LEARN Journal: Language Education and Acquisition Research Network
ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 17, No: 2, July – December 2024

Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

Hope You Guys Enjoyed this Vlog: Lexical Bundles and Moves within English Travel Vlog Discourse

Piyapong Laosrirattanachai[a], Piyanuch Laosrirattanachai[b,*]

[a] piyapong.l@ku.th, Faculty of Hospitality Industry, Kasetsart University, Thailand
[b] piyanuch.la@ku.th, Faculty of Hospitality Industry, Kasetsart University, Thailand
[*] Corresponding author, piyanuch.la@ku.th

APA Citation:
Laosrirattanachai, P. & Laosrirattanachai, P.(2024). Hope you guys enjoyed this vlog: lexical bundles and moves within English travel vlog discourse. *LEARN Journal: Language Education and Acquisition Research Network, 17*(2), 249-278.

Received 26/02/2024

Received in revised form 05/04/2024

Accepted 25/0/2024

ABSTRACT

Lexical bundles and moves are essential for vloggers to communicate clearly and purposefully within travel vlog discourse. It is crucial for L2 learners and practitioners aiming to enter the industry to master these bundles and understand the moves used in creating travel vlogs. This corpus-based study compiled a list of 239 four-word lexical bundles serving as fixed slots and their 98 variable slots from the Travel Vlog Corpus, which comprises 434,809 running words. These bundles were categorised by function: 79 as stance expressions, 75 as discourse organisers, 80 as referential expressions, and 5 as special conversational functions. The study also identified four move types and their 19 component steps necessary for creating travel vlogs. It emphasised that lexical bundles and moves are critical knowledge with important functions for generating travel vlog discourse. The study concluded by proposing pedagogical implications and discussing future research directions.

Keywords: corpus-based analysis, keyword analysis, lexical bundles, move analysis, travel vlog

Introduction

The vlog, a visual variant of a blog, has gained popularity among users and attracted scholarly interest from various disciplines (Ladhari et al., 2020; Lee & Watkins, 2016; Snelson, 2013). Among the different types of vlogs, the travel vlog is particularly popular. These are publicly accessible, tourist-generated videos, often found on social media platforms, that typically showcase the vloggers' travel experiences (Xu et al., 2021). Platforms like YouTube are common for creating travel vlogs, where some vloggers become celebrities with up to 7.5 million subscribers. The aspiration to become a renowned travel vlogger draws many beginners into the industry; however, few achieve their goals (Ruiz-Gomez et al., 2021). Success in vlogging requires various skills and qualifications, including self-presentation (Marwick, 2015), viewer engagement, travel content creation (Hou, 2018), and, from our perspective, language skills essential for conducting travel vlogs effectively.

In the English for Specific Purposes (ESP) classroom, learners increasingly focus on careers involving social media, such as those of an influencer, YouTuber, and vlogger. Neophytes interested in entering the travel vlog industry often begin with a trial-and-error approach, which frequently leads to failure. Beneficial guidelines should be prepared for ESP learners using teachers' language and discourse knowledge to empower learners and enhance their chances of becoming successful travel vloggers.

This situation poses a challenge for L2 learners or practitioners in the tourism field who aspire to become travel vloggers. To assist learners in becoming recognised travel vloggers and developing it into a profession, they need not only creative thinking but also the ability to communicate fluently and naturally in English. This requires producing language that achieves effective communication goals. Learners should possess essential vocabulary knowledge, lexical bundles (LBs) that facilitate smooth communication, and the ability to design content with components (referred to as 'moves' in this study) that help convey messages from the vlogger to the viewers, thereby achieving their communication objectives successfully. In the scenarios described, developing a list of LBs and move structures within travel vlog discourse is essential. Therefore, the current study proposes a list of the LBs and moves required to create travel vlogs, which should be beneficial for teaching and learning purposes, especially at educational institutions specialising in training hospitality personnel or L2 practitioners.

Keywords and lexical bundles

Vocabulary is crucial for language production and serves as a foundational element of knowledge. Technical travel vocabulary often comes

to mind initially. A review of the literature shows that numerous word lists related to tourism have been developed (Laosrirattanachai & Ruangjaroon, 2021). Therefore, redeveloping a technical tourism word list might be redundant, as learners could directly benefit from existing materials. Instead, teachers should focus on enhancing learners' knowledge of the LBs needed to create travel vlogs.

A lexical bundle is a sequence of words or a phrase that frequently appears together within a language, viewed as multi-word entities or segments with specific meanings or functions. LBs range from basic collocations to complex phrases or idiomatic expressions, playing a significant role in language analysis and corpus linguistics by illustrating language usage patterns and aiding in understanding word and phrase usage in context. LBs are often analysed to gain insights into language proficiency, discourse organisation, and linguistic diversity across different contexts or genres. In second language acquisition, these LBs are stored in long-term memory (Bolinger, 1976; Ellis & Simpson-Vlach, 2008; Wray, 2002; Wray & Perkins, 2000), proving highly beneficial for English communication in various contexts, including travel vlog discourse. Terms like building block of discourse, chunks, clusters, formulaic language, formulaic sequences, multi-word expression, multi-word unit, n-grams, phrasicon, recurrent sequences, recurrent word combination, and word sequences are also used to describe similar phenomena (Altenberg, 1998; Cortes, 2002; De Cock, 1998; De Cock et al., 1998; Hyland, 2008a; Schmitt & Carter, 2004; Schmitt et al., 2004; Scott, 1996; Staples et al., 2013; Stubbs, 2007a; 2007b). Language, both spoken and written, is recognised as being formulaic in nature (Ellis, 1996; Granger & Meunier, 2008; Pawley & Syder, 1983; Qin, 2014; Sinclair, 2004; Wray, 2002), making LBs essential for language learners due to their frequent use in everyday communication (Biber et al., 2004). However, there is no universal set of LBs applicable in all contexts; distinct communicative needs require different LBs (Biber & Barbieri, 2007), leading to numerous LB studies aimed at enhancing teaching and learning (Cortes, 2004; 2006; Li & Schmitt, 2009).

The number of words in LBs can vary from 3 to 6, depending on the research objectives and the researchers. To compile a list of LBs, researchers may use various methods, one common approach being to analyse the frequency and range values of LBs in the corpus (Biber et al., 2004; Cortes 2013; Hyland 2008a) through, using software like Sketch Engine (Kilgarriff et al., 2014) for facilitation. Frequency refers to how often a phrase recurs within a corpus to be considered a bundle, whereas range, another criterion for determining LBs, indicates the diversity of sources in which the phrase appears. In written genre studies, the frequency cut-off points for bundle identification vary, with thresholds set at a minimum of 10 times per million words (Biber et al., 1999), 20 times per million words (Byrd & Coxhead, 2010;

Cortes, 2004; Hyland, 2008a), or even 40 times per million words (Biber et al., 2004). In spoken registers, Benson and Coxhead (2022) determined that a bundle should appear at least five times in their rugby corpus. The frequency cut-off point may also depend on the LB's word count; the more words in the LB, the lower the frequency cut-off per million words can be set (Cortes, 2013). Regarding range, appearing in 3–5 texts or 10 per cent of the texts is commonly used as a criterion for identification (Biber & Barbieri, 2007; Biber et al., 1999; Cortes, 2013; Hyland, 2008b).

Another method for creating LBs involves using keywords from the corpus as a foundation to form LBs comprising those specific keywords (Grabowski, 2015). This approach necessitates conducting a keyword analysis before developing a list of LBs. Unlike absolute frequency, keyword analysis assesses word occurrence based on relative frequency, using statistical techniques like log-likelihood (LL). Keywords (KWs) are words used more frequently in a target corpus (often a smaller, specialised corpus) than in a reference corpus (a larger corpus, such as the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA)) (Bondi & Scott, 2010; Scott & Tribble, 2006). These keywords often constitute the central elements of longer word sequences (Grabowski, 2015).

