

English's Travel Advisories: Corpus-based Study in English's Word/Phrase Forms

Rio Alfinda*,¹

¹ *Applied Linguistics Study Program, Faculty of Language, Arts, and Cultures,
Universitas Negeri Yogyakarta, Indonesia*

**Corresponding email: rioalfinda.2023@student.uny.ac.id*

Abstract

Keywords:
*corpus linguistics,
lexical bundles,
academic wordlist,
official text,
n-grams*

This study explores the academic-based words frequency, lexical bundles appearances, and the structures of lexical bundles in an official travel advisory files issued by British Government. Utilizing corpus-based programs for analyses, there are 50 travel advisories of 50 countries consisting of 31,243 tokens. The results present that verb phrases as the most frequent structure of lexical bundles, closely followed by noun phrases. These findings contrast with previous studies where noun or prepositional phrases are dominant. In addition, the study finds out that academic terms possess 7.70% of percentage in the texts, higher than undergraduate texts, but lower than highly prominent texts. The study points out the different lexical features of travel advisories and offers suggestions for further investigation at the terms. Furthermore, the findings have possibilities to be analysed, implemented, and widened in other research studies as this study provides insight into vocabulary usage and complementary materials for teaching and learning development.

|| Received: 30/12/2024

|| Accepted: 09/07/2025

|| Published: 30/07/2026

1. Introduction

Analysing vocabulary forms by implementing corpus linguistics application possesses importance to figure out the vocabulary's frequency and distributions. These lexical patterns and their occurrences lead to the insights of the words usage and connectivity in the context of vocabulary items. The analyses assist to identify repetitive patterns and habits in a language which draws the details and complexities of words. Therefore, one of the purposes is to give benefits to students, specifically English learners, to determine vocabulary use (Hongyan, 2018). The analysis can also be referenced for developing terms into dictionary (Handayani, Sukerti, & Wicaksana, 2022), specifically for English

for Specific Purposes in accounting sector. In this case, it is accessible to analyse and find out the use of academic words and lexical bundles in texts.

Biber and Barbieri (2007) identify lexical bundles as groups of words implemented to be in analyses to convey ideas whether in an essay or a spoken discourse. Aligned with the theory, Borucinsky and Pritchard (2022) add that lexical bundles possess roles to assist scrutinizing the meaning of language in specialized discourse. In the context of specialized written discourse, numbers of current research have offered some results in terms of word bundles (Shirazizadeh & Amirfazlian, 2021; Hassanzadeh & Tamleh, 2023; Jalali & Moini, 2014; Borucinsky and Pritchard, 2022).

Shirazizadeh and Amirfazlian (2021) investigate the lexical bundles specifically in applied linguistics' theses, articles, and textbooks. At this point, they analyse lexical bundles in organizational and functional characteristics needed for tutors/teachers to educate students about the bundles for academic writing preparation. This analysis includes subcategorized groups of research-oriented, text-oriented, and participant-oriented functions. On the other hand, Hassanzadeh and Tamleh (2023) investigate the lexical bundles from texts written by the native English speakers. According to their research findings, there are the most apparent bundles from the English writers such as in the present study, in the current study, in the case of, it is possible that, the results of the, and on the other hand. The findings of two studies reflect the most used vocabulary in English writing in general. Meanwhile, the other studies focus on specific fields on bundle analysis.

Borucinsky and Pritchard (2022) explored varied lexical bundles in four corpora (MEAT, MAS, MALSE, and MALSC). They find the structure of four-word bundles, such as at the end of, in accordance with the, and Croatian-based bundle koji se odnose na. On the other side, Jalali and Moini's (2014) study results four-word lexical patterns of a corpus CIMRA, which consists of the aims of the, aim of the present, and study was to evaluate.

In conclusion, lexical bundles analyses obtain crucial insights from various fields of study. The results of study can be aimed to comprehend vocabulary use in any forms of writing. However, the study of bundles is better if followed by the exploration of academic words scrutiny.

Academics words are determined as a part of measurement that focuses on the vocabulary production which the analysis can be summarized. According to Maamujav (2021), the possession of rich and complex lexicons is beneficial as this delivers to the text quality, specifically in terms of academic goals. The context of lexical variation takes account, in common, of the theory of Coxhead (2000) which is universally known as Academic Word List 570 or AWL.

