A Comparative Investigation into Formulaic Expressions in EAP Textbooks: The Case of Social and Hard Sciences

Parastoo Alizadeh Oghyanous*

Department of Foreign Languages, Faculty of Letter and Humanities, Kharazmi University, Tehran, Iran

1. Introduction

Discourse analysis and text analysis reveal valuable information about the nature of the connected speech going beyond word and sentence level. It is now accepted that knowing how to read is not solely dependent on vocabulary knowledge and draws on various skills such as knowledge about registers and knowing how words and phrases are used together in particular text types. The more is known about the characteristics of

* Corresponding author
various text types the more informed decisions can be made about teaching reading and
writing of various text types. In this regard genre-based approaches to reading and
writing emphasize the requirements of discourse community (Swales 1990) and
disciplines (Kim & Kim, 2012) from academic texts.

Another research line that has received less attention in this regard is the use of
formulaic expressions in written texts and how such use of formulaic expression varies
across disciplines (Wolter & Gyllstad, 2013). Despite the fact that the literature on
formulaic expressions in written texts is not thick and extensive, more recently there is a
rise of research on formulaic expressions (e.g., Koprowski, 2005; Biber, 2006; Kim &
Kim, 2012; Wolter & Gyllstad, 2013) in academic context either in spoken or written
format that further points to the importance of formulaic expression in comprehension.
This line of research needs to be continued to non-native English learning context of
research to further enrich the knowledge about the use of formulaic expression it written
academic texts in social and hard sciences. It is noteworthy to mention that there are two
main branches of sciences, namely, natural and social. Natural sciences such as physics,
biology, etc., are considered “hard”, while the social sciences such as sociology,
economics, etc., are described as “social” or “soft” (Hedges, 1987).

There is no question that the number of graduate students is rising either in Iran
or internationally which means that more up-to-date information should be used for
helping the learners better read and write in English as an international and academic
language. This issue more elegantly highlights the importance of new research on texts
characteristic across disciplines. Such research should target both the original books
published for native and non-native speakers and also EAP books with the aim of
improving the reading competence of non-native speakers. After all, disciplines are
internationally recognized and this means that there are many students in non-English
countries interested in reading original academic books although they have limited
knowledge of English. Therefore, the quality of EAP books in terms of authenticity,
genre, etc. is important, because they should prepare the learners for reading original
academic books. In this matter, less attention has been paid to formulaic expressions in
academic books.

As academic textbooks are essentially regarded as the most important source of
language input for EAP students, investigating formulaic sequences in EAP textbooks
would obviously provide insights into the educational principles of formulaic
sequences, and the definite necessities of various registers (Miao, 2014). According to Swales (1990), every EAP genre has its own word choice, and learning to be influential in the genre basically requires learning the pertinent formulas. Genres are hugely important in the practical business of writing. They are, after all, “resources for getting things done, and we all have a repertoire of resources we can call on for recurring situations, from shopping lists to job applications” (Hyland, 2003, p. 19). The current study focuses on formulaic expression in EAP textbook genre in different disciplines which partly addresses the need for more research on text characteristics. The current study was justified and significant on certain grounds like its contribution to the theoretical framework on formulaic expression and written academic texts and its pedagogical and theoretical implications. This study may allow a review of the role of text analysis including analysis of formulaic expression in English for academic purpose. The review would tell us the current status of the role of formulaic expressions in academic texts. In addition, the actual analysis done in the current study about the frequency of formulaic expression would further enlighten us about the role of formulaic expression in academic texts and its relation to previous studies.

In addition to the theoretical contributions of the current study, the results of the study would give awareness about the use of formulaic expression in academic texts to the language teachers and material developers. For instance, the results of the study can tell the teachers how the formulaic expressions are used in the academic texts and accordingly what instructional strategies they may use to better help the leaners understand the role of formulaic expression in comprehension of the academic texts. Similarly, material developers may get awareness about the use of formulaic expression and accordingly make more informed decisions regarding the incorporation of activities and passages that take account of formulaic expressions.

