



FORMULAIC LANGUAGE OF TOURISM IN ENGLISH FOR ACADEMIC PURPOSE (EAP) COURSE BOOK: A CORPUS-DRIVEN APPROACH

Ehsan Namaziandost¹⁾, Meisam Ziafar²⁾, Dwiniasih³⁾

Islamic Azad University of Shahrekord, Iran¹⁾

Islamic Azad University of Ahvaz, Iran²⁾

Swadaya Gunung Jati University, Cirebon, Indonesia³⁾

e.namazi75@yahoo.com¹⁾, meisam-ziafar@iauahvaz.ac.ir²⁾, dwini6644@gmail.com³⁾

Received: March 3, 2020

Accepted: April 7, 2020

Published: May 27, 2020

Abstract

One approach to taking advantage of corpora in language teaching would be adding to a textbook through enriching it through employing corpus-based research. When it comes to using English for Academic Purposes (EAP) materials, the inclusion of corpora in teaching language becomes even more urgent. In the current study, the authors did their best to investigate and describe the presence of formulaic language in an EAP textbooks titled: *English for international tourism: Pre-intermediate students' book* written by Dubicka and O'keeffe (2003) through a case study, and corpus-driven method as a research methodology. Therefore, this study aims to investigate to what extent the EAP course book designed for tourism titled English for international tourism (EIT) is compatible with a corpus-driven formulaic approach. Findings show that this EAP textbook falls fairly short of presenting the necessary formulas as frequently employed in tourism English. Supplementing such materials with corpora and the formulaic they provide may boost the quality of EAP education and practice.

Keywords: Corpora, English for Academic Purposes (EAP), formulaic language, tourism English

INTRODUCTION

According to Richards (1993, cited in Harwood, 2005), the assumption that textbooks are prepared based on joint endeavor of both theoreticians and practitioners is not certain. Harwood (2005) maintains that course books writers have a focus on their intuition or “folk beliefs” (p. 150) in their descriptions of academic norms. What is needed is an understanding based on corpus-based research which proves that English for Academic Purposes (EAP) text book development is a highly heterogeneous process. The fact that textbook development is not aligned with research evidence does not obviate the need for textbooks as sources of systematicity, although such acceptance must be conditional and accompanied by criticism and mediated by corpus-based research. Harwood (2005) believes that “although EAP textbooks may claim to teach a style of writing which holds good across the academy, corpus-based research reveals the naivety of this claim. The differences in

academic discourse practices from discipline to discipline mean that a lack of specificity can mislead and distort” (p. 155).

One must be aware of the type of corpora used in determining the degree to which EAP course books comply with corpora findings. Harwood (2005) criticizes basing EAP textbook development on expert corpora on journal articles on the grounds that (a) student writing and teacher writing are not constant and change through time, (b) undergraduates and postgraduates are writing in genres different from the experts’ journal article genre.

As Harwood (2005) puts it “it is not that the practice of using corpus data to evaluate textbooks is erroneous. Rather, what is important is making sure that the appropriate corpora are selected to enable this evaluation to be methodologically sound” (p. 157). Harwood believes that a fair evaluation of textbooks involves a thorough comparison between the language taught through textbooks and a battery of generically diverse student and expert corpora.

