical importance in these coursebooks. For example, the terms 'GUEST' (F=882, rank 22), 'HOTEL' (F=872, rank 23), 'ROOM' (F=635, position 28), 'TRAVEL' (F=518, rank 36), and 'TOURISM' (F=514, rank 37) appear often, highlighting the instructional materials' emphasis on fundamental tourism concepts and functions.

4.2. Alignment with Established Corpora (Keyword Analysis)

The Tourism Corpus (TC) was compared to a general English reference corpus (such as the British National Corpus or BNC) and a specialized tourism corpus to perform a keyword analysis and answer the research question regarding the alignment of the coursebooks' lexical profile with established corpora. The purpose of this research was to determine which terms are statistically overrepresented in the coursebooks, thus emphasizing the specialist vocabulary, and to evaluate the degree to which this vocabulary corresponds to real language use in both general and particular tourist situations. The study of these high-frequency terms provides firsthand knowledge about the coursebooks' thematic focus. For example, the prevalence of the words 'GUEST' and 'HOTEL' indicates a high emphasis on customer service and lodging in the tourist industry. A common usage may be found in the concordance line below.

"GUEST"

one of the rooms we have ordinary blankets for	**guests**	who prefers them. So that's no problem at all.
as a personal touch, we like to welcome our	**guests**	with a vase of flowers in the room on arrival.
of the reception desk, the ashtray was not emptied.	**guests**	also kept waiting at reception for unacceptable
What would we do if terrorists attacked some of our	**guests**	hostage? 6. What if the computer system?
WAITRESS Good evening, sir, good evening, madam.	**Guest**	Good evening. WAITRESS: Do you have a

"HOTEL"

You need to send a map of the location of his	**hotel**	to Mr. Stuart Hosfield, a client.
Rainy beaches, cathedrals, carnival, restaurant,	**hotel**	train sunny desert museum folk dance bar motel
is free in business, not even a client booking a	**hotel**	room. Reservations made directly by the client
month, during which discounted air tickets and	**hotel**	charges will be offered to tourists. Rio de Ja
You can have a metropolitan	**hotel**	but you cannot have a metropolitan cottage.

Furthermore, the frequent use of the words "TRAVEL" and "TOURISM" clearly indicates the subject matter, as demonstrated by the concordance below.

"TRAVEL"

motivation and behavior: Listening Reasons for	**travel**	and money spent on travel, the pie charts show
on a business trip, you can claim things like	**travel**	and hotel costs but you must remember to
industry and it will give me the opportunity to	**travel**	to other countries. The sector of tourism I was in
Write a news item for a web page or a	**travel**	magazine describing your campaign. Use the
journey you would like to go on. Visit a local	**travel**	agency and ask for information. During your visit

"TOURISM"

The growth of adventure	**tourism**	market growth in the adventure holiday market
Gabriela Cometa, I am interested in	**tourism**	because it is an exciting industry, and it will
I have had some training and experience in	**tourism**	already. This includes an initial tourism training
in tourism already. This includes an initial	**tourism**	training certificate, and experience with childcare
to do in the future? What do you like about	**tourism**	What do you do to relax? You get to experience

The examples shown here demonstrate how the core principles presented in the coursebooks may be directly applied to the common situations and interactions that tourists encounter. In addition, several methods were very purposefully applied to ensure the legitimacy and trustworthiness of the conclusions of the lexical analysis. To guarantee that the fundamental data was clean and accurate, the rigorous procedure of compiling the corpus, which is described in full in the methodology, consisting of systematic processes for data conversion and error correction, was carried out. We utilized WordSmith to manage all the words, which methodically generates frequency lists and ensures consistent counting. A customized dictionary named "English for Tourism" also helped people understand domain-specific terminology. We created this lexicon using the official Oxford Dictionary of Travel and Tourism, which helped minimize misunderstandings that came from using general English meanings. These activities were all done to make the lexical insights from the study more reliable and rigorous in terms of analysis.

5. Discussion

This study aimed to analyze the lexical profile of English for Tourism coursebooks, revealing details about their alignment with established corpora and real-world language use in the

tourism industry. The lexical analysis of the Tourism Corpus (TC) provided answers to the study's research objective, revealing both strengths and areas for development in the coursebooks' vocabulary.