2. Literature Review

2.1 Definition of formulaic expression

There is not a complete consensus on the definition of formulaic expressions despite the fact that formulaic expressions have been given an important role in second language acquisition. There are various terms referring to the concept of formulaic expressions including lexical bundles, formulaic sequences, lexical phrases, and
phrasal verbs (Wray, 2002). Based on Wray’s (2002) definition, a formulaic expression is a:

“sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (p. 9).

Formulaic expressions can be different in terms of grammar, fixedness, semantic transparency, pragmatic function, phonology, frequency, and holistic versus individualistic nature in terms of storing in memory and use. Schmitt and Carter (2004) point out that formulaic expressions can be long, short, syntactically fixed, or less fixed. In this regard, Read and Nation (2004) proposed a more efficient classification of formulaic expressions for more accurate and reliable identification of formulaic expressions. In this regard, two common ways of dealing with classification of formulaic expressions has been classification based on phraseology and frequency based classification.

In phraseology based classification, formulaic expression is seen as continuum extending between free sequence of words to formulaic sequence of words (Conzett, 2000; Hill, 2000). Based on such classification, formulaic expressions can be free combinations, collocations, and idioms. These types of formulaic expressions can vary in terms of fixedness of syntactic and transparency of meaning or literalness. For instance, kick the bucket is more fixed and less literal than the formula expression of do homework.

**2.2 Formulaic expression in language acquisition**

In second language acquisition literature, formulaic expression including collocations, idioms, proverbs, etc. are considered linguistic devises helping the language learners to sound more fluent and accurate. That is the reason that Dechert (1983) called the formulaic expression islands of reliability to point to the efficacy of formulaic expression in helping the language learners to sound more native like. Pawley and Syder (1983) maintain that the use of formulaic expression compensates for lack of language proficiency and help the learners to perform more native like.

In the field of psycholinguistic, it has been tested whether formulaic expressions are learned and used holistically or in a bottom up manner. This issue has been
examined both in first and second language and some report that formulaic expressions are processed faster and more accurately than non-formulaic ones (e.g., Jiang & Nekrasova, 2007). Although much has been discussed regarding the advantages of formulaic expressions, there are still areas less known to the researchers. For instance, collocations as a subtype of formulaic sequences have not been adequately explored in terms of processing although there seem lots of studies on collocations (e.g., Altenberg & Granger, 2001; Webb & Kagimoto, 2009). In addition, two features of collocations that might affect their processing have been less investigated, i.e. the issue of the strength of collocations and the congruity of the collocations. Congruence refers to the correspondence between collocation in L2 and its meaning in L1. Another issue is that less is known about the speed of producing formulaic expressions in comparison with the speed of their comprehension. There are some studies on the use of formulaic expression in oral tasks (e.g. Arnon & Priva, 2013; Ellis & Simpson-Vlach, 2009; Janssen & Barber, 2012) but majority of the studies have been on receptive tasks like self-paced reading tasks and phrasal decision tasks (e.g., Jiang & Nekrasova, 2007; Tremblay, Derwing, Libber, & Westbury, 2011).

Studies on frequency of formulaic expressions have yielded different results with some pointing to the positive effect of frequency of formulaic expression (Kim & Kim, 2012; Wolter & Gyllstad, 2013). Jiang and Nekrasova (2007) found that formulaic expressions are processed faster by Korean ESL speakers. However, Valsecchi et al. (2013) did not report such benefits of the frequency of formulaic expressions.

### 2.3 EAP materials

The prominence of EAP materials has long been emphasized in EAP literature. It is believed that materials and textbooks have a decisive role in different learning situations (Dudley-Evans & John, 1998). ESP/EAP materials are supposed to have six different objectives: (a) incentive for learning, (b) establishing learning and teaching process, (c) exemplifying a view of the nature of language, (d) manifesting the essence of the learning tasks and activities, (e) developing the basis of educator training, and finally (f) offering models of proper language use (Hutchinson & Waters, 1987). However, regardless of the prominence and steady growth of ESP/EAP materials and textbooks and contrary to its worldwide position, regrettably they are margined in several ways.
Currently, EAP is an indisputable part of national curriculum in all fields of education in Iran. Moreover, university students of non-English fields have to pass different EAP courses in order to “bridge the gap between the learners’ general English reading competence and their ability to read discipline-based courses” (Atai, 2002, p. 269). Mostafaei Alaei and Ershadi (2016) argued that the material development for EAP and ESP courses in Iran was basically increased by the establishment of SAMT (Iranian Centre for Materials Development in Humanities) in 1984. According to Atai and Shoja (2011), in the 1980s that Ministry of Science, Research and Technology accepted the responsibility for ESP instruction at all Iranian universities, it could be “published eight EAP textbooks for students of sciences, humanities, sociology, engineering, medicine, and agriculture” (p. 306).

