



Academic vocabulary in sport tourism news: A corpus-based study

Darunee Yotimart ^{a1} 

^a *Thailand National Sports University, Chaiyaphum campus, Chaiyaphum, Thailand*

APA Citation:

Yotimart, D. (2021). Academic vocabulary in sport tourism news: A corpus-based study. *Journal of Language and Linguistic Studies*, 17(3), 1527-1535. Doi: 10.52462/jlls.110

Submission Date:01/06/2021

Acceptance Date:05/09/2021

Abstract

The objectives of the research is to study and analyze Academic Word List (AWL), which appeared in the corpus of sport tourism news wordlists collected from international online news. The framework of this study was based on Coxhead's (2000) 570- Academic Word List (AWL). A corpus of 112,623 running words taken from the international online news, against the AWL of 570 words by using the Concordance Program, AntWordProfiler, for comparing the vocabulary load of the corpus of sport tourism news and AWL. The data were analyzed and displayed as percentages and frequencies. The results revealed that words in the AWL covered 7.38 % of the total words in the corpus. Findings also revealed that some of the words found in the present study were not found in Coxhead's AWL. This suggests that vocabulary needs of students in sport or tourism are characteristically different from those of students in other disciplines. The findings indicate the usefulness of AWL for sport learners and suggest the need to produce field-specific word lists embracing all frequent academic vocabulary items necessary for the expression of the rhetoric of the specific discipline. The Sport Tourism News wordlists can guide instructors to prepare the sport or tourism students in setting their vocabulary learning goals.

Keywords: Specialized Academic Word Lists; Academic Word List (AWL); Sport Tourism News; Corpus Linguistics; English for Specific Purposes (ESP)

1. Introduction

Thailand's higher education aims at training students not only in new developments of graduates' knowledge, equipped with 21st century skill attainment pertaining digital age literacy, effective communication, innovative thinking and high efficiency, but also in their specific field knowledge. In the test summary launched by the Educational Testing Service for TOEFL Internet and paper-based test, Thailand's steadily left behind other Southeast Asia (Ulla, 2018). To equip them with the crucial tools for their highest performance, the need to master their English generally and specifically in their field has to be addressed. In keeping with the spirit, the importance of studying English for specific purposes in each career path could not be compromised so that learners are shaped to be well-prepared in utilizing English for their specific purposes or fields practically in more internationalizing society (It-ngam, T and Phoocharoensil, 2019). Vocabulary is considered the language element that is key to

¹Corresponding author

E-mail address: daruneeeyoti@gmail.com

successful language learning. Knowledge of specialized academic vocabulary is of importance for the academic success of EFL students in all fields. English for Specific Purposes (ESP) learners often come across the problem in learning academic words, in comparison with other word types in specialized texts (Bunyarat, 2020). Maamuujav (2021) pointed out that retaining advanced vocabulary is an important factor that leverages the academic writing quality. The analysis of multiple hierarchical regression show that the writing ability was shaped by the academic vocabulary usage. Moreover, Maamuujav, Olson, and Chung (2021) also revealed that sentence boundary, the low use of syntactic variety, the underuse of complexed subordinate clauses to link ideas, and the lack of complexity or sophisticated uses in vocabulary are commonly found in L2 students' writing. The regression analyses posited that the use of syntactically advanced lexical level prejudged the perceived high writing quality.

However, English for Specific Purposes (ESP) learners often come across the problem in learning academic words, based on the comparison with other word types in specialized texts such as textbooks, research articles, academic speech (Gholaminejad and Sarab, 2020; Heidari, Jalilifar and Salimi, 2020; Aldohon, 2018). Therefore, pointing out the list of academic words to ESP learners helps master their language in specific areas with minimal time spent in learning language or understanding specialized texts (Tongpoon-Patanasorn, 2018). Recent findings do not adequately generalize one universal academic wordlist which can be able to response the students' needs in a wide range of fields (Gholaminejad and Sarab, 2021). Tourism discipline possesses a high number of non-English speaking students who are involved in academic communication, deprived of an access to a ready-made resource (Salim, Ibrahim, & Hassan, 2012). Thus, this issue has motivated researchers to provide specific wordlists of this particular discipline. The additional reason to choose sport tourism in particular is that the present researcher was inspired by the latest curriculum for sport and sport communication students in Thailand National Sports University (TNSU) in that they are increasingly required to develop the academic skill to prepare for compulsory English academic test in their final year as an English exit exam prior to graduation. Moreover, the result of the present study will be adopted for the course content under English News Translation subject for sports communication students in TNSU, whereby they need to read and translate a lot of sport-related news, especially the sport tourism news which has gained popularity in sport field in Thailand (Meenoi & Sungruga, 2018).