As word lists are beneficial as its main position in English, especially English for Academic Purposes (EAP) (Nguyen & Coxhead, 2022), Academic Word List 570, or simplified Academic Word List (AWL), contains list of vocabulary in academic terms

which is beneficial for academic success across a lot of disciplines (Coxhead, 2000). In Studies on the AWL 570 have found out on how often the words emerge in academic texts, whereby they are applied in different subjects. These words are pivotal for examining and constructing academic writing as these lexicons are frequently applied in educational settings.

According to the definitions and studies above, it is more appropriate to find articles which have dealt with the combination of lexicon level, n-gram-based bundles, and academic words analyses. As most studies in lexical bundles focus on the frequency, structure, and functions, while the lexical level studies only focus on the number of classified words, this study combines two variables of lexical levels and bundles. In other words, this research concentrates on official texts correlated with travel advisory articles focusing on specific writing in the combined studies of border security, travel, and law. This hence leads to the development and comprehension of English vocabulary related with those fields of study.

Therefore, there are some questions emerging to find the purposes of study:

1. What are the most frequent lexical bundles appeared in the travel advisory texts?
2. What are structures of lexical bundles found?
3. What levels are the words explored in the text?

As the three questions have been constructed, this study tends to find analyses about lexicons which are useful for teachers, other researchers, and students to find vocabulary in context and level of the lexicons formed and implemented in official text.

2. The literature review

Understanding how lexical bundles and academic vocabulary operate across registers has been central to research on academic and professional English communication. This section reviews the main theoretical and empirical studies that inform the present analysis of lexical patterns and academic-word usage in official genres such as government travel advisories. It first summarizes previous findings on lexical bundles in academic and professional texts, followed by an overview of research into the Academic Word List (AWL) as an indicator of lexical sophistication. Then, it outlines the corpus-based tools and methodological approaches used to analyse bundle frequency and structure. Finally, it identifies research gaps in the current literature and establishes the rationale for investigating lexical bundles and AWL density in official public communication.

2.1. Studies on lexical bundles in academic and professional registers

This subsection reviews foundational and recent studies examining lexical bundles (Biber & Barbieri, 2007; Shirazizadeh & Amirfazlian, 2021; Borucinsky & Pritchard, 2022; Jalali & Moini, 2014). These works emphasize that bundle use varies across registers, disciplines, and communicative purposes, revealing distinctive structural and functional tendencies.

2.2 Academic vocabulary and the Academic Word List (AWL)

Research focusing on academic lexis (Coxhead, 2000; Maamuujav, 2021; Nguyen & Coxhead, 2022; Porwal & Devare, 2024) shows how AWL coverage indexes academicity and lexical sophistication. Studies demonstrate that AWL proportion and word-family range vary substantially by text type and discipline, providing quantitative measures of lexical density.

2.3 Corpus-based tool and analytical frameworks

Corpus studies typically integrate software such as AntConc and AntWordProfiler (Anthony, 2023) with frequency-based and manual coding approaches (Laufer & Nation, 1995). This subsection outlines those procedures and how they inform bundle extraction, functional classification, and academic-word profiling.

2.4 Gaps in literature and research positioning

Despite extensive work on bundles and AWL in scholarly writing, official genres—such as government travel advisories—remain understudied. The present study addresses this gap by combining bundle analysis and AWL profiling to describe the phraseological and lexical characteristics of this directive public register. In the ASEAN context, the approaching of the AEC (ASEAN Economic Community) 2015 is believed to have a considerable impact on academic and professional communication within a nation and across the nations.

3. Method

As the study found out the lexical bundles and the use of AWL 570 vocabulary in English's travel advisories, this research implemented document analysis by Creswell & Creswell (2023). They exemplify that this method is destined to analyse and to interpret the content of documents. Therefore, the study employed a qualitative method on texts.

3.1 Data Collection

The data were attained from the United Kingdom's travel advisory documents (<https://www.gov.uk/foreign-travel-advice>). There were 50 articles from 50 countries with 31,243 tokens gained to be analysed. The preference of British Government official text is considered pertinent as the language use is more varied and contains more decent structures than other English text from other countries. Hence, the text represents the congruous language use of English in general.