However, it was believed that the EAP textbooks developed and published by SAMT “were inadequate, ineffective, and in some cases with obvious typographical and content-related errors below the dignity of university education” (Farhady, 2005, p. 5). Several investigations have been done on EAP textbooks developed and published by SAMT usually outlining critical results. As aptly pointed out by Soodmand Afshar and Movassagh (2016, p. 144):

“The one-size-fits-all approach taken to EAP in Iran in the form of inflexible, tailor-made, already-outlined and unappealing materials of SAMT followed by monotonous translation-based instruction of EAP teachers- a vicious circle still compounded by traditional rather unimodal summative assessment procedures (i.e. use of translation and true-false and multiple-choice reading comprehension tests) lead us nowhere.”

According to Mostafaei Alaei and Ershadi (2016), mostly EAP/ESP materials are developed by the EFL teachers and in some cases subject specialists develop some sources. Mohebi, Hessamy and Karimkhanlouie (2013) scrutinized the views of EAP students and EFL teachers concerning their EAP textbook. The findings of their study indicated that lack of proper materials to the students’ requirements, overcrowded classes, lack of clear goals in the EAP environment, and inadequately of time were the main problems teachers and EAP learners confronted. In conclusion, although there were positive
perceptions of the textbook they use, most of the instructors did not think that it was enough for an EAP course.

2.4 Formulaic expression in EAP textbooks

Most of the studies have pointed to the benefits of formulaic expressions in the comprehension of L2 learners. Based on the claim by Biber et al. (1999) multi-word sequences comprise a large part of academic texts which indicates that examination of lexical sequences and the way they affect comprehension is significant in academic text research. Thanks to the advent of computerized analysis of text this area of research has been on the rise and many researchers like Biber, Conrad and Cortes (2004), Biber and Barbieri (2007), Cortes (2004), and Hyland (2008) have carried out text analysis using computer. In this regard, different studies have pointed to the prevalent of formulaic expressions in written mode of language use in discipline such as management and linguistic (e.g. Adel & Erman, 2012; Biber and Barbieri, 2007; Cortes, 2004; Hyland, 2008).

Koprowski (2005) found that formulaic expressions are practical in the textbooks and proposed that the use of formulaic sequences in textbooks. Simpson-Vlach and Ellis (2010) proposed the use of chunk of words based on the frequency of chunk of academic words. In the study by Biber (2006), difference between oral and written speech was examined in terms of category and functions of formulaic expression in college textbooks. Gray and Biber (2013) have pointed to the use of corpus and corpus technology in studying formulaic expressions in written and oral speech and how formulaic expression are used in different registers. Schmitt (2012) also expressed the usefulness of high frequency formulaic sequences in the textbook compilations. According to Yang (2012), textbook compilation and learning must highlight the importance of formulaic sequences that basically pave the way for native-like choices and also idiomatic expression. As in spoken discourse, the use of common expressions which act like prefabricated expressions can be an asset for reading fluency and comprehension. Taking this issue into account in EAP textbook preparation can lead to development of more authentic materials and consequently contribute to the academic reading competency of university students.

Several ESP/EAP textbooks evaluations have been conducted in recent decades. For instance, Farhady (2005) discussing different ESP parameters, such as materials,
context, etc., concluded that there is a necessity to carry out systematic need analyses, train competent educators, change the design of the materials, and finally improve the testing processes and procedures. Moreover, Tajeddin (2005) evaluating some ESP textbooks concerning their linguistic input, and output, argued that ESP textbooks which are developed and published by SAMT not merely lack one unified approach and lesson plan but also do not follow any definite goal concerning the materials provided for improving the EAP students’ comprehension or translation abilities. Tajeddin (2005) concluded that the future planning necessitates an essential and basic revision of the EAP materials and textbooks developed and published by SAMT.