Objectives of the study

The following research question is that what are the most frequently used academic words in a corpus of sport tourism? This study attempted to examine the generality of Coxhead's (2000) AWL across the sport tourism news corpus. The corpus provided students with a shortcut to expose to the words that they will see most frequent in their fields or subject areas. Instructors can also use these word lists to create the materials to encourage sport tourism related students to get more familiar with vocabulary they have to often see.

2. Literature Review

Corpus linguistics

Corpora allow researchers to analyze language in a number of different ways. Corpora can be built up by collecting related language materials and recording the target language, words, grammar, and collocations within. It can be used to show word frequency, phraseology, and collocation. The high-frequency words indicate the communal of word tokens in writing of different fields and contexts. As a result, these words are the handiest for language learners. However, technical words can't be used in

any discipline due to their limited occurrence in some specific fields. The low-frequency words are the words with rare usage. Academic words highly occur in academic texts from various text types and fields (Nation, 2017). Additionally, corpora can generate frequency lists which can be beneficial to witness the probable differences among the corpora. The categories of corpora are distinguished based on their particular purposes as stated below. A specialized corpus is a group of texts from a particular genre, namely, newspaper editorials, specialized textbooks, and academic articles in a particular field, lectures, informal talks, or students' essay, which is to be representative of these particular genres (Bunyarat, 2020; Therova, 2020). Given the importance of academic vocabulary as aforementioned, the academic wordlist can be created through the use of corpus linguistics. The use of corpus collection of texts is increasingly impactful technique in researching language (Hu, Gao, and Qiu, 2021; Khany and Kalantari, 2021; Zahra and Abbas, 2020).

Corpus-based vocabulary research for discipline-specific purposes

Corpus-based language teaching is gradually becoming mainstream, especially in ESP. Many corpora have been developed for different fields of specific disciplines based on its word frequency analysis and coverage against the academic word list (AWL) (Heidari, Jalilifar and Salimi, 2020; Khair, Firdaws and Ramdhani, 2019; Quero, 2017; Wang, 2017). For example, the coverage of AWL word forms in the Pharmacy Research Articles Corpus was 17.69 % (Heidari, Jalilifar and Salimi, 2020), in the Science Research Articles Corpus was 5.82 % (It-ngam, T and Phoocharoensil, 2019), in the Linguistic and Literature Abstract Corpus was 9.2% and 9.1 %, respectively (Khair, Firdaws and Ramdhani, 2019), in the Art Research Articles Corpus was 10.57 % (Wang, 2017), in the Medical Text Corpus was 8.23 % (Quero, 2017). Recently, Khany and Kalantari (2021) generated academic word list specifically for accounting to see the congruence between the word list and Coxhead's academic word list (AWL). The analysis of research article in accounting field identified 658 academic word families with the highest frequency in the corpus called Accounting Academic Word List (AAWL), accounting 10.16 % of the whole corpus with 354 words concurrent to the AWL. In addition, 3.98 % of the whole corpus came to 50 most frequently used words in the list. This study substantiates the importance of corpus-based word lists specially for a particular discipline. The findings of also propose the implication of AAWL in the accounting community for its reference.

However, limited studies in academic lexical of sport tourism are found using a corpus-based analysis. To name a few, Kongcharoen (2018) built a corpus of 280 research articles in physical education and sport science, approximately 1.1 million words, and then compiled a list of high-frequency words for the corpus and compared that list to that of Coxhead's (2000) Academic Word List (AWL) collected from research articles in a wide range of fields. The result showed that the coverage of AWL word forms in the Physical Education and Sport Science Research Articles Corpus was 21.15%. However, the previous studies did not examine academic words in the sport tourism field under the international news content. The present paper underlines existing gaps in the literature by proposing frequent academic words based on the corpus of academic word list in sport tourism. The reason of choosing sport tourism is that it helps raise an awareness of students in sport and tourism field in the newest trend of sport tourism around the globe due to its development can economically support country or at least community which has valuable touristic and sport resources or premises. To keeping up with the sport tourism news update around the globe, those students or sport tourism personnel should be able to read and understand international English news.