3.2 Data Analysis

To analyse data, this study implemented lexicon analysis apps Antconc and AntWordProfiler. Those apps are founded by Laurence Anthony (Anthony, 2023; see: <https://www.laurenceanthony.net/software>). The application of Antconc aims to the

exploration of lexical bundles. Meanwhile, AntWordProfiles is implemented to investigate academic words and percentage of word levels. As analysing data employ the steps proposed by Creswell & Creswell (2023), the analysis consists of (1) organizing and preparing data; (2) reading the entire data; (3) coding the data; (4) grouping into segments; and (5) representing/interpreting the data. The analysis provided followed the basis of Laufer & Nation (1995) and Biber & Barbieri (2007) and were discussed to compare with studies in the Introduction section.

4. Findings and discussion

4.1 Findings

This study investigated lexical bundles and level of English according to the official text retrieved from British's travel advisory documents. In accordance with the questions, the results are provided into three segments: most frequent lexical bundles; structures of the lexical groups; and words level.

1) The Most Frequent Vocabulary

Based on the data, lexical bundle results consist of four n-grams and five n-grams. Here is the data of four n-grams.

Table 1. Bundles of four-tail n-grams

<i>No</i>	<i>Data</i>	<i>Rank</i>	<i>Frequency</i>	<i>Range</i>
1	follow the advice of	1	43	1
2	and follow the advice	3	32	1
3	the advice of local	5	30	1
4	advice of local authorities	6	29	1
5	aware of your surroundings	6	29	1
6	try and carry out	71	11	1
7	the capability of (noun)	71	11	1
8	are likely to try	96	10	1
9	be aware of your	96	10	1
10	can in turn increase	96	10	1

The data provides four-tail n-grams implemented the most and the frequency the chunks in use. As displayed as examples, “follow the advice of” tops the list of four-word bundle in which the appearance rate is 43 times in the articles. Meanwhile, the other bundles that frequently come into sight are “and follow the advice” and “advice of local authorities”. In addition, there is a phrase that has specific mark “(noun)” in data number 47 (See *Appendix A1*). This mark is censored due to a possibly sensitive word as this study only focuses on linguistics feature, specifically the bundles used in English. Aside

from the four phrase forms, there is also form list of five-tail n-grams. Here is the data of five-word n-grams:

Table 2. *Bundles of five-tail n-grams*

<i>No</i>	<i>Data</i>	<i>Rank</i>	<i>Frequency</i>	<i>Range</i>
1	follow the advice of local	2	29	1
2	the advice of local authorities	2	29	1
3	keep up to date with	4	24	1
4	to carry out attacks in	4	24	1
5	are very likely to try	6	21	1
6	the advice of the local	80	10	1
7	this can in turn increase	80	10	1
8	could change at short notice	95	9	1
9	including in places frequented by	95	9	1
10	international shipping in the red	95	9	1

The data exemplifies that “follow the advice of local” and “the advice of local authorities” are the most implemented chunks with 29 times of frequency. Meanwhile “international shipping in the red” shows the least frequent phrase with nine times. Specifically, there are some phrases or chunks that have specific mark “(noun)” which the phrases are indicated sensitive as the use of language is the focus of this study.

2) Structures of Lexical Bundles

The following table displays the structures of 4-tail n-grams found in the dataset, indicating their frequency and type.

Table 3. *The Calculations of N-Grams Structures*

<i>Structures</i>	<i>Four-word</i>	<i>Five-word</i>	<i>Total Number of N-grams</i>	<i>Percentage</i>
VP	15	15	30	30%
NP	13	16	29	29%
ADJP	4	4	8	8%
ADV	3	3	6	6%
PP	7	4	11	11%
CLAUSE	7	7	14	14%
OTHER	1	1	2	2%
TOTAL	50	50	100	100%

The table offers data in which there is no significant differences between distributed numbers of verb phrases and noun phrases. In this case, additionally, clause places third as the most apparent structures of the lexical bundles as prepositional phrases, adjective

phrases, and adverb phrases. Meanwhile, the categorized “other” bundles are mostly conjunctions. As shown in **Table 3**, there is detail information about the groups of words with types of structures they possess in two additional tables. Here is table about the four-word n-Grams as provides:

Table 4. Structures of Four-tail N-Grams

<i>No</i>	<i>Data</i>	<i>Type of N-grams</i>	<i>Frequency</i>
1	follow the advice of local	VP	29
2	the advice of local authorities	NP	29
3	keep up to date with	VP	24
4	to carry out attacks in	PP	24
5	are very likely to try	VP	21
6	try and carry out	VP	11
7	the capability of (noun)	NP	11
8	are likely to try	VP	10
9	be aware of your	VP	10
10	can in turn increase	VP	29

The data explains about a frequency table with types of n-gram structures including verb phrases (VP), noun phrases (NP), adjective phrases (AdjP), adverb phrases (AdvP), prepositional phrases (PP), clause, and other. The dataset consists of 15 VP, as the most frequent type of four n-gram bundles, followed by 13 NP, seven clauses, seven prepositional phrases, four bundles of adjective, three adverb phrases, and one is categorized other (See **Appendix A4**).