However, to the best of the researcher's knowledge, despite the significance of formulaic expressions in EAP textbooks, it seems that little research, if any, has investigated the functions and frequencies of formulaic expressions in EAP textbooks in the context of Iran. Thus, to shed more light on this issue, the present study focused on the formulaic expressions in EAP textbooks in social sciences (i.e., English for the Students of Psychology, English for the Students of Sociology, and English for the Students of Political Science) and also hard sciences (i.e., English for the students of Engineering, English for the students of Biology, and English for the Students of Mechanization and Mechanics of Agricultural Machinery) compiled and published by SAMT. The focus on formulaic expressions was in terms of their functions and frequencies and the way they were used differently across hard and soft disciplines. The textbooks in social sciences were from psychology, sociology, and political sciences and in hard sciences from engineering, biology and mechanics. It should be noted that the texts were written by non-Iranian writers but Iranian writers had selected and modified them to be used for educational purposes. It seems that this localized line of research on formulaic expressions in academic texts have received less attention by scholars. Thus, the following research questions were formulated by the researcher:

1. What are the frequencies of formulaic expressions in EAP books in hard sciences?
2. What are the frequencies of formulaic expressions in EAP books in social sciences?
3. Is there any difference in frequency of formulaic expressions between EAP books in social sciences and EAP books in hard sciences?
3. Method

3.1 Corpus

In order to find out about the frequency of formulaic expression, first a corpus of texts was compiled. The corpus included EAP textbooks compiled and published by an Iranian publication, namely, SAMT, as a well-known publication of academic textbooks in Iran which makes it a good representative of Iranian EAP publications. The EAP textbooks were chosen randomly from the list of EAP textbooks in hard and soft sciences from SAMT publications. Roughly speaking, hard sciences are various disciplines in natural sciences like biology, chemistry, and physics while soft sciences are those disciplines associated with social sciences like economics, psychology, sociology (Palema, 2009). The EAP textbooks in social sciences were English for the Students of Psychology, English for the Students of Sociology, and English for the Students of Political Science. As for the EAP textbooks in hard sciences, the textbooks included English for the Students of Engineering, English for the Students of Biology, and English for the Students of Mechanization and Mechanics of Agricultural Machinery. The total size of the corpse was 30000 words (15000 for hard sciences and 15000 words for soft sciences). The texts were fed into concordance software for computerized analysis of the texts.

3.2 Analysis Procedure

The texts were analyzed by the concordances of four-word cluster formulaic sequences by AntConc 3.2.4. Four-word range for identification of formulaic expressions is the recommended range for bundle of words in academic texts (e.g. Adel & Erman, 2012; Biber, Conrad, & Cortes, 2004; Hyland, 2008). Hyland (2008) has specifically pointed out that four-word range can provide clearer picture of formulaic sequences in terms of their functions.

AntConc3.2.4 is a software developed by Anthony (2011) at the Center for English Language Education in Science and Engineering, School of Science and Engineering, Waseda University. AntConc3.2.4 allows for the both raw and normalized frequency analysis which is the standard procedure for identification of lexical bundles like formulaic expressions. Raw frequency is concerned with the frequency count of lexical bundles in the entire corpus while normalized frequency deals with frequency count of lexical bundles per limited number words within a text. In the current study,
initially the frequency count for identification of formulaic expressions was normalized by setting the cut-off point as occurrence of 10 four-word bundles per hundred words in line with previous studies on identification of frequency counts (e.g. Kashiha & Heng, 2014). However, after setting the AntConc to identify the lexical clusters within the normalized range, it was found that standard normalized frequency and bundles with range of 2 words do not lead to proper identification of lexical bundles. Therefore, the setting was compromised and normal frequency was set to a minimum of 10 occurrences per 400 words. Furthermore, the range of lexical bundles was set to 3 and 4 word bundles.

