

GLOBAL ACADEMIC RESEARCH INSTITUTE

COLOMBO, SRI LANKA



GARI International Journal of Multidisciplinary Research

ISSN 2659-2193

Volume: 09 | Issue: 01

On 31st March 2023

<http://www.research.lk>

Author: Nandarathna BGDS, Dhanapala RM, Nizam AAM,
Bandara SMC, Maduwanthi AAI
Rajarata University of Sri Lanka

GARI Publisher | Language | Volume: 09 | Issue: 01

Article ID: GARI/ICM/2022/101A | Pages: 220-231 (12)

ISSN 2659-2193 | Edit: GARI Editorial Team

Received: 27.11.2022 | Publish: 31.03.2023

USE OF CORPUS - BASED APPROACH TO CREATE A LIST OF KEYWORDS FOR TEACHING VOCABULARY TO AGRICULTURE UNDERGRADUATES IN SRI LANKAN STATE UNIVERSITIES

Nandarathna BGDS, Dhanapala RM, Nizam AAM, Bandara SMC, Maduwanthi AAI

Department of English Language Teaching,

Rajarata University of Sri Lanka

ABSTRACT

Corpus based teaching has become a useful approach in second language teaching with the advancement of technology. However, teaching is one of the fields where technology has a little impact. Corpus Linguistics is one of the technology - based tools which has not been adequately used or experimented yet, but could be useful in second language teaching. However, in recent decades, the use of corpora in classrooms has started to progress. For decades, a number of studies have been conducted on the application of corpus linguistics approach in pedagogical aspect. However, a little research has been done on the use of corpus-based approach in ESP classrooms up to date. There are hardly any researchers who have focussed on using corpus linguistic tools to create keywords lists which will be helpful for vocabulary teaching in the ESP (English for Specific Purposes) classrooms, specifically in Sri Lankan State Universities, and that void is addressed through the current study. The research is a quantitative study which is done by using AntConc Corpus Linguistic Tools to create a list of keywords by using the main corpus of the study. CAU (Corpus for Agriculture Undergraduates) is the main corpus of the study and BE06 Corpus compiled by Paul Baker serves as the reference corpus of the study. The analysis section commences with an exploratory process of identifying themes in the main corpus of the study and then, the keywords

were categorized into groups thematically. Finally, a brief description of how these keywords can be used in the ESP classrooms to teach and develop vocabulary lesson was also given. Although, using corpora in language teaching is challenging, it has a big potential in ESP classrooms. Accordingly, the study will open the doors to the field of Corpus Linguistics in ESP classroom in Sri Lankan university context.

Keywords: CAU, Corpus Linguistics, ESP classroom, Pedagogical aspect

INTRODUCTION

Past decades have been a dramatically accelerating pace in the development and also, the adoptions of new technologies. With the development of technology, almost all the fields of the world also got developed or modernized in numerous ways and education is one such field which has been highly influenced by the development of technology. Corpus-based teaching is one of the most important development which has taken place in the field of education with the advancement of technology. Corpus Linguistics, which is viewed as a methodology or research tool by some linguists while the others perceive it as a discipline or theory, has been progressing rapidly within the field of Linguistics since 1964. Corpus Linguists are mainly interested in

descriptive or functional interpretations of language (Meyer,2002), and they study linguistic phenomena through the empirical analysis of large computerized databases of language which are known as corpora. According to Biber, et al., a corpus is “a large and principled collection of natural texts” (Biber, Conrad, & Reppen, 1998, p4) which is compiled and a representation of the language in general, a dialect or other subset of the language. A corpus can be created with language based on transcribed speech, written texts or both. Then, these corpora are stored electronically and then analyzed using computer software programs. Accordingly, Corpus – based approach in language teaching allows to identify the specific words related to a particular field and the behavior of the language in different contexts, which enables teachers to teach language more effectively.