As we can see, few research are involved in the investigation of sport and tourism discourse community together. Therefore, research that center around the language in sport tourism related texts should have been conducted in order to familiarize students and people involved in sport field, especially TNSU students, to more words frequently appeared in the specific-area context. The result

of this study which is academic word list in sport tourism field can directly boost the vocabulary repertoire of English language learners and indirectly inform and raise awareness of the present trend of sport tourism via online news.

Framework

Three kinds of words that learners and teachers of English language should focus based on the derivation of their frequencies have been listed in vocabulary teaching and learning research. They are the General Service List (GSL) gathered by West (1953), mentioned as the 2,000 most frequent words in English; the 570-word Academic Word List (AWL) established by Coxhead (2000), consisting of words recurrently found across a number of academic texts than in non-academic texts; and the specialized or technical vocabulary. Academic vocabulary are words basically found in a wide variety of academic texts but rather unusual in other text types (Dang, Coxhead, and Webb, 2017; Gholaminejad and Sarab, 2021).

The Academic Word List (AWL) was developed by Coxhead (2000), and it contains 570 word families that account for 10.0% of the total words (tokens) in academic texts. The AWL was developed from different academic journals and university textbooks in 28 disciplines from four main areas: the arts, commerce, law, and natural science. It was found to be useful for further research pertaining to academic vocabulary (Therova, 2020). The reason to use Coxhead's (2000) AWL as a reference to academic vocabulary in the current study is that the AWL has been mainly adopted and utilized by ESP research to analyze the target words in written texts (Quero, 2017). Dang (2017) states that AWL is compiled by using a large amount of more balanced data and consistent criteria of frequency for being representative than other previous wordlists. Recently, Gholaminejad and Sarab (2021) also insisted that the AVL (the academic vocabulary list) should not replace the counterpart, AWL (the academic word list). Therefore, there have been no counterparts competing AWL.

3. Research Methodology

Samples

The Sport Tourism News Corpus was specifically created for this study. It consists of 150 sport tourism news from the prestigious news sources in Thailand, e.g. The Nation, and the world, e. g. CNN and BBC, collected from their websites with their date, picture, and sources deleted.

Research Instrument

Regarding the data processing, the research articles were in a PDF format and had to be converted into a *.txt* for further processing. After obtaining the corpus, the software called AntWordProfiler available from Laurence Anthony's Website was used for analyzing the vocabulary load of texts, how much and what vocabulary occurs in the particular group of AWL texts. The program enables the search and count of text components. It also analyzes the use of word forms and its contexts, phrases and morphemes, as well as calculating the word forms or compare them to other referenced texts (Kunanets, Levchenko and Hadzalo, 2018, September).

Corpus linguistics centers around the analysis of natural language through the use of computer technology to help analyze large volume and text corporuses. AntWordProfiler launched by Dr. Laurence Anthony is designed to analyze corpus linguistics and is considered to be powerful tool for linguistic research conduct (Kunanets, Levchenko, and Hadzalo, 2018, September). Thus using the program, the number of occurrences of each word as well as the words shared with the AWL word lists were determined.

Data collection

It should be noted that the approximate 300 recent sport tourism news were purposively selected and all of which were published in the year 2014-2020. The data were collected in October and November 2020. Due to the Covid 19 pandemic in 2020, I decided to add more news published in 2014. The final Sport Tourism News Corpus has approximately 112,623 words with 9,352 word types.

Data Analysis

After using the AntWordProfiler program, the number of occurrences of each word as well as the words shared with the AWL word lists were determined. The reason for adopting the frequency rates was supported by the Coxhead's (2000) selection of AWL words. Coxhead's corpus of the AWL was approximately 3.5 million words and the selection criteria of the AWL word forms was that each word form in the AWL should be found minimum 100 times in the whole corpus.