Due to the significant benefits of developing a list of LBs for teaching L2 learners, extensive research has focused on creating comprehensive LB lists. However, the literature review indicates a preference for studying and developing LBs in academic rather than professional contexts. There is also a tendency to prioritise written genres over spoken ones. Research on LBs in academic written genres includes studies on LBs in research articles (Shahriari, 2017), theses and dissertations (Hyland, 2008a), textbooks (Biber et al., 2004), and student writing (Cortes, 2008; Durrant, 2017), with some research investigating LBs across these written genres (Shirazizadeh & Amirfazlian, 2021). In contrast, research on LBs in spoken genres includes studies on university spoken registers (Biber & Barbieri, 2007), conversation and academic prose (Conrad & Biber, 2005), and spoken academic EFL genres (Wang, 2017). There are also studies on LBs in professional domains, such as legal genres (Breeze, 2013) and pharmaceutical discourse (Grabowski, 2015). This review highlights a research gap in the creation of LB lists for spoken genres, especially in professional domains, underscoring the need for further study and development in this area.

The presentation of LBs can vary in format. The first method involves presenting them in alphabetical order, offering simplicity and convenience, although it might pose challenges for self-directed learners without guidance. The second method entails presenting LBs according to their connection with different moves in the discourse (Cortes, 2013).

Another method presents LBs based on structural and functional classifications (Biber et al., 2004). For structural classifications, LBs are organised according to types of structures, such as verb phrase fragments (e.g., "*is based on the*"), noun phrase and prepositional phrase fragments (e.g., "*a little bit of*"), and dependent clause fragments (e.g., "*if you want to*"). Functional classifications, conversely, organise LBs according to their communicative functions. Biber et al. (2004) propose common functional categories in the taxonomy of LBs, including:

Stance expressions: These bundles have two sub-functions. The first is epistemic stance, indicating the level of knowledge about various information and expressing certainty, doubt, or probability, such as *"I don't know why"*. The second sub-function is attitudinal/modality stance, expressing the speaker's attitude, intention, command, obligation, or evaluation towards a topic or proposition, with examples like *"it is important to"*, *"I believe that"*, and *"it seems that"*.

Discourse organisers: These bundles also comprise two sub-functions. The first is topic introduction/focus, signalling that the speaker is about to introduce a new topic, as seen in examples like *"on the other hand*", *"let's have a look*", and "*as you can see*". The second is topic elaboration/clarification, used to expand on the current topic, with examples including *"as a result*", *"for example*", and "*in conclusion*".

Referential expressions: Consisting of four sub-functions, the first is identification/focus bundles, such as "*that's one of the*", focusing and emphasising the following noun phrase. The second is imprecision bundles, like *"and things like that*", indicating unclear or similar things. The third is specifying attributes, for example, *"have a number of"* and *"the nature of"*, specifying quantity, amount, or logical relationships. The final sub-function is time/place/text-deixis bundles, referring to specific places, times, or earlier content, such as "*according to the data*" and "*previous research has shown*".

Special conversational functions: These bundles serve specific conversational purposes like greeting, seeking clarification, and expressing politeness, with examples such as *"I said to him*", *"What are they doing?""*, and "*thank you very much*".

The final method presents LBs based on the degree to which the slots in the bundles are fixed or variable (Benson & Coxhead, 2022; Renouf & Sinclair, 1991; Staple et al., 2013), helping avoid redundancy in LB presentation. For instance, rather than listing *"what I want to*", *"I want to do*", and "*want to do is*" separately, they can be represented as "*what [I want to do] is*", where *"I want to do*" is the fixed slot, and *"what*" and "*is*" are variable slots.

Move analysis

Genres are defined as communicative events or typified forms of discourse that respond to specific communication needs (Paltridge, 2013). Genre analysis examines the contextualised linguistic behaviour within formal academic or professional settings. These genres represent distinct forms of communication with specific communicative intentions, recognised and shared among members of the relevant professional or academic communities. As a result, language usage within each genre adheres to established norms to achieve the communicative goals specific to a particular field or social context, manifesting in consistent structural patterns (Bhatia, 2004). Genres reflect the cultural norms and operational procedures within disciplinary and organisational frameworks, underscoring the social functions associated with disciplinary, professional, and institutional norms (Martin et al., 1987; Miller, 1984). Moreover, each disciplinary and professional genre exhibits unique characteristics, often shaped by a blend of textual, discursive, and contextual elements (Bhatia, 1993; Swales, 1990).

Move analysis is a method used to analyse a genre, with the term "move" referring to a component or section of a text, as proposed by Swales (1990). Each move has a specific communicative function, contributing to the text's overall purpose (Kanoksilapatham, 2007). Moves may contain one or several "submoves" or "steps" that support their purpose (Ho, 2017; Stoller & Robinson, 2013). Collectively, these moves and steps construct a text or discourse (Henry & Roseberry, 1998; Tardy, 2011). Genres can be categorised in various ways, such as written and spoken genres, or academic and professional genres, depending on the study's purpose and the researchers' approach. In language teaching and learning, move analysis serves as a tool to identify key sections within a communicative event, enhancing learners' understanding of effective communication across different genres.

Studies using move analysis have examined a range of genres, including research articles (Kanoksilapatham, 2005; Lu et al., 2021; Maswana et al., 2015; Swales, 1981), online review responses (Ho, 2018; Laosrirattanachai & Laosrirattanachai, 2024; Panseeta & Watson Todd, 2014; Thumvichit & Gampper, 2019), webchat exchanges (Xu & Lockwood, 2021), email responses (Van Herck et al., 2022), conference or thesis presentations (Hu & Liu, 2018; Rowley-Jolivet & Carter-Thomas, 2005), TED talks (Chang & Huang, 2015; Li & Li, 2021), job recruitment posters (Phattisiri et al., 2023), podcasts (Ye, 2021), and press conferences (Laosrirattanachai & Laosrirattanachai, 2023). Despite the breadth of genres analysed, research on travel vlog discourse remains limited. Given the increasing interest in travel vlog discourse among language learners and the nuanced understanding

experienced individuals in professional fields have over genres compared to newcomers, this study aims to apply move analysis to uncover the move structures within travel vlog discourse. This analysis will focus on travel vlogs produced by renowned vloggers with substantial followings, offering guidance to L2 learners or practitioners in the tourism field who aim to become proficient travel vloggers.

Research questions

The current study aims to explore the LBs and move structure patterns relevant to ESP students and L2 practitioners in the tourism field, particularly those aspiring to become travel vloggers. The findings are expected to offer valuable insights for L2 learners and practitioners of English in this domain. The research questions are:

1. Which LBs are frequently used in the Travel Vlog Corpus (TVC)?
2. What specific function does each LB serve in communication?
3. What is the move structure of English travel vlog discourse?

Methodology

Travel Vlog Corpus preparation

To develop a list of LBs and unveil the move structures in travel vlog discourse, a spoken corpus of travel vlogs was compiled. The top 20 travel vloggers with the highest number of followers on YouTube were selected. The follower count was based on the total number from when their channels were started until December 2023. The 10 vlogs from each vlogger with the most views were transcribed, resulting in 200 vlogs with durations ranging from 10.04 to 27.21 minutes. The selection of popular and highly viewed vlogs ensured the identification of quality and engaging content well received by viewers. These 200 vlogs were transcribed from YouTube, which offers features like "Show transcript" and "CC" for accessing subtitles. Initially, the "Show transcript" feature was used to export the transcripts. However, upon reviewing, it was found that the transcriptions were not 100% accurate. Therefore, to ensure complete accuracy, the authors manually proofread and edited all 200 travel vlog transcripts three more times (firstly by the first author, secondly by the second author, and finally jointly). Table 1 presents an overview of the TVC.