Corpus-based teaching has become useful tools in language teaching with the advancement of technology and computers. Bonelli says, "In the context of the classroom the methodology of corpus linguistics is congenial for students of all levels because it is a 'bottoms-up' study of the language requiring very little learned expertise to start with. Even the students that come to linguistic enquiry without a theoretical apparatus learn very quickly to advance their hypotheses on the basis of their observations rather than received knowledge, and test them against the evidence provided by the corpus". (Bonelli, 2001) According to Salma (2010), a lot of research has been conducted to investigate the effectiveness of using corpus – based approach as a teaching technique to identify the native speakers’ usage of certain vocabulary items, language forms and expressions. Maddalena (2001) says that the use of real life and authentic examples in language classrooms is more beneficial than using examples built by the teachers. Another advantage of this approach in second

language teaching classroom is, it helps second language learners to learn not only the meaning of words, word forms and usage of them, but also the real language use of those words according to the context they appear. Salma (2010) also agrees with the view of Maddalena (2001) that using corpus in language teaching classroom is highly beneficial for second language learners and further she says that teachers should be properly trained to use corpora in classrooms at first to get the maximum benefits of this approach.

Accordingly, Corpus – based language teaching approach can be presented as a useful and effective approach in second language teaching. But, unfortunately this approach is not used widely in second language teaching classrooms and it has not been experimented yet adequately. Even though a number of studies have been conducted on the application of corpus linguistics approach in pedagogical aspect for decades, a little research has been done up to date on the use of corpus-based approach in second language teaching classrooms, specifically in English for Specific Purposes (ESP) classrooms. Further, none of the studies have focused on using corpus-based approach in ESP classrooms in Sri Lankan state universities. The purpose of the current study is to address the gap that has not been addressed through the prevailing research and show that corpus-based approach can be used in ESP classrooms in state universities to teach subject related vocabulary for undergraduates effectively. The current study mainly focuses on the use of Corpus-based approach for vocabulary improvement of the Agriculture undergraduates of Sri Lankan State Universities. Accordingly, a corpus was created using some of the written texts which were given as reading materials for the First year students of the Faculty of Agriculture, Rajarata University of Sri Lanka. Then, a keyword list was created using that main corpus and BE06 corpus

compiled by Paul Baker was used as the reference corpus. Then, the keywords were categorized thematically and the way that they can be used in vocabulary teaching is explained descriptively in the discussion section while both the limitations of the study and methodology are discussed in the conclusion section of the study.

Although using Corpus - based approach in second language teaching is challenging, it has a big potential in ESP classrooms and there are so many advantages in this approach for the second language learners. Thus, if the teachers are trained adequately on how to design suitable corpus- based tasks, they can help their students get exposed to a broader framework of how certain words are used in appropriate contexts effectively and they can make students familiar with the vocabulary related to a specific field of study through which both students' language skills and subject related knowledge can be improved. Accordingly, the current study aims at investigating the possibility of using Corpus-based approach to teach vocabulary to Agriculture undergraduates in Sri Lankan state universities as Corpus- based approach in ESP classrooms is a field that should be developed in future.

LITERATURE REVIEW

The inception of the Computer Assisted Learning (CAL) developed the main sectors in the field of education without an exception to language learning by forming Computer-Assisted Language Learning. (CALL). This novel method of teaching using technology enhanced the efficiency of the teachers to cater to the needs of the learners while allowing more individual learning time for the learners. Giavrimis pointed out that the resources that are available on the internet develop the inquiry skills which are necessary for critical thinking and activate the schema in

order to absorb the necessary information and to reject if there's any by the learners. Accordingly, learners and teachers alike prefer to use online materials which are more easily accessible than other time-consuming resources. For example, the use of the online version of the dictionary has instant access to almost every word, and the use of corpus in language teaching. Among these, the corpora act as a reliable resource for Second language learners (L2) to identify and gather the vocabulary that is more likely to encounter in their general use of the language or in their academic lives.