One way to approach the issue would be to take corpora as a source for formulaic expressions (multiword units). Ellis, Simpson-Vlach and Maynard (2008) assert that research done in the field of corpus linguistics proves that natural language considerably draws on recurrent multiword expressions or formulas. What has been advocated as a sound position in word processing is the fact that “human production grammar must store probabilistic relations between words” (p. 377). Wray and Perkins (2000) assert that based on corpus studies it can be claimed that most written or spoken languages seem to be largely consisted of “collocational sets” or “frameworks”. They contend that formulaic is in contrast with productivity in producing novel utterances and in analytically understanding them. Such contention may not be based on sound grounds since many studies have proven the fact that formulaic sequences can be out into further analysis and creative use. Romer (2009) also contends that one of the major findings of modern corpus (computer) linguistics is the fact that languages are highly patterned. Corpus-driven linguistics denies the classification of corpus linguistics as an approach and demands instead that the corpus itself should be the particular authority of the hypotheses which focuses on language. It is thus claimed that the corpus itself embodies a theory of language (Tognini-Bonelli 2001: 84-5). Languages consist of rather fixed or semi-fixed units and through studies in collocation and phraseology one can predict the co-selection of language items. What corpus studies have offered based on large collections of authentic texts from a range of different sources is that lexis and grammar are highly interdependent, thus vocabulary and syntax as two major areas should not be separated from each other as it has been traditionally postulated. The author claims that grammar-lexicon dichotomy may hold water for sentences employed in order to illustrate such dichotomy, but when it comes to real world language data, such dichotomy collapses (Soori, Kafipour, & Soury, 2011).

In order to see how formulaic expressions, corpora, and EAP can be interrelated and employed simultaneously in textbook evaluation one needs to bear in mind that English for academic Purposes (EAP) research has a focus on deciding the “functional patterns and constructions of different academic genres” (Ellis, Simpson-Vlach and Maynard, 2008, p. 377). Within any genre there are special forms of expression and mastering any genre effectively requires mastery of this “phraseology” (p. 377). Ellis, Simpson-Vlach and Maynard (2008) refer to Michigan Corpus of Academic Spoken English (MICASE) which provides access to various academic fields and registers, and makes possible a qualitative analysis of concordances and collocations. Analyzing such academic corpora reveals the presence of a high frequency of common “lexical bundles” (p. 377). The point is that:

The learner has to know these [academic] idioms as a whole; a literal interpretation is no good. And they have to know the common collocations and lexical bundles, too, not

only to increase their reading speed and comprehension, but also to be able to write in a native-like fashion (p. 377).

One approach in taking advantage of corpora in language teaching would be enriching a textbook through supplementing it with a corpus-based research of its components. Harwood (2008) delineates the characteristics of a lexical approach in language teaching and tries to outline the major difficulties of classroom implementation of such an approach and provides some solutions in order to overcome such difficulties. Harwood advocates using corpora as a tool in informing pedagogical materials. He takes a vocabulary/grammar integration approach and questions Chomskyan notion of a native speaker's output which consists of an infinite number of creative utterances, and emphasizes the role of prefabricated items which consists a significant part of a native speaker's spoken and written output. Harwood proposes two major principles in incorporating a lexical approach in language teaching. The first principle is that authentic English should be used together with TEFL course books via corpora, but he also warns against a corpora-bound situation and calls for a corpora-based teaching practice. As the second principle he emphasizes the fact that lexis should be recycled (as a follow-up process) and revisited through practice and repetition since unlike native speakers, L2 learners are not benefiting from spontaneous reviewing after initial presentations. Harwood enumerates the problems of a lexical approach in language teaching. The first point is that corpora data needs to be adjusted before it serves pedagogical goals through teachability/learnability considerations and learners should not be overloaded by unnecessary including of every lexical variant. Harwood warns that learners and teachers may resist to corpus-based materials due to the fact that such materials are untraditional and also because some may have negative reactions towards computers. Another problem is that computers and even computer-based corpora may not be available to be employed by teachers. The second major problem is about teaching and learning real English. (Najafi Sarem, Hamidi, & Mahmoudie, 2013; Namaziandost, Hosseini, & Utomo, 2020). The main limitations in this regard are that learners may not be interested in learning L1 lexis and thus sounding like native speakers, and the fact that real lexis must be prioritized according to learners' needs and not teacher's intuitions. Other points are that non-native teachers are not adequately equipped with lexical knowledge and that there are varieties of real English. Other problems are that course books are not designed to properly recycle lexis systematically, variety and novelty in revisiting lexis practice, face validity for learners and teachers, the perceived operationalizability of a lexical approach by teachers which reduces face validity of such an approach due to limitations like lack of available commercial materials (Masoumpanah & Tahririan, 2013; Namaziandost, Rezvani, & Polemikou, 2020).