Table 1

Description of the TVC

N=200	Time (minutes)	Size (running words)
Max	27.21	6,474
Min	10.04	1,545
Average	16.14	2,174
Total	**3,228.58**	**434,809**

The average size of the vlogs in the TVC was 2,174 tokens, with a range from 1,545 to 6,474 tokens, totalling 434,809 tokens. Despite the varied sizes of these sub-corpora, they effectively represent the vlogging style and language use of the vloggers. While a larger corpus can provide a more comprehensive view of a specific field (Benson & Coxhead, 2022), the relatively small size of the TVC is deemed more suitable for identifying specific linguistic patterns and their relationship to the language used in this particular area (Koester, 2006).

Identifying lexical bundles and their functions within travel vlog discourse

In this study, the list of LBs was developed by initially identifying the keywords (KWs) used as core components, aiming to ensure that the resultant LBs are pertinent to KWs frequently used in travel vlog discourse. Understanding these LBs will enable learners to use them more naturally and effectively in travel vlog discourse.

A list of KWs was first created, with these KWs forming the basis of recurrent word combinations within the travel vlog discourse. To extract the KWs, the TVC was compared with the British National Corpus (BNC), using the LL value with the Key-BNC programme (Graham, n.d.). Using the BNC as a comparative corpus is deemed adequate for analysing and identifying linguistic differences between two corpora (Johnson & Esslin, 2006; Scott, 2001). Therefore, keyword analysis in this study involved a comparison with the BNC. Before this analysis, the TVC underwent editing to align its spelling with British English, necessitating changes such as altering "*traveler*" to "*traveller*". After generating the list of KWs, three additional criteria were applied to refine the KWs that would serve as core components of the LBs in travel vlog discourse:

After generating the list of KWs, three additional criteria were applied to refine the KWs that would serve as core components of the LBs in travel vlog discourse:

1. Keyword analysis: The top-100 KWs based on the LL value were identified as potential core components of the LBs.

2. Stop word deletion: Determiners, auxiliary verbs, pronouns, and proper names were removed from the top-100 KWs.

3. Dispersion: A minimum dispersion value of 0.6 across the 200 travel vlogs was required.

Given the high sensitivity of LL values to corpus size, determining a cut-off point for classifying words as keywords is a contentious issue (Pojanapunya & Watson Todd, 2016). For this study, which aimed to identify 50 potential KWs for developing a prospective list of LBs, the initial selection of the top-100 KWs was later refined based on subsequent criteria.

Function words are typically used as stop words in developing word lists because they are non-technical and considered less important. However, in this study, some function words, like "*like*", which may be frequently used by vloggers to encourage viewers to click the "*like* button", might hold enough significance to be deemed KWs of travel vlog discourse. Consequently, only determiners, auxiliary verbs, pronouns, and proper names were excluded from the top-100 KWs.

Dispersion is measured in various ways, but many scholars prefer Juilland and Chang-Rodríguez's (1964) method for developing word lists, as it is a common approach to assess dispersion (Gardner & Davies, 2014; Lei & Liu, 2016; Nation, 2006). The dispersion value ranges from 0 (indicating an extremely unbalanced distribution) to 1 (denoting a completely balanced distribution). This study adopted Dang et al.'s (2007) dispersion cut-off point of 0.6, considering words above this threshold as potential core components for the LBs. After applying the aforementioned criteria to the top-100 KWs, the number of KWs exceeded 50. Those ranked beyond the top-50 were considered as alternatives. It is possible that some KWs may not qualify as core components of the LBs. In such instances, the KW not serving as the core component of 'an LB would be omitted, and the KW ranked 51st could replace it.

After compiling the list of KWs, the list of LBs was identified using the following criteria:

1. Length and degree: Defined as clusters of four or more words, with the four-word unit serving as a fixed slot.

2. Unit of counting: Type.

3. Association: The fixed slot must contain a KW as its core component.

4. Frequency: An LB must appear four or more times in the TVC.

5. Range: An LB must be used by at least four vloggers.

6. Judgement-based method: LBs rated with mode values of 2 or 3 by three experts were included in the final list.

7. Functional classification: LBs were classified based on their functions.

Research has shown that shorter strings frequently occur within longer bundles (Byrd & Coxhead, 2010; Cortes, 2004, 2013; Hyland, 2008a; Staple et al., 2013). To prevent redundancy, this study presented the list of LBs by distinguishing between fixed and variable slots in the bundles (Benson & Coxhead, 2022; Renouf & Sinclair, 1991; Staple et al., 2013). Another crucial aspect was the length of the bundles. Studies indicate that two-word units are common, three-word units are also frequent but often form part of larger bundles, particularly four-word units, which occur significantly more often than five- or six-word units (Beng & Keong, 2014; Biber et al., 1999; Biber et al., 2004; Csomay, 2013; Hyland, 2008a; Kopaczyk, 2012; O'Keeffe et al., 2007). Consequently, four-word units were identified as the fixed slots in this study, with the possibility of extending these into longer sequences with variable slots.

The frequency and range cut-off points were established based on prior LB list development research (Benson & Coxhead, 2022; Biber & Barbieri, 2007; Cortes, 2013; Hyland, 2008b). Given the TVC's small size, the frequency cut-off was set at four or more occurrences. For the range criterion, to mitigate bias from one vlogger using certain bundles across multiple vlogs, the study used 20 vloggers as the unit for range consideration, instead of the 200 vlogs. Therefore, LBs had to be used by at least four different vloggers to be selected.

The LB rating scale, encompassing the LBs that met the previous five criteria, was evaluated by three experts in the current study. These experts had over five years of experience in teaching tourism, particularly in conducting tours and creating vlog content, using English as the medium of instruction. Given that the list of LBs was intended to facilitate travel vlog creation, the focus was not on the technical nature of the LBs but on their necessity for conveying messages in travel vlogs. Consequently, the developed list of LBs did not include technical multi-word units; rather, it featured general expressions frequently used in travel vlog production. The lexical bundle rating scale, detailed in Table 2, consisted of three levels.

Table 2

Description of lexical bundle rating scale

Scale	Description
1	A multi-word unit that is generally known by L2 users of English and require less effort or special training to use appropriately
2	A multi-word unit that is generally known by L2 users of English but requires special training or suggestion to use in the travel vlog creation

Scale	Description
3	A multi-word unit that is unique and is strictly related to travel vlog creation

After the LBs were rated, those meeting the criteria were classified into four functional categories: stance expressions, discourse organisers, referential expressions, and special conversational functions (Biber et al., 2004). This functional classification is intended to assist learners in using the LBs appropriately for communication when conducting travel vlogs, ensuring that they can effectively convey their messages within the vlog discourse.

Exploring moves within travel vlog discourse

The travel vlogs analysed in this study were full-length vlogs released on YouTube, recognised by the discourse community, with their use of English being appropriate for travel and educational purposes. To explore communicative discourse components in the 200 travel vlogs, the following procedures were adopted:

1. Judgement-based method by authors: the first and second authors collaborated to examine the moves and steps in the sample travel vlogs.

2. Judgement-based method by connoisseurs: the same three experts who had verified the travel vlog LBs in the previous stage were tasked with coding the travel vlogs using the moves and steps proposed by the authors.

To identify the moves and steps in the vlogs, one vlog was randomly selected from each vlogger, resulting in 20 vlogs as the sample. The first and second authors independently watched and then reviewed the transcripts of these 20 vlogs. Based on their analysis, the two authors developed a draft coding protocol separately. Subsequently, they compared and discussed their drafts and created a revised version of the coding protocol. This second draft was tested on another 20 vlogs. Following a thorough discussion of the pilot results, adjustments were made before finalising a master inventory of the moves and steps.