To ensure a rigorous word selection, the researcher followed Hyland and Tse (2007)'s criterion, which considered those frequent words as those occurred above the mean of the total number of academic words. The mean for full text was identified and families of those items above the mean were generated and counted. This criterion was justifiably replicated by the previous research (Ibrahim, Shah & Abudukeremu, 2018).

4. Results and Discussion

The results of Academic Word List (AWL)'s coverage in the corpus provide the total word token and the coverage of AWL in the corpus of sport tourism news from the international news database as in the following table.

Table 1. The AWL's Coverage in the corpus

| LEVEL | FILE | TOKEN | TOKEN% | CUMTOKEN% | TYPE | TYPE% | CUMTYPE% | GROUP | GROUP% | CUMGROUP% |
|--------------|-------------|--------|--------|-----------|------|-------|----------|-------|--------|-----------|
| 1 | awl_570.txt | 8312 | 7.38 | 7.38 | 1051 | 11.24 | 11.24 | 449 | 5.13 | 5.13 |
| 0 | - | 104311 | 92.62 | 100 | 8301 | 88.76 | 100 | 8301 | 94.87 | 100 |
| TOTAL | | 112623 | | | 9352 | | | 8750 | | |

Following the table 1, a corpus of 112,623 running words taken from the international online newspaper, against the AWL of 570 words by using the Concordance Program. The results revealed that words in the AWL were 8,312, covering 7.38 % of the total word tokens in the corpus. In short, the AWL coverage in the corpus was less than that of Coxhead's (2000) coverage of Academic Corpus (10%).

Based on the table 1, the word families or group in this corpus were generated, totaling 449 academic words. Following (Hyland and Tse, 2007) frequency criterion, hence, the researcher first identified the mean for these 449 academic word occurrences which was 17. This was the criterion that was used in developing the frequent word families or group of the corpus. Words occurring fewer than 17 times were omitted from the list, thus 114 academic words out of 570 academic words were derived

to be further examined. The examples of the frequent words are *globe*, the most frequently used academic word, occurred 306 times, accounting for 3.6 %, followed by *partner* (253 times or 3.0 %) and *team* (236 times or 2.8 %). There were 8,312 academic word tokens or 449 academic word families or groups found in the corpus, but 121 academic word families did not occur in the corpus. The disappearance of these words can be explained that some words may be frequently found in a particular field, for example, in scientific texts, or sport news. Kamasak and Soyaltin (2021) pointed that some academic words which potentially occur in academic texts across different genres and academic text type did not seemingly show any occurrence in other kinds of texts. Therefore, it can be concluded that these words found in special contexts should be attentive by teachers (Jamalzadeh and Chalak, 2019).

In a nutshell, the academic words from the AWL have 7.38 coverage of the corpus of sport tourism news from the international news database. The findings showed that the percentage of academic words was lower than the previous related studies of AWL in other genres, namely, 10.16 percent in Khany and Kalantari (2021), 9.2 percent in Khair, Firdaws and Ramdhani (2019), 8.23 percent in Quero, (2017), 10.57 percent in Wang (2017), and 5.82 percent in It-ngam and Phoocharoensil (2019). However, this is not unexpected since academic words, as the name suggests, naturally occur in academic texts. However, the present study can be considered as useful instructional resources for this discipline. For example, there were 12 words which occurred more than 100 times. They were *globe*, *partner*, *team*, *economy*, *major*, *facilitate*, *culture*, *participate*, *region*, *promote*, *create*, and *community*. Furthermore, all 114 considered frequent academic words can be selectively incorporated in the vocabulary recognition Bingo games or Charades activities due to its utmost useful for sport or tourism discipline students.

From the findings, AWL word forms do not cover all the reading materials in the field of sport tourism. In line with Maamuujav, Olson and Chung (2021), the difference between academic vocabularies and high frequency words is that the former is less frequent in informal conversations and speeches. Given their high frequency in written texts, learning these words repeatedly is vital for language learners or beginner learners of English for specific purposes (ESP) (Nation, 2017).