Chen and Baker (2010) believe that the importance of corpus-extracted word combinations as building blocks in forming discourse has been widely recognized, although it has not encouraged ELT publishers or practitioners to put more emphasis on computer-retrieved formulaic language in the curriculum and materials. The authors have discovered that the use of lexical bundles between expert academic writing and university student writing (native and non-native alike) shows a significant gap. They suggest that the frequency-driven formulaic expressions in native expert writing can be highly beneficial to learner writers to achieve a more native-like style of academic writing, and as a result calls for inclusion into EFL/ESL curricula.

On the other hand, Koprowski (2005) believes that one of the problems which course book writers face is the huge number of lexical chunks present in the language. Although such problem is pertinent to teaching general English, it seems that such problem is less relevant

when it comes to English for Academic Purposes (EAP) course books. As Koprowski (2005) asserts in order to support language learners the most useful and relevant chunks.

The present study takes corpus-mediated teaching more as a priori research to be done before putting course book into classroom practices, although it does not refute the usefulness of enriching textbook contents through a posteriori research (action research) which dynamically promotes the usefulness of pre-designed course books. This study tries to investigate to what extent the EAP course book designed for tourism titled English for international tourism (EIT) is compatible with a corpus-driven formulaic approach. Attempt has been made to demonstrate to what extent academic formulaic expressions have been presented and taught and whether such formulas have been chosen from the most frequent academic expressions through a corpus-based research. British National Corpus (BNC) as an online corpus has been employed in order to find answers to the research questions just mentioned. Two criteria emphasized by Koprowski (2005) i.e. frequency and range have been taken as two factors to judge the usefulness of each formulaic expression.

METHOD

In the current study the authors did their best to investigate and report the presence of formulaic language in an English for Academic Purposes (EAP) textbooks titled: *English for international tourism: Pre-intermediate students' book* written by Dubicka and O'keeffe (2003) through a case study and a corpus-driven method as a research methodology. Three chapters are taken as a sample in this study. Each chapter is taken as a unit of study to report formulaic language incorporated in the book since Corpus-driven linguistics denies the classification of corpus linguistics as an approach and demands instead that the corpus itself should be the particular authority of the hypotheses which focuses on language. It is thus claimed that the corpus itself embodies a theory of language (Tognini-Bonelli 2001: 84-5). Corpus analysis analyzed model of language choices in abundant collection of texts by preference that have been composed under whatever circumstances would be typical of that genre. Ideally, the collection is machine determined for regularity and reliability, and any decisions drawn from the resulting data are grounded in theoretically comprehensible interpretive judgments (Biber, Conrad, & Reppen, 1998, pp. 4-5). In Lee's (2008) taxonomy of methods to corpus analysis, this study deceits between the essentially qualitative and intuitive corpus-informed approaches and the essentially quantitative corpus-induced studies used in, say, machine learning (p. 88).

RESULTS AND DISCUSSIONS

In general, EIT seems to be wanting in regard to the presence of highly frequent formulaic expressions at first glance. Scant numbers of formulaic expressions have been presented for each chapter and other related highly frequent formulaic expressions relevant to each chapter have also been given which could highly promote the quality of this English for Academic Purposes (EAP) textbook.

Table 1. Formulaic expression presented in chapter 1

Formulaic expressions presented in chapter 1	Frequency in BNC							Total per million
	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	

What is the name of the guest in room 212?	4	7	5	8	3	4	6	37	.43
Can you spell that name for me?	1	1	-	-	-	-	1	18	.57
Where is he/she from?	6								
What is his passport name?	1	7	2	2	3	3	23	52	.51
	2								