The refined coding protocol was then verified for reliability. Three trained coders coded the final sample of vlogs, which comprised 3 vlogs from each vlogger, totalling 60 vlogs (30% of the entire collection). They used a codebook with yes-no categories. The identification agreement rate for the moves and steps was 92.13%, indicating excellent agreement. Once this rate was established, the coding protocol was applied to the entire set of vlogs.

Results

Lexical bundles and their functions within travel vlog discourse

Generating keywords

To complement the development of the LB list, the top-50 keywords were identified. The study revealed that out of the top-100 keywords, based on the LL value, 29 were determiners, auxiliary verbs, pronouns, and proper names, leaving 71 words for dispersion value analysis. Each of these 71 words had a dispersion value of at least 0.6 across 200 vlogs. Despite these findings, the aim was to compile 50 keywords as core components of the LBs. Therefore, only the top-50 keywords by LL value from the 71 were selected as specific keywords. Examples of these keywords, which met the required criteria and ranked within the top 50 by LL value, include "*enjoyed*", "*favourite*", "*video*", "*vlog*", "*travelling*", "*subscribe*", and "*hotel*".

Lexical bundles

Once the top-50 keywords were identified, the LB list was created. Using the first five criteria to determine the LBs, the analysis yielded 1,278 bundles. Final assessments were conducted by three experts who rated the necessity of the LBs for vloggers. The rating scale results indicated that 239 four-word units, functioning as fixed slots, and their 98 variable slots, forming longer bundles, were considered essential for vlog creation, based on a mode value of 2 or 3.

Table 3

Examples of lexical bundles, including fixed and variable slots for the keywords

Variable slot	Fixed slot	Variable slot
so	*subscribe* if you want	to go, to see
hope you, if you	guys enjoyed this *vlog*	

Table 3 presents examples of LBs derived from the keywords "subscribe" and "vlog." It is clear that the LB generated from "subscribe" is "*subscribe if you want*", which can be expanded with other variable slots to form longer LBs. Similarly, the LB from "vlog" is "*guys enjoyed this vlog*", which can be extended to longer LBs like *"hope you guys enjoyed this vlog*" and "*if you guys enjoyed this vlog*", as demonstrated in the following excerpt:

*"**Hope you guys enjoyed this vlog**, and if you did, it would mean the world to me if you subscribed to my YouTube channel and checked out my Instagram."*

Functional taxonomy of the LBs within travel vlog discourse

After all LBs were collated, they were categorised into functional groups, divided into four main functions and eight sub-functions, as identified by Biber et al. (2004). These functions include stance expressions (epistemic stance and attitudinal/modality stance), discourse organisers (topic introduction/focus and topic elaboration/clarification), referential expressions (identification/focus, imprecision, attribute specification, and time/place/text-deixis), and special conversational functions. The comprehensive list of LBs, classified according to their functions, is presented in Table 4.

Table 4

Lexical bundles and their functions within travel vlog discourse

No.	Lexical bundles		
	Variable slot	Fixed slot	Variable slot
Stance expressions: Epistemic stance bundles			
1		no idea what is	going
Stance expressions: Attitudinal/Modality stance bundles			
1		and hit the subscribe	button
2		go get some food	
3		make sure you subscribe	
4	do	not forget to subscribe	
5		subscribe if you have	not, not already
6		subscribe if you like	
7	so	subscribe if you want	to go, to see
8		subscribe to my channel	
9	so	excited to be here	
10	hope you, if you	guys enjoyed this vlog	
11		happy to be here	
12		have to come here	
13	so	I am really excited	to
14		I am so excited	
15		I am super excited	
16		I am very excited	
17		I hope you guys	
18		it is absolutely beautiful	
19		it is actually pretty	
20		it is pretty cool	
21		it is pretty good	
22		it is really beautiful	
23		it is really cool	
24		it is really good	
25		it is really nice	
26		it is so beautiful	
27		it is so cool	
28		it is so good	
29		it is so nice	
30		it is so so	
31		it looks so good	
32		it looks so nice	
33		kind of hard to	
34		and we will see	
35	and we	are about to go	
36		are going to check	
37		are going to get	
38		are going to go	
39		are going to see	
40		are ready to go	
41		going to be exploring	
42		going to show you	the
43		going to spend the	
44		going to take a	
45		gonna have a quick	

No.	Lexical bundles		
	Variable slot	Fixed slot	Variable slot
46		gonna show you a	
47		great way to get	
48		my all time favourite	
49		of the most amazing	
50	one of	of the most beautiful	
51		of the most incredible	
52		one of my favourite	
53		one of our favourite	
54		so much better than	
55		that is pretty good	
56		that is really good	
57		that is really nice	
58		that is so cool	
59		that is so good	
60	and	then we are gonna	check, get, go, see
61		this is my favourite	
62		this is pretty cool	
63		this is really good	
64		this is so cool	
65		this is so good	
66	nice	to meet you guys	
67	I wanted, want	to show you guys	
68		to watch the sunset	
69		we are actually going	
70	and then, now	we are going to	check, get, go, see, be travelling
71		we need to go	
72	I	will see you guys	
73		will show you guys	
74		you are gonna see	
75	if	you enjoyed the video	
76		you have to go	
77	if	you like this video	
78		you need to know	
Discourse organisers: Topic introduction/focus bundles			
1		and we are here	
2		as soon as you	get
3	and, so	as you can see	there
4		here you can see	
5		I am just here	
6		I have been here	
7		I have been travelling	
8		I have seen so	
9		if you can see	
10		if you come here	
11		just look at the	
12		let me give you	a
13		let me show you	
14		let us check it	out
15		let us do it	
16		let us do this	
17		let us get on	
18		let us get out	
19		let us get some	
20		let us go back	
21	so	let us go check	it, it out, out
22		let us go for	
23		let us go get	
24		let us go inside	
25		let us go see	
26		let us go to	
27		let us have a	look
28		let us head to	
29		let us see how	
30	so	let us see if	we
31		let us see this	
32	so	let us see what	
33		let us take a	
34		let us try it	
35		let us try this	
36		look at all the	
37		look at all these	
38		look at the size	of
39		look at the view	
40		look at this place	
41		look at this view	
42		now it is time	for, to
43		now we are at	
44		now we are back	
45		now we are going	
46		now we are gonna	check, get, go, see
47		now we are heading	to
48		now we are here	
49		now we are in	
50		now we are on	
51		now we are walking	
52		now you can see	
53		see if we can	get, get a
54		see what else we	
55		see what we can	
56		so check this out	
57		so let us go inside	

No.	Lexical bundles		
	Variable slot	Fixed slot	Variable slot
58		so let us try	
59		so make sure you	
60		we have been here	
61		we have been walking	around
62		we made it here	
63		when you are travelling	
64	if	you are new here	
65		you can actually see	
66		you can come here	
67		you can see there	
68		you guys can see	
Discourse organisers: Topic elaboration/clarification bundles			
1		know how to say	it
2		know what I mean	
3		know what it is	
4		know what that is	
5		know what to do	
6		know what to expect	
7		so that we can	
Referential expressions: Identification/focus bundles			
1		is the most beautiful	
2		it is a good	place
3		there is a beautiful	
Referential expressions: Imprecision bundles			
1		and it is like	
2		and it looks like	
3		it does not look	like
4		it is almost like	
5		it is literally like	
6		it is more like	
7	I have	never seen anything like	this
8	see	what it is like	
9		what it looks like	
Referential expressions: Bundles specifying attributes			
1		a lot of food	
2		a lot of people	here
3		a lot of you	guys
4		a lot to see	
5		so many of them	
6		so many things to	
7		so much more to	
8	there is	so much to see	
9		so one of the	
10		there are a lot	of
11		there are so many	
12		there is a big	
13		there is a bunch	of
14		there is a couple	