In addition, the function of formulaic expressions was also taken into account by using the classification framework proposed by Biber, Conrad and Cortes (2004). After finding the formulaic expressions, their functions were identified using the classification framework work by Biber, Conrad and Cortes (2004) and their frequencies were counted by the help of AntConc3.2.4 and made ready for analysis using independent samples t-test. It should be noted a Chi Square may not suitable for this study as Chi Square requires only categorical variables while in the current study there was one categorical variable (type of science; Iranian EAP books vs international EAP books) and one interval variable (the frequency of formulaic expression). Since the frequency of lexical bundles in either hard or soft disciplines were accumulated to from a total score showing the total use of formulaic expressions they could no longer be considered categorical data. Furthermore, since the accumulative frequency scores were not normally distributed non-parametric test of Mann Whitney U test was chosen for statistical analysis.

4. Results

The identification and sorting process of the formulaic structure was conducted in two phases. The actual analysis was started after feeding the texts into the AntConc 3.2.4 and setting the parameter for the the N-gram function of the software. The software was set to find the lexical bundles within the range of four-word bundles are recommended by literature. After making the initial preparations, the software was run the list of lexical bundles was produced. This was the first phase of the study. In the second phase, the list of identified lexical bundle was surveyed manually to remove any meaningless bundle or lexical bundles that were proper nouns like names of journals,
places, institutes, etc. In addition, field specific clusters such as chemical reaction engineering or in chemical engineering were also removed because, only common structures in hard sciences subjects were the target structures. Otherwise, a long list of subject specific structures that are commonly used in each subject had be listed. Figure 1 shows the raw output of the AntConc run on the hard science corpus.

Figure 1 The raw output of the AntConc run on the hard science corpus

As seen in Figure 1 some lexical clusters such as a a a and to kw m were structures that were erroneously identified by the software. Table 1 shows the list of lexical bundles and their frequencies in EAP books in hard sciences after surveying and refining the output of the AntConc 3.2.4.

Table 1 list of lexical bundles and their frequencies in EAP books in hard sciences

<table>
<thead>
<tr>
<th>Frequency</th>
<th>lexical bundles</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>283</td>
<td>shown in figure</td>
<td>textual</td>
</tr>
<tr>
<td>244</td>
<td>is given by</td>
<td>ideational</td>
</tr>
<tr>
<td>240</td>
<td>the number of</td>
<td>ideational</td>
</tr>
<tr>
<td>Line</td>
<td>Sentence</td>
<td>Function</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>205</td>
<td>a function of ideational</td>
<td></td>
</tr>
<tr>
<td>203</td>
<td>can be used ideational</td>
<td></td>
</tr>
<tr>
<td>181</td>
<td>in terms of ideational</td>
<td></td>
</tr>
<tr>
<td>179</td>
<td>the rate of ideational</td>
<td></td>
</tr>
<tr>
<td>152</td>
<td>as well as textual</td>
<td></td>
</tr>
<tr>
<td>139</td>
<td>the use of ideational</td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>be used to ideational</td>
<td></td>
</tr>
<tr>
<td>134</td>
<td>the value of ideational</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>based on the ideational</td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>area of application ideational</td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>as a function ideational</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>as a function of ideational</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>as shown in textual</td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>depends on the ideational</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>for a given ideational</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>the effect of ideational</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>of the liquid ideational</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2* Comparison between functions of formulaic expression in hard sciences in terms of frequency occurrence.
As seen in Figure 2, ideational function of formulaic expression has the highest frequency occurrence among other function of formulaic expression in hard science EAP books.

The same analysis of formulaic expression presented above was repeated on the social sciences’ texts. In other words, first the raw output of AntConc was surveyed and then refined. Figure 3 shows an extract from raw output of AntConc.

![Figure 3](image-url)  
*Figure 3* The raw output of the AntConc run on the social science corpus

After surveying the output of AntConc, it was found that certain expressions were proper nouns, and topic specific expression such as the united states and theory of personality respectively. Some other expressions were meaningless such he or she which were removed from the list of formulaic expression in social sciences. The final list of formulaic expressions in social sciences is found in Table 2.