Table 2. Possible formulaic expression presented in chapter 1

Possible formulaic expressions to be presented in chapter 1	Frequency in BNC								Total per million
	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	Total	
How may I help you?	-	6	1	1	-	-	2	10	.1
How may I help you?	-	4	-	-	-	-	-	4	.25
Could you sign your name here?	1	1	1	-	-	-	-	3	.1
Have you got a single room for tonight?	2	30	2	7	17	-	307	365	3.13
I'd like to stay in a double room.	2	12	5	5	1	2	19	46	.46
Sorry, we're fully booked for tonight.	203	203	82	363	50	27	323	1251	14.19
I'd like to check out, please									
Do you have any vacancies									
What's the room rate for a single room?		1		1		1	7	10	.14
How long will you be staying?	-	2	-	-	-	-	-	2	.13

As can be seen in table 1 and 2 that little attention has been paid to the role of formulaic expressions prevalent in hotel jobs. The chunks presented also are not the most frequent ones in hotel conversations. It seems that EIT puts tourism English learners at a disadvantageous position in regard to high frequent expressions needed in a hotel conversation. The following examples have been provided as instances of the use of formulaic expressions in hotel conversations (taken from <http://2ndnature-online-eikaiwa.com/index.htm>). What EIT lacks is exactly the same conversations which include high frequent relevant formulaic expressions in hotel conversations.

Examples:

Dialog 1: (Booking a Room over the Phone)

- Clerk: Royal Inn. How may I help you?
- Jim : Hello, **I'd like to reserve a single room** for next week.
- Clerk: Certainly, sir. When will you be arriving?
- Jim : Well, I'll be arriving on June 15.
- Clerk: And **how many nights will you be staying?**
- Jim : Two nights. I'll be leaving on June 17.
- Clerk: OK. I'll check to see if there are any vacancies. Please hold on.

- Jim : Thanks.
- Clerk: Hello, sir. There's no problem. There are rooms available on June 15.
- Jim : Great! By the way, **what is the room rate?**
- Clerk: US\$75 per night for a single room.
- Jim : OK. Does it come with a bath?
- Clerk: Yes, all of our rooms have bath or shower.
- Jim : OK. Can I reserve a room then?
- Clerk: Sure. May I have your name and telephone number, please?

Dialog 2: (Asking For a Room at a Hotel 1)

- Clerk: Good evening. May I help you?
- Jim : **I need a single room**, please.
- Clerk: **Do you have a reservation?**
- Jim : No, I'm afraid I don't.
- Clerk: I'm sorry, **we are fully booked**.
- Jim : Oh! Do you know where I can find another hotel in this area?
- Clerk: There is a Holiday Inn across the street.
- Jim : OK. I'll try there. Thank you.

Dialog 3: (Asking For a Room at a Hotel 2)

- Clerk: Good evening. May I help you?
- Jim : **I need a single room**, please.
- Clerk: **Do you have a reservation?**
- Jim : No, I'm afraid I don't.
- Clerk: **How long will you be staying with us?**
- Jim : Just one night.
- Clerk: Would you prefer a non-smoking room?
- Jim : Yes, please.
- Clerk: OK. We have a single non-smoking room on the 10th floor.
- Jim : Great! **What's the room rate?**
- Clerk: US\$65 plus tax. Breakfast included.
- Jim : Excellent!
- Clerk: **May I ask you to fill out this form for me, please?**
- Jim : Sure

Table 3. Formulaic Expressions Presented in Chapter 2

Formulaic expressions presented in chapter 2	Frequency in BNC								Total per million
	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	Total	
Are you taking any holidays in July?	2	2	1	-	1	1	5	12	.08
Could you put me through to Gabriella, please?	2	6	-	1	1	-	1	11	.15
I'm calling about the holiday on page 84 of your brochure.	2	1	-	1	-	-	-	4	.12

Table 4. Possible Formulaic Expressions Presented in Chapter 2

Possible formulaic expressions to be presented in chapter 2	Frequency in BNC								Total per million
	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	Total	
Can you keep an eye on my bag?	40	153	30	23	31	8	54	339	3.56
Have you ever been to Iran?	10	7	-	1	-	1	1	20	.33
Is this seat taken?									
East or west, home is the best									
I'd like to reserve two seats to New York.	503	180	197	169	127	59	363	1598	19.14
Will that be one way or round trip?	4	10	21	28	7	4	41	115	1.32