No.	Lexical bundles		
	Variable slot	Fixed slot	Variable slot
15		there is a few	
16		there is a little	
17		there is a lot	more, of
18		there is a ton	
19		there is a whole	bunch of
20		there is lots of	
21		there is no way	
22		there is not much	
23		there is so much	to, to see
Referential expressions: Time/place/text-deixis bundles			
1		all over the place	
2		around the world and	
3		back to my hotel	
4		back to our hotel	
5		back to the hotel	
6	the	centre of the city	
7		down to the beach	
8		get out of here	
9		go to a place	
10		go to the beach	
11		here in the city	
12		for the next vlog	
13	as you all know	in every travel vlog	
14	I will see you, see you	in the next video	
15	I will see you, I hope to see you, hope to see you, see you	in the next vlog	
16	welcome, and welcome	to another travel vlog	
17		vlog series out here	in
18		is the only place	
19		it is right there	
20		just in front of	
21		just walking around the	
22		just walking through the	
23	the	last time I was	here
24		get back to the	
25		going to be heading	
26		gonna head back to	the
27		just arrived at the	
28		just got out of	
29	the, in the	middle of the city	
30		my first time here	
31		outside of the city	
32		part of the city	
33		part of the island	

No.	Lexical bundles		
	Variable slot	Fixed slot	Variable slot
34		right across the street	
35		right here in the	
36		right in front of	
37		right in the heart	of
38		right in the middle	of, of the
39		right on the beach	
40	so	stay tuned for that	
41		this place is so	
42	last	time I was here	
43		way to get around	
44		which you can see	
45		you are coming here	
Special conversational functions (for travel vlogs)			
1		oh my god look	
2		oh my gosh guys	
3		thank you for watching	this vlog
4		thank you so much	for watching this vlog
5		what is up guys	

Out of the total 239 developed LBs, when categorised by functions, they were distributed as follows: 79 were stance expressions, with 1 epistemic stance bundle and 78 attitudinal/modality stance bundles; 75 were discourse organisers, which included 68 topic introduction/focus bundles and 7 topic elaboration/clarification bundles; 80 were referential expressions, comprising 3 identification/focus bundles, 9 imprecision bundles, 23 attribute specification bundles, and 45 time/place/text-deixis bundles. Additionally, there were 5 bundles under special conversational functions, serving purposes such as politeness, exclamations, and greetings. Examples include "*thank you for watching this vlog*", *"oh my gosh guys*", and "*what is up guys."*

It is noteworthy that, although contractions are common in speech, they are not consistently used. This study presents the LBs in their full form rather than as contractions, allowing users to decide on their practical application.

Moves within travel vlog discourse

The move analysis identified the communicative discourse components, specifically the moves and steps used in English travel vlog discourse. This analysis primarily focused on the frequency of these moves and steps across 200 vlogs. It is important to note that some moves and steps recurred within a single vlog, necessitating a clear definition of "frequency". In this study, irrespective of the number of occurrences within a vlog, each move or step was counted once, setting the maximum possible frequency at 200. Additionally, the study determined the obligatory nature of moves and steps. According to Kanoksilapatham's (2005) criteria, a move or step appearing in at least 60% of the corpus (120 vlogs) was considered obligatory. Table 5 presents the moves and steps identified in travel vlogs.

Table 5

Moves and steps within travel vlog discourse

Move Step	Frequency (N=200)	%	Remark
1 Vlogger presentation	*190*	*95*	*Obligatory*
A Greeting viewers	113	56.5	Optional
B Announcing the tourist destination	190	95	Obligatory
C Introducing vlogger	67	33.5	Optional
2 Drawing viewers' attention	*140*	*70*	*Obligatory*
A Setting the scene	80	40	Optional
B Outlining vlogs	56	28	Optional
C Providing tourist destination background	94	47	Optional
3 Starting the journey	*200*	*100*	*Obligatory*
A Giving large-scale transportation information	60	30	Optional
B Describing accommodation	130	65	Obligatory
C Giving small-scale transportation information	158	79	Obligatory
D Presenting surrounding atmosphere	197	98.5	Obligatory
E Introducing famous landmarks	186	93	Obligatory
F Recommending where to eat and interesting dishes	173	86.5	Obligatory
G Participating in tourist activities	101	50	Optional
H Short talks with locals	84	42	Optional
4 Closure	*167*	*83.33*	*Obligatory*
A Evaluating the tourist destination	99	49.5	Optional
B Recommending other interested vlogs	38	19	Optional
C Thanking and farewell	77	38.5	Optional
D Requesting viewers' reaction	104	52	Optional
E Making future appointment	154	77	Obligatory

Table 5 shows that the move analysis identified four move types and their 19 component steps. At the move level, all four were considered obligatory. At the step level, seven steps were deemed obligatory, while the remaining 12 were classified as optional.

The *vlogger presentation* (1) move aims to orient viewers using three steps. The *greeting viewers* (1A) step typically occurs at the start of the vlogs, where the vlogger greets the viewers. The *announcing the tourist destination* (1B) step informs viewers about the vlog's main tourist destination. This is followed by the *introducing vlogger* (1C) step, where the vlogger introduces themselves and shares their experience with the tourist destination.

(1) 1A: *What is up? Welcome to today's video.* (Vlog 128)
(2) 1B: *We are now in Kiev Ukraine.* (Vlog 128)
(3) 1C: *My name is XXX and in 2015 I had the extreme pleasure of travelling to nine Asian countries and spending almost seven and a half months abroad.* (Vlog 23)

The *drawing viewers' attention* (2) move seeks to captivate the audience through various methods. This move comprises three steps. The *setting the*

scene (2A) step includes: 1) the vlogger asking viewers if they are familiar with the tourist destination, 2) the vlogger engaging with pedestrians to gather their knowledge about the place, and 3) highlighting the reasons that make the tourist destination interesting. Another step within this move is the *outlining vlogs* (2B), where the vlogger compiles highlight shots to summarise the vlog. Lastly, the *providing tourist destination background* (2C) step involves giving essential information about the tourist destination before the journey begins.

(4) 2A: *What do you think the least visited country in the world is? Well, after a quick Google search, it turns out that neither myself nor most people on the street had ever even heard of this extremely tiny country in the South Pacific called Tuvalu.* (Vlog 1)

(5) 2B: *Today in Alaska, we are cruising up the Endicott Arm. I woke up this morning to see chunks of ice floating past my balcony, and there are seals on these chunks of ice. This afternoon, I'll be getting into one of the Zodiac boats with the Ventures team, getting a closer look at the ice, and hopefully a closer look at some seals as well.* (Vlog 71)

(6) 2C: *Switzerland is perfect from having the cleanest water to the most unbelievable landscapes. Switzerland is a real-life fairy tale. It's constantly at the top of the list as the country with the highest quality of life, but also at the top of the list for being the most expensive. On top of all this, Switzerland is a neutral country. They don't pick sides in wars, and they remain peaceful by any means necessary.* (Vlog 102)

The *starting the journey* (3) move is the core of travel vlog discourse, being more extensive than the other moves. This move consists of eight steps, each contributing detailed information about the tourist destination. The journey typically commences from the vlogger's hometown, leading to the *giving large-scale transportation information* (3A) step, where the vlogger discusses travel methods from the origin to the tourist destination, often covering international or interregional transit. Upon reaching the destination, the *describing accommodation* (3B) step occurs, showcasing the vlogger's lodging. The *giving small-scale transportation information* (3C) step then provides details on local transportation options. The *presenting surrounding atmosphere* (3D) step offers viewers a glimpse of the panoramic scenery.