*Table 2* list of lexical bundles and their frequencies in EAP books in social sciences

<table>
<thead>
<tr>
<th>Frequency</th>
<th>lexical bundles</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>157</td>
<td>in terms of</td>
<td>ideational</td>
</tr>
<tr>
<td>114</td>
<td>the experiential system</td>
<td>ideational</td>
</tr>
<tr>
<td>111</td>
<td>experimental social psychologies</td>
<td>ideational</td>
</tr>
<tr>
<td>111</td>
<td>the psychology of</td>
<td>ideational</td>
</tr>
<tr>
<td>107</td>
<td>the role of</td>
<td>ideational</td>
</tr>
<tr>
<td>101</td>
<td>the study of</td>
<td>ideational</td>
</tr>
</tbody>
</table>
Based on the frequency occurrence of formulaic expressions in social sciences, it was found that ideational formulaic expressions were the most frequent ones. Textual expressions had much lower frequency than ideational expressions and interpersonal expressions were not found in the social science corpus.

Up to this point, the use of formulaic expression had similar distributions in social and hard sciences. To be more specific, in both social and hard sciences, ideational expressions had the highest frequency and interpersonal expressions were
non-existent. This comparison between hard and social sciences reveals the prevalence of formulaic expression in both hard and social sciences. In order to make a robust comparison between hard and social sciences in terms of frequency occurrence of formulaic expressions, total frequency count in each discipline was computed that then was analyzed using the statistical test of Mann Whitney. Table 3 shows the descriptive statistics of the frequency counts and Table 4 shows the result of Mann Whitney.

Table 3 descriptive statistics of the frequency counts

<table>
<thead>
<tr>
<th>discipline</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
<th>Kolmogorov-Smirnov Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>hard</td>
<td>20</td>
<td>157.25</td>
<td>52.70461</td>
<td>.235</td>
<td>20</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>social</td>
<td>16</td>
<td>96.000</td>
<td>21.11240</td>
<td>.172</td>
<td>16</td>
<td>.200</td>
</tr>
</tbody>
</table>

Based on descriptive statistics, the mean of total frequency of formulaic expressions was 157.25 (SD=52.70) in hard sciences and 96.00 (SD=21.11) in social sciences. Apparently, there was much difference between hard and social sciences in terms of frequency of formulaic expression. Since the frequency scores in hard sciences was not normally distributed (D=0.23, P≤0.05), Mann Whitney U test was used to compare the frequency count of formulaic expressions between hard and social sciences.

Table 4 Results of Mann Whitney

<table>
<thead>
<tr>
<th>discipline</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>hard</td>
<td>20</td>
<td>25.28</td>
<td>505.50</td>
<td>24.500</td>
</tr>
<tr>
<td></td>
<td>social</td>
<td>16</td>
<td>10.03</td>
<td>160.50</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to the results of Mann Whitney U test, there was a significant difference between hard and social sciences in terms of frequency of formulaic expressions (U=24.50, P=0.00). Based on this result and by taking account of mean of
frequency counts of formulaic expression, it can be claimed that hard science EAP books included significantly higher use of formulaic expressions than social science EAP books.

5. Discussion

The present study aimed at investigating the frequency occurrence of formulaic expressions in EAP books in hard and social sciences. The corpus of the study was a selection of EAP books in hard and social sciences published by SAMT publication. The identification and frequency counts of the formulaic expressions were achieved by the help of AntConc software. The software was set to identify and count the frequency of formulaic expression within the range of 3 to 4 words. Based on the output of the software, it was found that in both hard and social science EAP books the most frequently used expression were ideational followed by textual formulaic expressions. Furthermore, it was revealed that there was a significant difference between the frequency occurrence of formulaic expressions between hard and social science EAP books. To be more specific, formulaic expression had much higher frequency in hard science EAP books.

The results of the study suggest that formulaic expression is an essential part of EAP texts. This essential role is implied by the prevalence of formulaic expressions and their different structures and functions. Previous studies have also pointed to the prevalence of formulaic expression in written texts of certain disciplines like management and linguistics (e.g. Adel & Erman, 2012; Biber and Barbieri, 2007; Cortes, 2004; Hyland, 2008). The fact that formulaic expression was prevalently found in both hard and social science texts may indicate that formulaic expression has a universal nature in texts rather than unique and specific nature. In other words, the use of formulaic expression may have nothing to do with the topic and content of discourse. By considering the characteristics and features of formulaic expression, it can be concluded that formulaic expression has an economical function that gives the language use fluency and reliability. In the definition of formulaic expression, memorized nature of formulaic expression has been a common element. For instance, Wray (2002) referred to the prefabricated nature of formulaic expression which means that they are stored and retrieved whole from memory. This memory based and prefabricated nature of formulaic expression can give the mind more space for reasoning and content
production. Dechert (1983) called the formulaic expression islands of reliability to point to the efficacy of formulaic expression in helping the language learners to sound more native like.