Table 5. Formulaic Expressions Presented in the Other Chapters

Formulaic expressions presented in the other chapters	Frequency in BNC								Total per million
	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	Total	
Can I have some juice?	72	4	1	2	-	1	1	81	1.32
Would you like some fruit?	26	36	-	1	-	-	-	63	1.65
Are there any tables free?									
He always puts a lot of salt in his food	10	3	2	1	2	-	-	18	.33
There's been a mistake	4	49	3	8	5	3	7	79	.78
I'll be with you in a minute	4	9	-	-	-	-	-	13	.57
I'll ask the chef to head it up									
We apologize for any inconvenience	1	4	3	9	2	-	13	32	.39
It's a good idea to book hotels in advance.	-	1	5	3	-	-	14	23	.43
Make sure you drink lots of fluids when it is hot	154	97	148	35	54	12	206	706	8.46
It's wise to avoid uncooked food	-	2	10	4	11	12	41	80	.88
I recommend you use your own card	4	2	3	1	-	-	1	11	.21
It gets very humid in summer	382	155	110	73	95	30	98	943	11.8

Table 6. Possible Formulaic Expression Presented in the Other Chapters

Possible formulaic expressions to be presented in the other chapter	Frequency in BNC								Total per million
	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	Total	

Can I get you something to drink?	5	37	-	-	-	-	1	43	.96
I'll come back to take your orders in a few minutes.	28	73	4	39	4	4	24	176	1.9
If you look up you will notice.....	11	4	14	-	22	3	21	78	1
Do you have anything to declare?	-	1	-	1	1	1	1	5	.06
Would you like any help with your luggage?	611	598	62	25	23	12	102	1433	16.7
Do you mind if I sit here?	80	106	3	-	4	-	11	204	1.6
What's the purpose of your visit?	1	4	-	-	2	1	2	10	.12
Could I have a wake-up call at seven o'clock?	48	84	6	5	1	5	10	159	1.75
Would it be possible to have a late check-out?	9	13	6	4	5	3	18	58	.61
We need a few more minutes to decide.									
Is this seat free?									
What time do I need to check out?									

Results show that there are highly frequent and practical expressions which have been ignored by EIT writers as shown in table 4 and 6. The total per million section of each table represents both the frequency and range of each expression and is a further proof for the weakness of EIT with regard to taking formulas as pragmatic and practical building blocks of a tourism discourse. There are other famous and useful chunks which could not be extracted from BNC and have been presented at the end of each table.

Results show that English for International Tourism (EIT) is highly ignorant of the facilitative role of formulaic expressions in teaching English for tourism. It seems that the sparse use of such formulas has also been the result of intuition and folk belief rather than a corpus-based research. Where formulas have been used, there has been no explicit effort to teach them systematically and there has been no concern to promote learners' awareness about the presence of such formulas. It confirming to Dryer (2013) that the corpus is even more selective in the places where it acknowledges readers, scorers, or raters. It is not astonishing to examine them frequently in the characteristic of "audience/rhetorical awareness," but they are extremely demonstrated when offering affirmative assessments of style and in their robust reactions to the appearance of error. In chapter four, EIT could present formulas useful for tour guides. In chapter seven the writers do a good job in presenting the necessary formulaic expressions for giving advice and recommendation in the section titled "Language Focus", although such expression have not been contextualized. Within each chapter there are highly frequent chunks which have not been included and ignored. Formulas have been sparsely scattered throughout EIT without any attempt to contextualize them within conversations and natural discourse. If taken into consideration, formulaic expressions prevalent in the field of tourism have the potential to yield enormous numbers of ready-made formulas which can be taken as sources for deeper manipulation and contextualization.