(7) 3A: *Just a quick stop in Fiji. And this's the domestic airport, pretty small. Boarding flight number two.* (Vlog 1)

(8) 3B: *Welcome to our hotel room. This is so cool. We're staying at like a traditional Japanese hotel or like inn. Okay, yes, so the rules are this little level up here is the half so you take your shoes off here and I believe if you open up these, there should be some sandals.* (Vlog 46)

(9) 3C: *All right, now this boat is gonna take us 30 minutes to our home. So far, it's looking pretty good.* (Vlog 12)

(10) 3D: *There's some pretty cool buildings. All the doors and kind of the framework, all made out of wood. Beautiful architecture! It looks awesome but there's no stalls here. I guess you have to come here during the nighttime. That's the best time to visit when all the hustle and bustle is going on.* (Vlog 66)

The *introducing famous landmarks* (3E) step is central to the vlogs, detailing key tourist sites for the audience. The *recommending where to eat and*

interesting dishes (3F) step highlights renowned eateries and local culinary specialties. To engage the audience, the *participating in tourist activities* (3G) step depicts various leisure activities, capturing the vlogger's experiences. Finally, the *short talks with locals* (3H) step provides deeper insights into the destination's culture and lifestyle.

(11) 3E: *Now, the Red Square is the main central square here in Moscow, with festivals, demonstrations, parades, you name it. It's actually home to the Kremlin, which is behind that red wall behind me. That is where the government of Russia is, the president.* (Vlog 92)
(12) 3F: *Okay. So, we're back at the cafe and we're having the best Koshary in all of Egypt. Koshary is like the national dish of Egypt. What's in Koshary? Rice, macaroni, pasta, tomato sauce, cool. All right. Wow! Looks good. And this is the hot sauce. This is the sauce of the Koshary.* (Vlog 120)
(13) 3G: *And now, Taewon's going to show me how they use this machine to separate the rice from the stem.* (Vlog 13)
(14) 3H: Vlogger: *So, you live here with your mom?*
Locals: *Yeah, my mom, my dad, my sister. Just a simple life. That's all we need though.*
Vlogger: *You grew up here? This is your childhood home, right here? That's amazing.* (Vlog 1)

The *closure* (4) move typically occurs at the end of vlogs. In this move, the vlogger wraps up the vlog by sharing their thoughts and feelings about the trip in the *evaluating the tourist destination* (4A) step. To promote other vlogs, the vlogger uses the *recommending other interested vlogs* (4B) step which involves promoting additional vlogs available on the channel. The vlogger then expresses gratitude and bids farewell to the audience in the *thanking and farewell* (4C) step. Additionally, the *requesting viewers' reaction* (4D) step is frequently used to encourage viewer engagement, such as soliciting comments, hitting the like button, sharing vlogs, and channel subscription. Last, the vlogger engages the audience for future content in the *making future appointment* (4E) step.

(15) 4A: *Even though I only had four days, Afghanistan was truly the most impactful country I've ever visited.* (Vlog 4)
(16) 4B: *Okay, I'm signing out the video now. If you'd like to know about more things that you can do in the city or in the state of Rio, check out my playlist that I'll put in the description and in the comment section because there I have around like 20 videos from this trip and a previous trip doing all sorts of like cool stuff here.* (Vlog 145)
(17) 4C: *Thanks guys so much for watching this vlog. I hope you enjoyed it.* (Vlog 32)
(18) 4D: *Guys, if you enjoyed today's video, please hit that like button.* (Vlog 21)
(19) 4E: *Let's get lost* again in the next one.* (Vlog 28) (* The vlogger uses *'lost'* to refer to *'travel'* in his channel.)

Discussion

Compared to the spoken genre, the written genre appears to have received more attention from researchers. However, with the advancement

of corpus techniques, the importance of corpus-based studies on the spoken genre has been increasingly recognised. Access to more state-of-the-art and digitalised materials has become easier, leading to new studies on specific discourses, such as the travel vlog discourse in the current study.

In this study, using keywords as fundamental elements for generating LBs has proven to be highly beneficial for learners. This approach facilitates the development of LBs related to keywords frequently encountered in travel vlogs. Keywords are crucial in steering the analysis towards specific linguistic patterns, thus enhancing efficiency by focusing on linguistically relevant units. Consequently, researchers can identify LBs quickly and effectively, ensuring their contextual relevance and reflecting language usage in the TVC. Furthermore, the study and analysis of LBs based on keywords reveal recurrent language usage patterns, illuminating the functional aspects of language in travel vlog discourse. Additionally, learners can gradually expand their knowledge from individual words to bundles, understanding the interrelation between these levels, which promotes a deeper comprehension and aids in smoother communication.

To present the LBs developed in this study, a functional classification was used. The functional taxonomy of LBs offers insights into how language functions to convey meaning, express attitudes, organise discourse, and maintain cohesion in communication. Analysing LBs within these functional categories enhances the understanding of language use across different contexts and genres, aiding language learners in their effective utilisation for communication.

The move analysis findings reveal several interesting aspects. First, the step "*Announcing the tourist destination*" is identified as obligatory in the "*Vlogger presentation*" move, likely because the tourist destinations vary in each episode, necessitating this information for the viewers. "Conversely, steps like *Greeting viewers*" and "*Introducing vlogger*" may be omitted, as viewers might already be familiar with the channel's regular vlogger. Second, while *Drawing viewers' attention*" comprises a single move without an obligatory step, it is deemed essential overall, as it plays a crucial role in capturing the viewers' interest and preparing them for the journey. Lastly, the frequent appearance of steps "*Presenting surrounding atmosphere*", "*Introducing famous landmarks*", and "*Recommending where to eat and interesting dishes*" underscores their importance in offering a shared experience of the journey and conveying valuable information and insights about significant and popular locations and culinary recommendations.

Pedagogical implications

LBs are considered challenging for non-native English speakers (De Cock 2002; Granger 1998). It is argued that mastering these expressions is crucial for achieving proficiency and fluency comparable to native speakers, making them essential in language education and acquisition (O'Keeffe et al., 2007). While the LBs developed in the current study may seem straightforward, not all necessitate explicit teaching. However, in line with the research objectives of developing LBs to aid learners in producing travel vlogs efficiently, these LBs have been included not due to their technicalness, but because they frequently occur in travel vlogs and can provide valuable guidance for learners creating such content. Moreover, incorporating these LBs in teaching can foster both incidental and intentional learning. Instructors might focus on teaching the more challenging LBs through intentional learning methods, allowing learners to acquire the simpler LBs incidentally. This approach will enable learners to engage with both learning types, familiarise themselves with new LBs, and improve their understanding and use of known LBs in communication.

The study suggests several recommendations for ESP classrooms. Initially, learning by memorising and analysing should be introduced to ESP learners, who must recognise all the moves and possible steps. Subsequently, they should analyse each move and step, including its purpose and function, and identify where LBs could be applied.

The next step involves learning by observing and comparing. Teachers should assign learners to view approximately 3–5 vlogs both autonomously and in class, ensuring that they pay close attention to the details. After observing authentic vlogs, teachers should engage learners with questions to encourage them to share their observations and how these relate to the study's findings previously discussed.

The final step is crucial as it can significantly enhance the learning experience and potentially inspire learners to pursue a career as travel vloggers. Learning by practicing and doing is essential. Producing a vlog involves budgeting for funding, time, ideas, and equipment, making practice a practical starting point. Group work is advised, as members can monitor each other's use of LBs and scripts and exchange opinions on the team's ideas, moves, and steps. Initially, practice should involve acting as a vlogger in front of the class with a hand-made storyboard to help students become comfortable using LBs in travel vlog contexts. Subsequently, the storyboard should be replaced with PowerPoint slides that learners can use while role-playing a vlogger in various scenes. At the course's conclusion, having become adept at using LBs and creating vlogs, learners should produce their

finest vlogs as a final project. These vlogs should be uploaded to YouTube for public viewing and feedback.