On the other hand, the positional distributions of five-tail n-gram structures obtained in the present study are discussed to provide insights into their occurrence patterns and potential implications. Here is the dataset:

Table 5. Structures of Five-tail N-grams

<i>No</i>	<i>Data</i>	<i>Type of N-grams</i>	<i>Frequency</i>
1	follow the advice of local	VP	29
2	the advice of local authorities	NP	29
3	keep up to date with	VP	24
4	to carry out attacks in	VP	24
5	are very likely to try	VP	21
6	the advice of the local	NP	10
7	this can in turn increase	CLAUSE	10
8	could change at short notice	VP	9
9	including in places frequented by	PP	9
10	international shipping in the red	NP	9

In terms of five-gram structures, the data is divided into 16 noun phrases, 15 verb phrases, seven clauses, four prepositional phrases, four adjective phrases, three adverb phrases, and one other. According to the data, it indicates that the texts have a diverse range of phrase structures provided with some frequencies.

3) AWL 570

The Academic Word List 570 is crucial for academic writing as it covers a wide range of topics and is used across various disciplines. Understanding the distribution patterns of these words can help educators enhance students' academic language proficiency and improve curriculum development and language assessment in academic settings. Here is the data table of distribution percentage of AWL based on AntWordProfiler app:

Table 6. *The Stance of Academic Word List 570*

Lexical Category	Percentage
GSL Basic 1st 1000	68.20
GSL Basic 2nd 1000	8.15
AWL 570	7.70
Not-in-the-list	15.95
Total	100

In accordance with the **Table 6**, the distribution of AWL shows a significant portion with a moderate percentage of 7.70% academic terms. Another table offering additional details could be as follows:

Table 7. *The lexicon distribution of Academic Word List 570*

No	Data	Frequency
1	areas	132
2	authorities	127
3	security	115
4	demonstrations	82
5	aware	67
6	routes	12
7	conflict	12
8	civil	12
9	previous	11
10	involved	11

The data offers a list of academic words repeatedly applied in the texts according to the corpus analysis. The words “areas” (132 occurrences), “authorities” (127), and “security” (115) are implemented generally. These terms are significant in number of

appearances compared with the other terms such as “demonstrations” (82), “aware” (67), and “military” (55). The data shows how often certain topics are talked in the context of travel advisory texts. In addition, if deeply analysed, these academic terms can be found the level of words according to the Cambridge dictionary, here are ten words of the data that the level can be classified:

Table 8. *The level of academic words*

<i>No</i>	<i>Data</i>	<i>Frequency</i>
1	areas	B1
2	authorities	C1
3	security	B1
4	demonstrations	B2
5	aware	B2
6	sexual	B2
7	military	B2
8	transport	B1
9	media	B2
10	occur	B2

Based on the **Table 8**, the distribution of academic words shows a significant portion of B2 levels (six words), as followed by B1 and C1. It can be concluded that the level of words possesses higher quality of language in context of official website.

4.2 Discussion

1) The Structures of Bundles

The study has investigated that verb phrases (VP) dominate (30%) the structures of lexical bundles of travel advisory documents in corpus analysis using Antconc. However, the number of percentages of the dominating phrases shows no significant distinction as the noun phrases follow behind with 29% appearances. Therefore, this study possesses the differences with the other manuscripts. For example, the study conducted by Shirazizadeh & Amirfazlian (2021) results noun phrases and prepositional phrases as the main findings whereby those data lead the top list with 75% data.

There are two other studies showing similarity in terms of slight difference results between dominating bundles and the other. However, the exact results do not have similarity, specifically in detail of phrases distribution. Borucinsky & Pritchard (2022) find that the distribution verb phrases, and noun phrases are not significantly different with number of bundles of 626 and 618 for both verb phrases and noun phrases. Nevertheless, the bundles do not excel the frequency. Shockingly, prepositional phrases top the distribution of frequency with 657 total bundles. Their study is totally different with the other research studies as is this study. On the other hand, Jalali & Moini (2014)

state that the bundles between noun and verb phrases are different with percentage of 24.02% and 19.89%.