5. Conclusion

The findings of the current research underline the necessity to build field-specific academic word lists to response the needs of non-native English learners of various discourse backgrounds. The word lists can be served as a pedagogic map to prepare English teachers for the English language content or syllabus design to achieve the students' useful vocabulary items. Thus, subject specificity of academic word lists is of importance and usefulness.

Contribution

Pedagogical implications are made for researchers, students, and material designers in sport and tourism field, who can use The Sport Tourism News word lists as a lexical repertoire to set their vocabulary learning/teaching goals. For example, English teachers can have an activity for boosting the recognition of the discipline specific academic word lists by selecting high frequent words in the given corpus and present them through fun word recognition games and activities, such as Heads up/Charades and Bingo games, etc. This will help them pick up these academic vocabularies without pressure from academic atmosphere. Moreover, English teachers can cooperate with other field subject instructors to incorporate these vocabulary word lists into the subject content to help the sport students at least come across and gradually recognize them. The findings of this study also assist ESP and EFL teachers with academic material and syllabus preparation, especially in the English News Translation

subject of TNSU's curriculum or in other sport or tourism related fields, as well as prepare them for English academic English types of test, namely TNSU English Test.

Suggestion

Though students need to focus on AWL, field specific word lists are encouraged to further focus. Field specific lists can help students learn necessary words which are important for their field of study. Therefore, the field specific words in this field of study should be further conducted, followed by grouping these words into semantic groups in order to equip students with which words are needed to study first in their field.

Limitation

Finally, this research is limited by the amount of online-available sport tourism news because most of the news sources do not provide the published sport tourism news section. As a result, a larger scale of word tokens from news reports and more various specific fields are necessary for further studies due to the appropriate-sized vocabulary dependent on the technical vocabulary knowledge of a given field.

References

- Aldohon, H. I. (2018). A Corpus-Based Investigation of Academic Vocabulary and Phrasal Verbs in Academic Spoken English (Doctoral dissertation, Carleton University). Ontario, Canada Bunyarat, P. (2020). A corpus-based study of specialized vocabularies in industrial work for developing industrial word lists of English for science and technology class. *Journal of Language and Culture*, 39(2), 122-151.
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213-238.
- Dang, T.N.Y. (2017). Investigating vocabulary in academic spoken English: Corpora, teachers, and learners (Unpublished PhD Thesis). Victory University of Wellington, New Zealand.
- Dang, T.N.Y, Coxhead, A. and Webb, S. (2017). The Academic Spoken Word List. *Language Learning*, 67(4), 959–997.
- Gholaminejad, R. and Sarab, M. R. A. (2021). A comparison of the academic word list and the academic vocabulary list: should the AVL replace the AWL?. *TEFLIN Journal*, 32(1), 167-182
- Gholaminejad, R. and Sarab, M.R. (2020). Academic vocabulary and collocations used in language teaching and applied linguistics textbooks: A corpus-based approach. *Terminology*, 26(1), 82-107.
- Heidari, F., Jalilifar, A. and Salimi, A. (2020). Developing a Corpus-Based Word List in Pharmacy Research Articles: A Focus on Academic Culture. *International Journal of Society, Culture & Language*, 8(1), 1-15.
- Hu, J., Gao, X. and Qiu, X. (2021). Lexical Coverage and Readability of Science Textbooks for English-Medium Instruction Secondary Schools in Hong Kong. *SAGE Open*, 11(1), 15-30.
- Ibrahim, E. H. E., Shah, M. I. A., & Abudukeremu, M. (2018). A Corpus-Based Lexical Study of the Frequency, Coverage and Distribution of Academic Vocabulary in Islamic Academic Research Articles. *The Journal of Social Sciences Research, special issue 2*, 570-577.
- It-ngam, T. and Phoocharoensil, S. (2019). The development of science academic word list. *Indonesian Journal of Applied Linguistics*, 8(3), 657-667.