Limitations

The primary limitation of this study was the use of the Key-BNC programme for keyword analysis. This program compares the target corpus with the BNC, thus the keyword results are inherently based on British English vocabulary. Although the researchers converted the spelling to British English *"*(e.g., "*favorite"* to "*favourite*"), they did not alter the vocabulary that differs between British and American English (e.g., "*lift*" vs "*elevator*"). This oversight might have excluded words specific to American English from the keyword list, potentially affecting the comprehensiveness of the LBs identified in this study. Another limitation relates to the selection of the top 20 travel vloggers based on their follower counts. Since the count was measured from the inception of their channels, travel vloggers with longer-running channels are likely to have higher follower counts, thus favouring their inclusion in the top 20. This methodology disadvantages newer travel vloggers. However, it is also notable that experienced travel vloggers, having refined their content through trial and error, are likely to produce higher-quality vlogs.

Recommendations for Future Studies

Future studies should explore vlogs from diverse perspectives. Research could focus on comparing vlogs created in different languages and by vloggers of various nationalities, considering how cultural characteristics may influence viewer attraction across different global regions.

Conclusion

The corpus-driven approach in this study revealed a substantial number of factors previously unexplored in travel vlog discourse, with LBs and moves being paramount. The investigation highlighted that approximately 239 four-word bundles serve as fixed slots, and their 98 variable slots are crucial for L2 learners or practitioners in travel vlog discourse. Presenting LBs as four-word bundles with variable slots for expansion, similar to this study and the approach of Benson and Coxhead (2022), offers a more manageable and less overwhelming format for learners, compared to a comprehensive list encompassing all forms. Additionally, understanding the structure and communicative functions within travel vlog discourse necessitates familiarity with four move types and their 19

component steps. Both bundles and moves are essential, serving as valuable resources in educational course design.

Acknowledgements

This research work was supported by a research grant from Kasetsart University, Kamphaeng Saen Campus, Thailand (grant number: KPS-RDI 2022-019).

About the Authors

Piyapong Laosrirattanachai: an Assistant Professor at Kasetsart University in Thailand. His areas of interest are corpus linguistics, vocabulary, word list, multi-word units, move analysis, and English for hospitality.

Piyanuch Laosrirattanachai: a lecturer in the English for Service Industry programme at Kasetsart University, Thailand. Her current research interests include vocabulary, multi-word units, move analysis, and sociolinguistics.

References

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 101–122). Oxford University Press.

Beng, C. O. S., & Keong, Y. C. (2014). A corpus study of structural types of lexical bundles in MUET reading texts. *3L: The Southeast Asian Journal of English Language Studies, 20*(2), 127–140. http://dx.doi.org/10.17576/3L-2014-2002-11

Benson, S., & Coxhead, A. (2022). Technical single and multiword unit vocabulary in spoken rugby discourse. *English for Specific Purposes, 66*, 111–130. https://doi.org/10.1016/j.esp.2022.02.001

Bhatia, V. K. (1993). *Analysing genre - Language use in professional settings.* Longman.

Bhatia, V. K. (2004). *Worlds of written discourse: A genre-based view*. Bloomsbury Publishing.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*(3), 263–286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Longman.

Bolinger, D. (1976). Meaning and memory. *Forum Linguisticum, 1*, 1–14.

Bondi, M., & Scott, M. (Eds.). (2010). *Keyness in Texts*. John Benjamins. https://doi.org/10.1075/scl.41

Breeze, R. (2013). Lexical bundles across four legal genres. *International Journal of Corpus Linguistics, 18*(2), 229–253. https://doi.org/10.1075/ijcl.18.2.03bre

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL, 5*, 31–64.

Chang, Y., & Huang, H. (2015). Exploring TED talks as a pedagogical resource for oral presentations: A corpus-based move analysis. *English Teaching & Learning, 39*(4), 29–62.

Conrad, S., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica, 20*(2004), 56–71. https://doi.org/10.1515/9783484604674.56

Cortes, V. (2002). Lexical bundles in Freshman composition. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131–145). John Benjamins Publishing Company.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397–423. https://doi.org/10.1016/j.esp.2003.12.001

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education, 17*(4), 391–406. https://doi.org/10.1016/j.linged.2007.02.001

Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora, 3*(1), 43–57.

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes, 12*, 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

Csomay, E. (2013). Lexical bundles in discourse structure: A corpus-based study of classroom discourse. *Applied Linguistics, 34*(3), 369–388. https://doi.org/10.1093/applin/ams045

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning, 67*(4), 959–997. https://doi.org/10.1111/lang.12253

De Cock, S. (1998). A recurrent word combination approach to the study of formulae in the speech of native and non-native speakers of English. *International Journal of Corpus Linguistics, 3*(1), 59–80. https://doi.org/10.1075/ijcl.3.1.04dec

De Cock, S., Granger, S., Leech, G., & McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. In S. Granger (Ed.), *Learner English on computer* (pp. 67–79). Longman.

De Cock, S. (2002). Pragmatic prefabs in learners' dictionaries. In *Proceedings of the Tenth EURALEX International Congress, Copenhagen, August 2002, Denmark*, 13–17.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics, 38*(2), 165–193. https://doi.org/10.1093/applin/amv011

Ellis, N. C. (1996). Sequencing in SLA: phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91–126.

Ellis, N., & Simpson-Vlach, R. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *Teachers of English to Speakers of Other Languages Quarterly, 42*(3), 375–396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Grabowski, Ł. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes, 38*, 23–33. https://doi.org/10.1016/j.esp.2014.10.004

Graham, D. (n.d.). *Key-BNC*. [Computer software]. Centre for Research Support. http://crs2.kmutt.ac.th/Key-BNC/

Granger, S. (1998). The computer learner corpus: A versatile new source of data for sla research. In S. Granger (Ed.), *Learner English on computer* (pp. 3–18). Longman.

Granger, S., & Meunier, F. (Eds.). (2008). *Phraseology: An interdisciplinary perspective*. John Benjamins.

Henry, A., & Roseberry, R. L. (1998). An evaluation of a genre-based approach to the teaching of EAP/ESP writing. *TESOL Quarterly, 32*(1),

Ho, V. (2017). Achieving service recovery through responding to negative online reviews. *Discourse & Communication, 11*(1), 31–50. https://doi.org/10.1177/1750481316683292

Ho, V. (2018). Exploring the effectiveness of hotel management's responses to negative online comments. *Lingua, 216*, 47–63. https://doi.org/10.1016/j.lingua.2018.10.004

Hou, M. (2018). Social media celebrity and the institutionalization of YouTube. *Convergence: The International Journal of Research into New Media Technologies, 25*(3), 534–553.

Hu, G., & Liu, Y. (2018). Three-minute thesis presentations as an academic

genre: A cross-disciplinary study of genre moves. *Journal of English for Academic Purposes, 35*, 16–30. https://doi.org/10.1016/j.jeap.2018.06.004

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41–62. https://doi.org/10.1111/j.1473-4192.2008.00178.x

Johnson, S., & Esslin, A. (2006). Language in the news: Some reflections on keyword analysis using WordSmith Tools and the BNC. *Leeds Working Papers in Linguistics and Phonetics, 11*.

Juilland, A. G., & Chang-Rodríguez, E. (1964). *Frequency dictionary of Spanish words*. Mouton.

Kanoksilapatham, B. (2005). Rhetorical move structure of biochemistry research articles. *English for Specific Purposes, 24*(3), 269–292. https://doi.org/10.1016/j.esp.2004.08.003

Kanoksilapatham, B. (2007). Introduction to move analysis. In D. Biber, U. Connor, & T. A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 23–41). John Benjamins.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography, 1*(1), 7–36. doi:10.1007/s40607-014-0009-9

Koester, A. (2006). *Investigating workplace discourse*. Routledge.