The results of the study showed that out of the top 100 keywords most often used in coursebooks, function phrases were the most prevalent. In addition, this finding aligns with instructional strategies in English for Specific Purposes (ESP) that prioritize high-frequency general vocabulary and fundamental grammar, especially for beginning learners (Hyland, 2006; Nation, 2013). This strategy makes knowledge more accessible and instantly relevant for students by promoting basic conversational skills and language use in practical settings across contexts. Crystal (2003) suggests that such basic vocabulary may boost students' confidence and linguistic readiness, especially in initial tourism-related English-language exchanges, which is in support of the greater requirement for extensive communication skills in current globalization. According to Paltridge and Starfield (2013), the observed overlap between tourism-specific and general English vocabulary further promotes language learning. Such overlap is a trait that is often seen in early English for Speakers of Other Languages (ESP) curricula.

According to the study's findings, the primary themes in these textbooks consist of content phrases commonly used along with function words. Hospitality-related issues, such as accommodation and customer service, appear to be highly valued, as demonstrated by the frequent use of phrases like "guest" (882 times), "hotel" (872 times), and "room" (635 times). This approach is consistent with other studies that emphasize the importance of these topics in establishing a rapport with first-time visitors (Bhatia, 1993; Dann, 1996). The often-used industrial phrases "travel," "tourism," "employee," "flight," and "tourist" all emphasize a practical approach to common transactional and interactional situations. The study also emphasized the importance of certain words, such as "ask," "please," "listen," "practice," and "speaking," in the regular service and communication tasks performed by tourist employees. This emphasis on speaking and customer service skills is a pedagogical objective often highlighted in tourist English education, providing students with a solid language foundation that is crucial for future industry professionals (Ellis & Johnson, 1994; Louhiala-Salminen, 1999). Additionally, the inclusion of verbs that are not exclusive to tourism-related settings, such as "good," "job," "make," and "want," suggests a goal to provide students with a wider range of communication skills. While the coursebooks did a good job with the basics and communication skills, the research showed that they were severely lacking in vocabulary, especially when it related to modern business trends. This includes terms like "inclusive travel," "sustainable tourism," and "online booking," which are relevant to the more modern or expanding areas of the business.

While the coursebooks did a good job with the basics and communication skills, the research showed that they were severely lacking in vocabulary, especially when it related to modern business trends. These terms, such as "inclusive travel," "sustainable tourism," and "online booking," relate to the more modern or growing parts of the business.

The absence of these regions is concerning, as they represent significant and evolving aspects of the modern tourist industry (UNWTO, 2021; Lane & Kastenholz, 2015). In contrast to the increasing necessity for culturally sensitive communication in international travel, the

coursebooks only contained a limited number of location-specific or culturally rich nouns, such as "Paris" or "country" (Hofstede, 2001; Spencer-Oatey, 2008).

This lexical gap may prevent students from learning the more varied and up-to-date words that describe how the tourist industry is changing. Richards and Rodgers (2001) stress how important it is to slowly introduce specialized and advanced vocabulary. This approach is based on the idea that using many broad and common words can be a deliberate educational strategy, making it easier for beginners to understand and quickly utilize the material. This gradual introduction would help students learn the language of the tourist industry today, which is what people want in an ESP curriculum (Dudley-Evans & St. John, 1998).

This work contributes significantly to the field by demonstrating how a corpus-based analysis can objectively assess the strengths and shortcomings of coursebook material in English for Tourism. In contrast to anecdotal findings from previous techniques, our study provides data-driven insights into the lexical profile of these textbooks. As a result of the fact that this research provides educators, curriculum designers, and textbook writers with information that is verifiable and supported by statistics on the language that is utilized in existing resources, the findings have a substantial impact on a variety of different individuals.In addition, the materials that are used in the classes ought to be more applicable and tailored to the requirements of students who are interested in pursuing careers in the tourism industry. This is because the language skills needed for these jobs are getting harder. Our results also show that the current textbook content is effective at teaching basic communication skills. However, they also show that changes are needed, such as adding more modern and specialized industrial terms. The pupils' language skills will determine how well they can handle the complicated and always-changing global tourism industry. This backs up Flowerdew and Peacock's (2001) general argument for making updates to VLE materials all the time.