The Definition of the Corpus and Corpus Linguistics

A corpus, as described by Kennedy, is a body of written texts or transcribed speech that can serve as a basis for linguistic analysis and description. (2014). It is a vast area that studies the structure of authentic speech or text to enable the exposure to the language mainly for the language teachers and learners; vital for those who use the language as a second language. Thus, the corpus is not confined to conducting research but it provides the backbone for the language teachers and learners to grasp the structure of the language, and what a language is in a more reliable way rather than just relying on the native speakers' intuition on the language. (Houston, 2018).

Corpus linguistics, on the other hand, means the use of technology or machine-readable texts to conduct either quantitative or qualitative linguistic analyses. (Conrad, 2000). Recent studies have used corpus linguistics in the field of education especially employed to explore language-related areas like language teaching and learning concerning a second language. Accordingly, most of the studies focussed on the effectiveness of the use of corpus in various language phenomena such as language acquisition, grammatical constructions, discourse patterns, and

vocabulary analyses. (Teubert & Cremakova, 2007).

Language Education in Corpus Linguistics

Corpus Linguistics has a vital application in education pedagogy, especially in the context of learning and teaching a language as a second language because the more the learners are exposed to an ample authentic text, it enables them to learn new vocabulary through immersing themselves in the language which helps for greater memorization. (Donesch-Jezo, 2013). Hence, the use of corpora to teach vocabulary is proved to be a supportive approach for L2 learners mainly because the corpora are authentic examples of real language use. Hou pointed out that “both dictionaries and computerized corpora are key resources for L2 learners” (2014) yet Johns and Winchmann believed that corpora are more effective to immerse the learners in a particular language rather than using the dictionaries given the authenticity of the language use. Thus, they make a significant contribution to the whole L2 learning process while learning the accurate use of the language. These contributions can be categorized into three based on their advantages: “corpus-based vocabulary teaching can help to avoid the negative transfer of the native language effectively; the statistics of the word frequency helps save the efforts in vocabulary teaching; the corpus can represent the natural contexts where the words are used in, which helps them acquire the usages of the words accurately”. (Lu, 2018).

The advantage of using corpus linguistics to avoid the negative transfer

When learning another language other than the native tongue, the learners will be somehow negatively affected by the native language which ends with high-frequency errors mainly because of the transferring of habits unsuccessfully from the

speaker’s L1 causing a negative transfer. (Dechert, 1983 and Ellis, 1997). This negative transfer obstructs the accurate understanding of the new knowledge due to adopting the L1 rules in learning another language. However, in this study, this phenomenon of the negative transfer distracts ESL learners when studying the new vocabulary as this study focuses on using wordlists or the corpus to teach a specific set of vocabulary.

Further, when teaching the new words to the learners in the target language, the learners may use a bilingual dictionary which may or may not explain all the usages in it yet it never be able to explain the usage explicitly thus making confusion for the learners of the target language. (Lu, 2018). However, by considering corpus linguistics when teaching target language vocabulary, the teachers can provide the students with the explicit and accurate usage of the word because a corpus is a collection of authentic usage of words or texts. Thus, corpus linguistics has the ability to proliferate the understanding of the learners avoiding the negative transfer caused by their L1.

ESL vocabulary teaching based on the frequency

Since the analysis of the corpus is objective, it is more accurate and reliable than personal judgments. Thus, the learners can know about a list of frequent vocabulary from the scientific perspective rather than a personal, subjective one. What’s more interesting is that a corpus can estimate the most frequently needed vocabulary for a context. Taking it into consideration, corpus analysis leads scholars to the notion of the lexical syllabus; an approach to language teaching that directly uses corpus findings to organize the content for teaching around the most frequent verbs based on the subject. Sinclair and Renouf (1988) were the first authors who suggested this innovative lexical syllabus following their work on the COBUILD project in which

later it was expanded to Collins COBUILD English Course by Willis and Willis' (1988). Also, John proposed that corpus data can be used in the classroom directly in a way of data-driven learning (DDL) which means learners themselves explore the authentic language materials to identify the common grammatical patterns as well as necessary lexis. According to him, "the use of computer-generated concordances in the classroom to get students to explore regularities of patterning in the target language, and the development of activities and exercises based on concordance output." (qtd in. Rana & Amin, 1991).