Kopaczyk, J. (2012). Applications of the lexical bundles method in historical corpus research. In P. Pezik (ed.), *Corpus data across languages and disciplines* (pp. 83–95). Peter Lang.

Ladhari, R., Massa, E., & Skandrani, H. (2020). YouTube vloggers' popularity and influence: The roles of homophily, emotional attachment, and expertise. *Journal of Retailing and Consumer Services, 54*, 1–11. https://doi.org/10.1016/j.jretconser.2019.102027

Laosrirattanachai, P., & Laosrirattanachai, P. (2023). Analysis of vocabulary use and move structures of the World Health Organization Emergencies press conferences on Coronavirus Disease: A corpus-based investigation. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 121–146.

Laosrirattanachai, P. & Laosrirattanachai, P. (2024). The triangle of language use: A corpus-based analysis of hotel responses to reviews. *TESL-EJ, 28*(1), 1–22. https://doi.org/10.55593/ej.28109a3

Laosrirattanachai, P., & Ruangjaroon, S. (2021). Corpus-based Creation of Tourism, Hotel, and Airline Business Word Lists. *LEARN Journal: Language Education and Acquisition Research Network, 14*(1), 50–86.

Lee, J. E., & Watkins, B. (2016). YouTube vloggers' influence on consumer luxury brand perceptions and intentions. *Journal of Business Research, 69*(12), 5753–5760. https://doi.org/10.1016/j.jbusres.2016.04.171

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*, 42–53. https://doi.org/10.1016/j.jeap.2016.01.008

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing, 18*(2), 85–102. https://doi.org/10.1016/j.jslw.2009.02.001

Lu, X., Yoon, J., & Kisselev, O. (2021). Matching phrase-frames to rhetorical moves in social science research article introductions. *English for Specific Purposes, 61*, 63–83. https://doi.org/10.1016/j.esp.2020.10.001

Martin, J. R., Christie, F., & Rothery, J. (1987). Social processes in education: a reply to Sawyer and Watson (and other), In I. Reid (ed.), *The place of genre in learning: Current debates* (pp. 46–57). Deakin University Press.

Marwick, A. (2015). You may know me from YouTube: (Micro)-Celebrity in social media. In: P. David Marshall & S. Redmond (Eds), *A Companion to Celebrity* (pp. 333–350). Wiley.

Maswana, S., Kanamaru, T., & Tajino, A. (2015). Move analysis of research articles across five engineering fields: What they share and what they do not. *Ampersand, 2*, 1–11. https://doi.org/10.1016/j.amper.2014.12.002

Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech, 70*, 157–178.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

Paltridge, B. (2013). Genre and English for specific purposes. In B. Paltridge & S. Starfield (Eds.), *Handbook of English for specific purposes* (pp. 347-366). Wiley-Blackwell.

Panseeta, S., & Watson Todd, R. (2014). A genre analysis of 5-star hotels' responses to negative reviews on TripAdvisor. *rEFLections, 18*, 1–13.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory native like selection and native like fluency. In J. C. Richards, & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–230). Longman.

Phattisiri, K., Laongchinda, K., Prakot, N., & Laosrirattanachai, P. (2023). Scrutinising moves used in job recruitment posters for English for

specific purposes classrooms: A case study of hotel job recruitment. *THAITESOL JOURNAL, 36*(2), 61–81.

Pojanapunya, P., & Watson Todd, R. (2016). Log-likelihood and odds ratio: Keyness statistics for different purposes of keyword analysis. *Corpus Linguistics and Linguistic Theory, 14*(1), 133–167. https://doi.org/10.1515/cllt-2015-0030

Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System, 42*, 220–231. https://doi.org/10.1016/j.system.2013.12.003

Renouf, A., & Sinclair, J. M. (1991). Collocational frameworks in English. In K. Aijmer, & B. Altenberg (Eds.), *English corpus linguistics* (pp. 128–143). Longman.

Rowley-Jolivet, E., & Carter-Thomas, S. (2005). The rhetoric of conference presentation introductions: Context, argument and interaction. *International Journal of Applied Linguistics, 15*(1), 45–70. https://doi.org/10.1111/j.1473-4192.2005.00080.x

Ruiz-Gomez, A., Leaver, T., & Abidin, C. (2021). Playing YouTube: How the Nancy YouTuber doll and app position children as aspiring YouTube influencers. *International Journal of Cultural Studies, 25*(2), 121–140. https://doi.org/10.1177/136787792110632

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1–22). John Benjamins.

Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (Ed.), *Formulaic Sequences* (pp. 127–152). John Benjamins Publishing.

Scott, M. (1996). *Wordsmith Tools 4*. Oxford University Press.

Scott, M. (2001). Mapping key words to problem and solution. In M. Scott & G. Thompson (Eds.), *Patterns of text: In honour of Michael Hoey* (pp. 109-128). Benjamins.

Scott, M., & Tribble, C. (2006). *Textual patterns: Keywords and corpus analysis in language education*. John Benjamins.

Shahriari, H. (2017). Comparing lexical bundles across the introduction, method and results sections of the research article. *Corpora, 12*(1), 1–22.

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

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse.* Routledge.

Snelson, C. (2013). Vlogging about school on YouTube: An exploratory study. *New Media & Society, 17*(3), 321–339. https://doi.org/10.1177/1461444813504271

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes, 12*(3), 214–225. https://doi.org/10.1016/j.jeap.2013.05.002

Stoller, F., & Robinson, M. (2013). Chemistry journal articles: An interdisciplinary approach to move analysis with pedagogical aims. *English for Specific Purposes, 32*(1), 45–57. https://doi.org/10.1016/j.esp.2012.09.001

Stubbs, M. (2007a). An example of frequent English phraseology: Distribution, structures and functions. In R. Facchinetti (Ed.), *Corpus Linguistics 25 years on* (pp. 89–105). Radopi.

Stubbs, M. (2007b). Quantitative data on multi-word sequences in English: The case of word 'world'. In M. Hoey, M. Mahlberg, M. Stubbs & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis* (pp. 163–189). Continuum.

Swales, J. (1981). *Aspects of article introductions.* University of Aston.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings.* Cambridge University Press.

Tardy, C. M. (2011). Genre analysis. In K. Hyland & B. Paltridge (Eds.), *The Continuum companion to discourse analysis* (pp. 54–68). Continuum.

Thumvichit, A., & Gampper, C. (2019). Composing responses to negative hotel reviews: A genre analysis. *Cogent Arts & Humanities, 6*, 1–21. https://doi.org/10.1080/23311983.2019.1629154

Van Herck, R., Decock, S., & Fastrich, B. (2022). A unique blend of interpersonal and transactional strategies in English email responses to customer complaints in a B2C setting: A move analysis. *English for Specific Purposes, 65*, 30–48. https://doi.org/10.1016/j.esp.2021.08.001

Wang, Y. (2017). Lexical bundles in spoken academic ELF genre and disciplinary variation. *International Journal of Corpus Linguistics, 22*(2), 187–211. https://doi.org/10.1075/ijcl.22.2.02wan

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge University Press.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication, 20*(1), 1–28. https://doi.org/10.1016/S0271-5309(99)00015-4

Xu, D., Chen, T., Pearce, J., Mohammadi, Z., & Pearce, P. L. (2021). Reaching audiences through travel vlogs: The perspective of involvement. *Tourism Management, 86*, 1–15. https://doi.org/10.1016/j.tourman.2021.104326

Xu, X., & Lockwood, J. (2021). What's going on in the chat flow? A move analysis of e-commerce customer service webchat exchange. *English for Specific Purposes, 61*, 84–96. https://doi.org/10.1016/j.esp.2020.09.002

Ye, Y. (2021). From abstracts to "60-second science" podcasts: Reformulation of scientific discourse. *Journal of English for Academic Purposes, 53*, 1–13. https://doi.org/10.1016/j.jeap.2021.101025