6. Conclusion

This study examined the lexical content of English for Tourism coursebooks used in Thai higher education using a corpus-based technique. We found that high-frequency function words and common content phrases associated with basic service interactions and hospitality were highly valued in our study of a corpus of 3.4 million words. This finding implies that the coursebooks effectively teach the English language skills required for casual conversations and common tourist situations. However, the survey also found a significant lack of specialized, contemporary, and culturally diverse language relevant to the evolving travel and tourism sector.

The findings of this study reveal that a number of major empirical insights into the current situation with English for Tourism teaching resources have been discovered. These insights have been demonstrated by the findings of this investigation. When seen from an objective standpoint, it demonstrates both educational capabilities in the construction of fundamental vocabulary as well as serious lexical gaps relating to the most recent advancements in the industrial sector. This verification is performed by the utilization of a method that is founded on corpora, and the validation that is based on data is utilized by curriculum designers, textbook writers, and educators to build educational materials that are more pertinent to the requirements of students and that are responsive to their demands. The study's results

confirm the importance of present material while also outlining a clear path for improvement. This effort ensures that future tourism professionals are equipped with language skills that truly reflect the industry's evolving global landscape while also making a significant contribution to the field of English for Specific Purposes (ESP) by providing a detailed lexical profile of a widely used educational resource.

7. Recommendations for Future Research

Building on these findings, future research could explore several avenues. First, a longitudinal study tracking the lexical evolution of tourism coursebooks over time would reveal how quickly these materials adapt to industry shifts, particularly concerning emerging concepts like sustainable tourism and digitalization. Second, a comparative analysis of coursebooks from different cultural or regional contexts could teach us about varied pedagogical approaches and their impact on lexical selection. Third, integrating qualitative data, such as interviews with tourism professionals and students, might offer deeper insights into the perceived utility and real-world application of the vocabulary found in these coursebooks. Finally, exploring the receptive and productive vocabulary knowledge of students using these coursebooks could further evaluate the effectiveness of the current lexical emphasis and inform future instructional design. These future directions will help refine our understanding of effective English language education for tourism.

Acknowledgements

References

Al-Busaidi, S. S. (2022). *Language Proficiency and Academic Performance*. In: Mohebbi, H., Coombe, C. (eds) Research Questions in Language Education and Applied Linguistics. Springer Texts in Education. Springer, Cham. https://doi.org/10.1007/978-3-030-79143-8_39

Afzalia,K. , Rezapoorianb,D.(2014). *Pragmatic Aspects of English for Tourism Course Books and ESL Learners Pragmatic Needs: A Speech Act Theory Perspective. Social and Behavioral Sciences 98 (2014) 52 − 59.* https://doi.org/10.1016/j.sbspro.2014.03.388

Basturkmen, H. (2010). *Developing courses in English for specific purposes*. Palgrave Macmillan.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Blanka, K. (2015). Teaching and Assessing Meeting Skills Procedia - Social and Behavioral Sciences Volume 186, 13 May 2015, Pages 894-897. DOI:10.1016/j.sbspro.2015.04.111

Blue, G. M., & Harun, M. (2000). Hospitality language as a professional skill. *English for Specific Purposes, 22 (1)*, 73–91. doi:10.1016/S0889-4906(01)00031-X

Brooks, M.A. (2014). *The Role of Vocabulary in English for Specific Purposes(ESP) Teaching and Learning*: Considerations for Asia University.

Bühler, K. (2011). *Theory of language: The representational function of language*. John Benjamins Publishing Company.

Chang, W. (2024). ESP Instruction Design Based on Digital Literacy for Tourism English Learners. *Journal of English Teaching through Movies and Media, 25*(3), 2024. doi: 10.16875/stem.2024.25.3.60

Coxhead, A. (2021). Vocabulary in English in tertiary contexts: connecting research and learning. *LEARN Journal: Language Education and Acquisition Research Network, 14*(1), 1-14.

Crosthwaite, P., & Cheung, L. (2019). *Learning the language of dentistry: Disciplinary corpora in the teaching of English for specific academic purposes*. John Benjamins Publishing Company.