The domination of verb phrases is in accords with pragmatics purpose in which the directive illocutionary is applied for the readers. Based on study shown by Plaza (2014) about EU maritime legislation's speech acts, directive verbs have several criteria: addressee-oriented and obligation and duty statements. In other English-speaking country's travel advisories, directive acts are able to be found, such as the U.S. travel advisory (see: <https://travel.state.gov/content/travel/en/traveladvisories/traveladvisories/australia-travel-advisory.html>) and the Canadian Government advisory (see: <https://travel.gc.ca/destinations/indonesia>).

5) The Academic Word List (AWL) Percentage

The academic word list includes 570 groups of academic-related vocabulary implemented in texts for academic purposes (Coxhead, 2000). Coxhead (2000) add that the list is made up from 10% words of the Academic Corpus in which those words possess appearances in the 20 of 28 fields. Aside from Coxhead's, the list of academic words is found in term of University Word List (UWL), Which Laufer & Nation (1995) find out, whereby 836-word families are constructed. However, the Coxhead's AWL has been widely used as reference, including for the digitalized analysis app for corpus linguistics study such as Anthony's AntWordProfiler program (Anthony, 2023). Therefore, this study implements the references, both Coxhead's finding and Anthony's app, to analyse the lexical bundles. According to the results, the study explores the percentage of AWL appeared. As 7.70 percent of the result are found, this study shows the difference with other study. Firstly, the findings of this study are higher with the results of study done by Ha (2019). On average, the finding provided by Ha, which focuses on the AWL, reveals percentage of 3.56 percent. However, as the study searches for the analysis on undergraduate's written texts, it is common if the comparison between this research and Ha's finding is quite significant. On the other hand, this study possesses lower rates of percentage with study conducted by Porwal & Devare (2024) whereby they explore lexical density that includes academic word list. Have they conducted analyses on the H-index $\geq$ 40-based articles, their data result shows 14.7 percent academic words in high quality papers. Further, the low-quality paper also shows higher appearance of academic word with 11.5% academic words. It is hence concluded that the study, if compared with undergraduate-level and prominent writer texts, vary as the travel advisory documents possess higher percentage than undergraduate text level.

5. Conclusion

This study investigated distinct data results on the frequency of lexical bundles and academic word list (AWL) vocabulary as the structures of the vocabulary bundles /

groups included. In structure of bundles, the verb phrases top the list in this study (30%) as noun bundles are slightly followed behind with 29% of percentage. Therefore, this study contrasts with other research articles underlining noun phrases or prepositional phrases as the most apparent bundles. In terms of AWL, the number of AWL percentage explored in this study differs from previous research. According to the data discussion, it is pointed out that this study is higher than undergraduate texts, but lower than analysis on prominent articles.

As the study on lexical bundles on official texts are hardly to be explored, it is suggested that the deepened study in this context can be widened. Furthermore, the study results can be implemented for English learners and teachers for further research or teaching materials. The teaching of English in the HE is characterised by its nature that closely relates to particular academic and professional settings. The teaching of English in the HE is characterised by its nature that closely relates to particular academic and professional settings.

References

- Alasmarty, A. (2024). Sustaining vocabulary knowledge growth through corpus-generated lists of lexical bundles and keywords in the Law of Contracts. *Heliyon* 10(9). <https://www.sciencedirect.com/science/article/pii/S2405844024059759>
- Anthony, L. (2023). *AntConc (Version 4.2.4) [Computer Software]*. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Anthony, L. (2023). *AntWordProfiler (Version 2.1.0) [Computer Software]*. Waseda University. Available from <https://www.laurenceanthony.net/software>
- Biber, D. & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes*, 26(3), 263-286. https://jan.ucc.nau.edu/biber/Biber/Biber_Barbieri_2007.pdf
- Borucinsky, M. & Pritchard, B. (2022). Lexical bundles in maritime texts. *ICAME Journal*, 46, 5 - 17. <https://doi.org/10.2478/icame-2022-0001>.
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213-238.
- Gezegin, B. B. (2019). Lexical bundles in published research articles: A corpus-based study. *Journal of Language and Linguistic Studies*, 15(2), 520-534. <https://doi.org/10.17263/jlls.586188>
- Ha, H. S. (2019). Lexical Richness in EFL Undergraduate Students' Academic Writing. *English Teaching*, 74(3), 3-28. <https://doi.org/10.15858/engtea.74.3.201909.3>
- Handayani, L. N. C., Sukerti, G. N. A., & Wicaksana, K. A. B. (2022). Integrating dictionary on taxation accounting terminology as self-learning kit in English for specific purposes. *Journal of Applied Studies in Language*, 6(2), 144-153. <https://doi.org/10.31940/jasl.v6i2.668>