Corpora and concordances prove the formulaic nature of language units and thus can be employed as research tools and materials in deciding on the content and sequence of EAP textbooks. Where formulas have been presented in EIT they have been of a general English rather than EAP nature (e.g. what is the name of the guest....?, what is his passport name?

etc.). Specialized formulas for tourism in many fields such as hotel management, ecotourism, restaurant, etc., have been neglected and this textbook is in need of a fundamental revision to be representative of English terms employed within a tourism field of study. Native-like style of academic writing, which Chen and Baker (2010) believe happens through frequency-driven formulaic expressions, is never achieved without developing materials prepared based on computer-retrieved formulaic language in the curriculum. All in all, it seems that EIT more resembles a general English textbook with contents about tourism and not an English for Academic Purposes (EAP) textbook with a genuine concern for teaching professional English for tourism.

CONCLUSION

A corpus-driven research on the role of formulas within EAP textbooks is advantageous in that it makes teachers and learners cognizant of the richness of such textbooks with regard to the presence of highly frequent and necessary expressions within any field of study. EIT has shown to be wanting in being based on sound corpus-based research and reduces an EAP material to a general English textbook embellished with topics and contents relevant to international tourism. A true EAP textbook needs to be much more representative and comprehensive in presenting specialized expressions and phraseology in tourism. All in all, it may be claimed that successful EAP teaching involves adapting classroom materials in order to provide further information about the formulaic as present in each specific field. Corpora may serve as a highly beneficial tool in naturalizing EAP formulaic language and practice.

REFERENCES

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, England: Cambridge University Press.
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30-49.
- Dryer, D. B. (2013). Scaling Writing Ability. *Written Communication*, 30(1), 3–35. Doi: 10.1177/0741088312466992. sagepub.com/journalsPermissions.nav
- Dubicka, I., O'keeffe, M. (2003). *English for international tourism: Pre-intermediate students' book*. Essex, England: Longman.
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42(3), 375-396. doi:10.1515/CLLT.2009.003
- Harwood, N. (2005). What do we want EAP teaching materials for? *Journal of English for Academic Purposes*, 4, 149-161. doi:10.1016/j.jeap.2004.07.008
- Harwood, N. (2008). Taking a lexical approach to teaching: Principles and problems. *International Journal of Applied Linguistics*, 12(2), 139-155. doi:10.1111/1473-4192.00028
- Koprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary course books. *ELT Journal*, 59(4), 322-332. doi:10.1093/elt/cci06
- Lee, D. Y. W. (2008). Corpora and discourse analysis: New ways of doing old things. In V. K. Bhatia, J. Flowerdew & R. H. Jones (Eds.), *Advances in discourse studies* (pp. 86-99). Oxford, England: Routledge.

- Masoumpanah, Z., & Tahririan, M. H. (2013). Target Situation Needs Analysis of Hotel Receptionists. *English for Specific Purposes World*, 1-19. Available at: http://www.esp-world.info/Articles_40/Masoumpanah.pdf
- Najafi Sarem, S., Hamidi, H., & Mahmoudie, R. (2013). A critical look at textbook evaluation: a case study of evaluating an ESP course-book: English for international tourism. *International Research Journal of Applied and Basic Sciences*, 372380. Available at: http://www.irjabs.com/files_site/paperlist/r_684_130209103353.pdf.
- Namaziandost, E., Hosseini, E., Utomo, D.W. (2020). A comparative effect of high involvement load versus lack of involvement load on vocabulary learning among Iranian sophomore EFL learners. *Cogent Arts and Humanities*, 7(1).
- Namaziandost, E., Rezvani, E., Polemikou, A. (2020). The impacts of visual input enhancement, semantic input enhancement, and input flooding on L2 vocabulary among Iranian intermediate EFL learners. *Cogent Education*, 7(1).
- Romer, U. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*, 7, 141-163.
- Soori, A., Kafipour, R., & Soury, M. (2011). EFL textbook evaluation and graphic representation. *European Journal of Social Sciences*, 26 (3), 481-493.
- Tognini-Bonelli, T. (2001). *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, 20, 1-28.