Curry, N., & McEnery, T. (2025). Corpus linguistics for language teaching and learning: A research agenda. *Language Teaching: Surveys and Studies. Language Teaching, 58*(2), 232-251 doi: https://doi.org/10.1017/S0261444824000430.

Crystal, D. (1997). *English as a global language*. Cambridge University Press.

Dann, G. M. S. (1996). *The language of tourism: A sociolinguistic perspective*. CAB International.

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.

Flowerdew, L. (2013). *Corpus-based studies of language use, language learning, and language teaching*. John Benjamins Publishing Company.

Flowerdew, L. (2013). Needs analysis and curriculum development in ESP. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 325–346). Wiley-Blackwell.

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching, 40*(2), 97–118. doi:10.1017/S0261444807004144

Harwood, N. (2005). What do we want EAP teaching materials for? *Journal of English for Academic Purposes, 4*(2), 149–161. doi: 10.1016/j.jeap.2004.07.008

Hsu, Wenhua. (2009). College English Textbooks for General Purposes: A Corpus-based Analysis of Lexical Coverage. *Electronic Journal of Foreign Language Teaching, 6*(1)

Huynh, Thu-Nguyet & Nguyen, Long. (2020). Applying corpus linguistics to English textbook evaluation: A case in Viet Nam. *VNU Journal of Foreign Studies, 37*(4). doi: 10.25073/2525-2445/vnufs.4676

Kacetl, J., Klímová, B. (2015). English Vocabulary for Tourism – A Corpus-Based Approach. In: Christiansen, H., Stojanovic, I., Papadopoulos, G. (eds) *Modeling and Using Context. CONTEXT 2015. Lecture Notes in Computer Science*, vol 9405. Springer, Cham. https://doi.org/10.1007/978-3-319-25591-0_37

Kennedy, G. (1998). *An introduction to corpus linguistics*. Longman.

Kim, D., & Gilman, D. A. (2008). Effects of Text, Audio, and Graphic Aids in Multimedia Instruction for Vocabulary Learning. *Educational Technology & Society, 11* (3), 114-126.

Laosrirattanachai, P. & Ruangjaroon, S. (2020). The word lists of hospitality service review construction. *Journal of Studies in the English Language, 15*(1), 107-158.

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. DOI: 10.1016/j.jeap.2016.01.008

López-Rocha, S. (2020). Refocusing the development of critical intercultural competence in higher education: challenges and opportunities. *Language and Intercultural Communication, 21*(1), 118–131. https://doi.org/10.1080/14708477.2020.1833900

Mackay, R., & Mountford, A. (Eds.). (1978). *English for specific purposes: A case study approach*. Longman.

McEnery, A. M., & Wilson, A. (1996). *A corpus linguistics*. Edinburgh University Press.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Namtapi, T. (2022). Needs Analysis of English for Specific Purposes for Tourism Personnel in Ayutthaya. *LEARN Journal: Language Education and Acquisition Research Network, 15*(1),409-439. https://so04.tci-thaijo.org/index.php/LEARN/index

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*(1), 59–82. doi: https://dx.doi.org/10.1353/cml.2006.0049.

Nation, P., & Newton, J. (1997). Teaching vocabulary. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 238–254). Cambridge University Press.

Trinant, K. & Kijpoonphol, B. (2021). Lexical collocations in a sample corpus of tourism research articles (SCTRA). *NKRAFA Journal of Humanities and Social Sciences, 9*, 94–108.

Wagner, A. (2023). An Analysis of Curriculum and Materials Design in ESP for Tourism Courses in Japan and Thailand. *Journal of Humanities Research*, 14, 2021

Wang, W. (2024). Recent Research on L2 Vocabulary Collocations Acquisition. What are the Implications for Language Teaching?. *International Journal of Education and Humanities, 16*(2), 305-308. https://doi.org/10.54097/16mmbj24

Wang, H., & Wang, Y. (2017). Jiyu yuliaoku de daxue yingyu zhuanye jiaocai pinggu fangfa de queli (The establishment of corpus-based evaluation approach of college ESP textbooks). *Contemporary Education Research and Teaching Practice, 12*, 205.