- Hassanzadeh, M., & Tamleh, H. (2023). The use of lexical bundles by native English authors in applied linguistics: A corpus-driven study. *Language Related Research*, 13(6), 541-569. <http://lrr.modares.ac.ir/article-14-50310-en.html>
- Hongyan, L. (2018). A STUDY ON CORPUS-BASED EFL VOCABULARY TEACHING. *ISLLAC: Journal of Intensive Studies on Language, Literature, Art, and Culture*. <https://doi.org/10.17977/UM006V2I22018P02>
- Jalali, Z. S., & Moini, M. R. (2014). Structure of lexical bundles in introduction section of medical research articles. *Procedia-Social and Behavioral Sciences*, 98, 719-726. <https://doi.org/10.1016/j.sbspro.2014.03.473>
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, 16(3), 307-322. https://openaccess.wgtn.ac.nz/articles/journal_contribution/Vocabulary_size_and_use_Lexical_richness_in_L2_written_production/12560396/files/23412143.pdf
- Maamuujav, U. (2021). Examining lexical features and academic vocabulary use in adolescent L2 students' text-based analytical essays. *Assessing Writing*. <https://doi.org/10.1016/J.ASW.2021.100540>.
- Nguyen, T., & Coxhead, A. (2022). Evaluating multiword unit word lists for academic purposes. *ITL - International Journal of Applied Linguistics*. <https://doi.org/10.1075/itl.21041.ngu>.
- Porwal, P., & Devare, M. H. (2024). Scientific impact analysis: Unraveling the link between linguistic properties and citations. *Journal of Informetrics*, 18(3), 101526. <https://doi.org/10.1016/j.joi.2024.101526>
- Plaza, S. M. (2014). Teaching performative verbs and nouns in eu maritime regulations. *Procedia-Social and Behavioral Sciences*, 141, 90-95. <https://doi.org/10.1016/j.sbspro.2014.05.017>
- Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and variation. *Journal of English for Academic Purposes*, 49, 100946. <https://doi.org/10.1016/j.jeap.2020.100946>

Appendices

This part contains extended data from the results. To differentiate these groups with those in the result section, the tables are indicated with letter A that is followed by numerical orders (e.g. A1, A2).

Appendix A1

Table A1. Bundles of four-tail n-grams

No.	Type	Rank	Freq.	Range
1	follow the advice of	1	43	1
2	and follow the advice	3	32	1
3	the advice of local	5	30	1
4	advice of local authorities	6	29	1
5	aware of your surroundings	6	29	1
6	up to date with	9	25	1
7	keep up to date	10	24	1
8	very likely to try	11	22	1
9	are very likely to	13	21	1
10	could be indiscriminate including	13	21	1
11	indiscriminate including in places	13	21	1
12	likely to try to	13	21	1
13	try to carry out	13	21	1
14	in places visited by	20	20	1
15	local media reports and	20	20	1
16	to date with local	20	20	1
17	places visited by foreigners	25	19	1
18	with local media reports	25	19	1
19	cannot be ruled out	28	18	1
20	including in places visited	28	18	1
21	stay aware of your	28	18	1
22	in the red sea	34	16	1
23	make substantive concessions to	35	13	1
24	the advice of the	35	13	1
25	British government considers that	45	12	1
26	for your presence is	45	12	1
27	increase the risk of	45	12	1
28	of the British government	45	12	1
29	of the local authorities	45	12	1
30	policy of the British	45	12	1
31	presence is unlikely to	45	12	1
32	reason for your presence	45	12	1
33	the British government considers	45	12	1
34	the reason for your	45	12	1
35	the risk of further	45	12	1
36	there is a risk	45	12	1