- Jamalzadeh, M. and Chalak, A. (2019). A Corpus-based Study of Academic Vocabulary in Physiotherapy Research Articles. *Language Teaching and Research Quarterly*, 9, 69–82.
- Kamasak, R., and Soyaltin, T. E. (2021). *Promoting the Effectiveness of Language Teaching Through Online Academic Word Lists and Extensive Reading in the COVID-19 Pandemic*. In *Digital Pedagogies and the Transformation of Language Education* (pp. 165-182). IGI Global.
- Hyland, K., & Tse, P. (2007). Is There an "Academic Vocabulary"? *TESOL Quarterly*, 41(2), 235-253. doi:10.2307/40264352
- Khair, I., Firdaws, A.F. and Ramdhani, R. A. (2019, December). *Vocabulary Profile of the Abstracts Written by Students Majoring in Linguistics, Literature, and Cultural Studies: A Comparative Analysis*. In *Social Sciences, Humanities and Education Conference (SoSHEC 2019)*. Atlantis Press. Indonesia
- Khany, R. and Kalantari, B. (2021). Accounting Academic Word List (AAWL): A Corpus-Based Study. *Journal of Foreign Language Teaching and Translation Studies*, 6(1), 35-58.
- Kongcharoen, P. (2018). Basic Physical Education and Sport Science English Word List for Physical Education Students. *rEFLECTIONS*, 25(2), 120-147.
- Kunanets, N., Levchenko, O. and Hadzalo, A. (2018, September). *The application of AntConc concordancer in linguistic researches*. In *2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, 2, 144-147). IEEE.
- Maamuujav, U. (2021). Examining lexical features and academic vocabulary use in adolescent L2 students' text-based analytical essays. *Assessing Writing*, 49, Retrieved from <https://doi.org/10.1016/j.asw.2021.100540>
- Kisa, N. (2019). The skills of principals to manage the generational differences: A scale development study. *Educational Administration: Theory and Practice*, 25(4), 745-772.
- Maamuujav, U., Olson, C. B. and Chung, H. (2021). Syntactic and lexical features of adolescent L2 students' academic writing. *Journal of Second Language Writing*, 53, Retrieved from <https://doi.org/10.1016/j.jslw.2021.100822>
- Meenoi, P., & Sungrugsa, N. (2018). *Model and Strategies for Sports Tourism Management of The Western Region to Prepare for Thailand 4.0* (Doctoral dissertation, Silpakorn University).
- Nation, P. (2017). How vocabulary is learned. *Indonesian Journal of English Language Teaching*, 12(1), 1-14.
- Quero, B. (2017). A corpus comparison approach for estimating the vocabulary load of medical textbooks using the GSL, AWL, and EAP science lists. *TESOL International Journal*, 12(1), 177-192.
- Salim, M. A. B., Ibrahim, N. A. B., & Hassan, H. (2012). Language for tourism: A review of literature. *Procedia-Social and Behavioral Sciences*, 66, 136-143.
- Therova, D. (May, 2020). Review of Academic Word Lists. *TESL-EJ*, 24(1), 1-15. Retrieved from <https://www.tesl-ej.org/wordpress/issues/volume24/ej93/ej93a5/>
- Tongpoon-Patanasorn, A. (2018). Developing a frequent technical words list for finance: A hybrid approach. *English for Specific Purposes*, 51(1), 45-54.
- Balcı, A., Akar, F., & Öztürk, İ. (2019). Exploring the meaningful work level of academics in terms of different variables. *Educational Administration: Theory and Practice*, 25(2), 241-284.

- Wang, P. (2017). A corpus-based study of English vocabulary in Art research articles. *Journal of Art and Humanities (JAH)*, 6(8), 47-53.
- West, M. (1953). *A General Service List of English Words*, London: Longman, Green and Co.
- Ulla, M. B. (2018). English language teaching in Thailand: Filipino teachers' experiences and perspectives. *Issues in Educational Research*, 28(4), 1080-1094.
- Zahra, T. and Abbas, A. (2020). Development of Gender-based Pakistani Academic Word List (G-PAWL). *Journal of Social Sciences & Humanities*, 28(2), 114-128.

AUTHOR BIODATA

Darunee Yotimart, Ph.D. is an English lecturer at the Faculty of Arts, Thailand National Sports University, Chaiyaphum campus, Chaiyaphum, Thailand. Her research interests include Corpus linguistics, English for Specific Purposes (ESP) and identity in academic writing. She is the translator of Thailand National Sports University's Thai Boxing curriculum (Thai to English version).