From another perspective, the frequent prevalent of formulaic expression can be seen as a sign of naturel language use and learning. In other words, the use of formulaic expression is related to the process of language acquisition and language use in general rather to be specific to a particular context of language use. Skehan (1998) in his cognitive approach to language learning pointed out that language learning does not occur solely based on rule system but it makes use of memory based system for more efficiency and fluency.

The results of the study also showed that there was a significant difference between hard and social science EAP books in terms of frequency counts of formulaic expression. Some hypothetical reasons can be named which need further proof of evidence in later studies. For instance, the fact that formulaic expression was more frequently seen in hard science books could be related to nature of hard sciences. In hard sciences, texts are heavily loaded with scientific information about mechanism, processes, and cause-effect relationship (Hutchinson and Waters, 1987) and most of technical words are low frequent words (Gablasova, 2015). Maintaining the relationship between such amounts of information may require more use of formulaic expression to lower the rate of information and to add more fluidity and smoothness to the technical language. Assuming this explanation does not suffice other assumptions can be made. For instance, it is possible that texts in EAP books by SAMT publications have not been modified systematically and based on a coherent framework and this might have caused distortion in the use formulaic expressions. In other words, it is possible that EAP texts analyzed in the current study could not represent the original texts in a systematic way. Aside from this explanation, it can put forward that the texts might not be a true representative of the EAP books by SAMT and more care should have been exercised in sampling the hard and social EAP books. To the knowledge of the researcher, since no study was found to directly scrutinize the difference in frequency of formulaic expressions between EAP books in social sciences and EAP books in hard sciences, this result cannot be compared.
6. Conclusion

The findings of the study led to the conclusion that formulaic expressions are found prevalently in EAP books either in hard sciences or social sciences. Although formulaic expressions were found more frequently in hard sciences both social and hard sciences included frequent use of formulaic expression. It should be noted ideational function of formulaic expressions was more frequently seen in both hard and social sciences. Another point was that no clear explanation can be put forward for justifying the more frequent use of formulaic expressions in hard sciences. Some hypothetical explanations can be made which require further investigations. For instance, the more frequent use of formulaic expressions in hard science EAP books can be due to nature of hard science texts and use of technical words. Therefore, use of formulaic expressions may mitigate the rate of delivery and harshness of technical texts in hard sciences. Based on the results of the present study, a number of implications are assumed. ESP/EAP syllabus designers and materials developers are the two groups that might profit from the findings of the present study by make efforts to design lessons, practices and tasks in hard and social science EAP books, which can enhance EAP students’ knowledge and use of formulaic expressions. Moreover, teachers and instructors are recommended to shed more light on improving the use of formulaic expressions and writing quality of the EAP students.

It is worth noting that in the present study merely four-word bundles have been taken into account. Therefore, further research is deemed essential to be conducted on three- and two-word bundles in order to retrieve a larger number of recurring patterns and consequently get a fuller picture of the use of formulaic expressions in hard and social science EAP textbooks. Moreover, it would be a great help if future investigations were to provide clear criteria specially for the functional classification of bundles, due to the fact that the functional classification was found in the present study to be the utmost challenging part, mainly because the different categories and subcategories need to be well-defined and more in general agreed upon in the literature. In addition, similar studies need to be replicated to reach a more conclusive result about the frequency of formulaic expressions in hard and social sciences. Such studies can eliminate certain shortcomings like sampling and measurement errors and would allow more converging findings regarding the use of formulaic expressions in hard and social science EAP books.
Bibliography


**How to cite this manuscript:** Parastoo Alizadeh Oghyanous. A Comparative Investigation into Formulaic Expressions in EAP Textbooks: *The Case of Social and Hard Sciences, 2018*, 1(1), 36-58.