37	your presence is unlikely	45	12	1
38	act makes payments to	71	11	1
39	advice of the local	71	11	1
40	at short notice you	71	11	1
41	builds the capability of	71	11	1
42	likely to try and	71	11	1
43	standing policy of the	71	11	1
44	the long standing policy	71	11	1
45	to try and carry	71	11	1
46	try and carry out	71	11	1
47	the capability of (noun)	71	11	1
48	are likely to try	96	10	1
49	be aware of your	96	10	1
50	can in turn increase	96	10	1

Appendix A2

Table A2. Bundles of five-tail n-grams

No.	Type	Rank	Freq.	Range
1	follow the advice of local	2	29	1
2	the advice of local authorities	2	29	1
3	keep up to date with	4	24	1
4	to carry out attacks in	4	24	1
5	are very likely to try	6	21	1
6	(noun) could be indiscriminate including	6	21	1
7	be indiscriminate including in places	6	21	1
8	likely to try to carry	6	21	1
9	(noun) are very likely to	6	21	1
10	try to carry out attacks	6	21	1
11	local media reports and follow	14	20	1
12	up to date with local	14	20	1
13	local media reports and follow	14	20	1
14	media reports and follow the	14	20	1
15	date with local media reports	18	19	1
16	to date with local media	18	19	1
17	with local media reports and	18	19	1
18	aware of your surroundings keep	21	18	1
19	including in places visited by	21	18	1
20	of your surroundings keep up	21	18	1
21	stay aware of your surroundings	21	18	1
22	in places visited by foreigners	27	16	1

23	indiscriminate including in places visited	27	16	1
24	very likely to try to	27	16	1
25	British government considers that paying	35	12	1
26	follow the advice of the	35	12	1
27	for your presence is unlikely	35	12	1
28	or secure your safe release	35	12	1
29	policy of the British government	35	12	1
30	reason for your presence is	35	12	1
31	the British government considers that	35	12	1
32	the reason for your presence	35	12	1
33	the risk of further (noun)	35	12	1
34	your presence is unlikely to	35	12	1
35	at short notice you should	56	11	1
36	builds the capability of (noun)	56	11	1
37	likely to try and carry	56	11	1
38	long standing policy of the	56	11	1
39	standing policy of the British	56	11	1
40	the long standing policy of	56	11	1
41	try and carry out attacks	56	11	1
42	advice of the local authorities	80	10	1
43	can in turn increase the	80	10	1
44	increase the risk of further	80	10	1
45	is not to make substantive	80	10	1
46	the advice of the local	80	10	1
47	this can in turn increase	80	10	1
48	could change at short notice	95	9	1
49	including in places frequented by	95	9	1
50	international shipping in the red	95	9	1

Appendix A3

Table A3. The Calculations of N-Grams Structures

STRUCTURES	Four-word	Five-word	Total Number of n-grams	Percentage
VP	15	15	30	30%
NP	13	16	29	29%
ADJP	4	4	8	8%
ADV	3	3	6	6%
PP	7	4	11	11%
CLAUSE	7	7	14	14%
OTHER	1	1	2	2%
TOTAL	50	50	100	100%

Appendix A4

Table A4. Structures of 4-tail N-Grams

No.	Type	Type of N-grams	Freq.
1	follow the advice of	VP	43
2	and follow the advice	OTHER	32
3	the advice of local	NP	30
4	advice of local authorities	NP	29
5	aware of your surroundings	AdjP	29
6	up to date with	PP	25
7	keep up to date	VP	24
8	very likely to try	AdvP	22
9	are very likely to	VP	21
10	could be indiscriminate including	VP	21
11	indiscriminate including in places	AdjP	21
12	likely to try to	AdjP	21
13	try to carry out	VP	21
14	in places visited by	AdvP	20
15	local media reports and	NP	20
16	to date with local	PP	20
17	places visited by foreigners	NP	19
18	with local media reports	PP	19
19	cannot be ruled out	VP	18
20	including in places visited	PP	18
21	stay aware of your	VP	18
22	in the red sea	AdvP	16
23	make substantive concessions to	VP	13
24	the advice of the	NP	13
25	British government considers that	CLAUSE	12
26	for your presence is	CLAUSE	12
27	increase the risk of	VP	12
28	of the British government	PP	12
29	of the local authorities	PP	12
30	policy of the British	NP	12
31	presence is unlikely to	CLAUSE	12
32	reason for your presence	NP	12
33	the British government considers	CLAUSE	12
34	the reason for your	NP	12
35	the risk of further	NP	12
36	there is a risk	CLAUSE	12
37	your presence is unlikely	CLAUSE	12

38	act makes payments to	CLAUSE	11
39	advice of the local	NP	11
40	at short notice you	PP	11
41	builds the capability of	VP	11
42	likely to try and	AdjP	11
43	standing policy of the	NP	11
44	the long standing policy	NP	11
45	to try and carry	VP	11
46	try and carry out	VP	11
47	the capability of (noun)	NP	11
48	are likely to try	VP	10
49	be aware of your	VP	10
50	can in turn increase	VP	10

Appendix A5

Table A5. Structures of five-tail n-grams

No.	Type	Type of N-grams	Freq.
1	follow the advice of local	VP	29
2	the advice of local authorities	NP	29
3	keep up to date with	VP	24
4	to carry out attacks in	VP	24
5	are very likely to try	VP	21
6	(noun) could be indiscriminate including	NP	21
7	be indiscriminate including in places	VP	21
8	likely to try to carry	VP	21
9	(noun) are very likely to	NP	21
10	try to carry out attacks	VP	21
11	local media reports and follow	NP	20
12	up to date with local	AdjP	20
13	local media reports and follow	NP	20
14	media reports and follow the	NP	20
15	date with local media reports	NP	19
16	to date with local media	PP	19
17	with local media reports and	AdvP	19
18	aware of your surroundings keep	AdjP	18
19	including in places visited by	PP	18
20	of your surroundings keep up	PP	18
21	stay aware of your surroundings	VP	18
22	in places visited by foreigners	AdvP	16
23	indiscriminate including in places visited	AdjP	16

24	very likely to try to	AdvP	16
25	British government considers that paying	CLAUSE	12
26	follow the advice of the	VP	12
27	for your presence is unlikely	CLAUSE	12
28	or secure your safe release	OTHER	12
29	policy of the British government	NP	12
30	reason for your presence is	CLAUSE	12
31	the British government considers that	CLAUSE	12
32	the reason for your presence	NP	12
33	the risk of further (noun)	NP	12
34	your presence is unlikely to	CLAUSE	12
35	at short notice you should	CLAUSE	11
36	builds the capability of (noun)	VP	11
37	likely to try and carry	AdjP	11
38	long standing policy of the	NP	11
39	standing policy of the British	NP	11
40	the long standing policy of	NP	11
41	try and carry out attacks	VP	11
42	advice of the local authorities	NP	10
43	can in turn increase the	VP	10
44	increase the risk of further	VP	10
45	is not to make substantive	VP	10
46	the advice of the local	NP	10
47	this can in turn increase	CLAUSE	10
48	could change at short notice	VP	9
49	including in places frequented by	PP	9
50	international shipping in the red	NP	9

Appendix A6

Table A6. The stance of Academic Word List 570 in AntWordProfiler

Lexical Category	Percentage
GSL Basic 1st 1000	68.20
GSL Basic 2nd 1000	8.15
AWL 570	7.70
Not-in-the-list	15.95
Total	100

Appendix A7

Table A7. The lexicon distribution of Academic Word List 570 (Coxhead, 2000)

No	Item	Frequency
----	------	-----------

1	areas	132
2	authorities	127
3	security	115
4	demonstrations	82
5	aware	67
6	sexual	57
7	military	55
8	transport	53
9	media	50
10	area	39
11	occur	38
12	incidents	35
13	region	33
14	targeted	31
15	targets	30
16	monitor	30
17	sites	28
18	major	28
19	secure	27
20	vehicle	26
21	target	25
22	instructions	25
23	illegal	25
24	injured	24
25	contact	23
26	significant	19
27	policy	19
28	credit	18
29	access	18
30	restrictions	17
31	vehicles	16
32	accommodation	16
33	financial	15
34	capability	15
35	release	14
36	periods	14
37	items	14
38	isolated	14
39	aid	14
40	tensions	13
41	response	13

42	releasing	13
43	regions	13
44	individuals	13
45	targeting	12
46	routes	12
47	conflict	12
48	civil	12
49	previous	11
50	involved	11

Appendix A8

Table A8. The level of academic words

No	Item	Level
1	areas	B1
2	authorities	C1
3	security	B1
4	demonstrations	B2
5	aware	B2
6	sexual	B2
7	military	B2
8	transport	B1
9	media	B2
10	occur	B2