According to the study of Brown Corpus by Francis and Kucera, the most frequently used 1000 words show 72% of the content in general texts while the top 2000 words cover 79.7% which asserts that approximately 80% of the content is covered through common words while affirming that the first 2000 words act as the most frequently used words thus confirming that those words are needed by a beginner at the very basic level for the purpose of acquiring adequate knowledge of the target language. However, as pointed out by many linguists, a corpus can aid to comprehend new additions to these 2000 words based on the learner's purpose of learning the target language. As pointed out by Lu,

"For those who are going to further their education after the basic school years, they need to acquire at least 836 academic words. To know which are the relatively high frequent words is rather meaningful for vocabulary teaching, for it basically decides the objects, or the important points of vocabulary teaching, and can also help to arrange the sequence of the words which are going to be taught, which can make vocabulary teaching more effectively." (2018).

The use of context clues to infer the meaning is one of the strategies used by language learners to learn vocabulary. In corpus-based research also, a corpus can be used to teach necessary vocabulary to the learners. What's specific in using a corpus is that it can be used to show all the contexts where a particular word occurs unlike inferring the meaning based on the context. Thus, corpus linguistics is more reliable to rely on to learn the several occurrences of the same word rather than relying on intuition based on contextual clues. (Lu, 2018). Accordingly, the importance of vocabulary teaching in foreign language teaching using a corpus has aroused more and more attention. As a result, Liang introduced different types of corpora and emphasized the necessity of using them in the context of teaching English because he discovered the relationship between corpora and vocabulary teaching through analyzing the theoretical foundation of English vocabulary learning strategies. (2015).

Furthermore, Wang found out the importance of using the corpora to teach vocabulary effectively by incorporating the Schema theory, lexical approach, and lexical grammar. He adopted Gui's approach (2003) with respect to Schema theory and came up with the notion that learners have the ability to construct language schema in order to enhance English proficiency through acquiring new knowledge and Corpora could cater adequate English materials for the process of constructing and reconstructing the schema related to gain new English knowledge. (2012). The use of Corpora shows its value in English vocabulary teaching because of the great convenience in searching and the authenticity that it provides for the target word as there are specialized corpora specific to particular contexts that have been developed recently. Thus, for the purpose of this study, the selected corpus contains the

Use of the context to teach the vocabulary

agriculture context to serve better for these learners.

Use of specialized vocabulary for ESP teaching and learning

Nation pointed out that the specialized vocabulary is “recognizably specific to a particular topic, field or discipline” because he noted that this specialized vocabulary comes from high-frequency or low-frequency words from its related context. (2001). However, acquiring the specialized vocabulary is essential for learners to get the necessary academic knowledge and literacy related to their own discipline. In these recent years, a number of studies have focussed on finding out what determines a word that is specialized in a certain context or a textbook, and also further studies commenced to address the relationship of acquiring this specialized vocabulary by the L2 learners of the target language. (Chung & Nation, 2004; Fraser, 2004; Kaewphanngam, 2002; Rizzo, 2010; Sutarsyah et al., 1994).

Waring and Nation stated that a learner should have the ability to know approximately 95% of the surrounding vocabulary related to a particular context to successfully comprehend the gist of the text. If the learners are unable to do so, they need to continuously stick to use the dictionaries which are frustrating and less fruitful. Besides, 30% of a particular context constitutes the specialized vocabulary that the readers are not capable of understanding without the dictionary. (2004). Furthermore, research by Chung and Nation found out that there are around 4,270 specialized words in an anatomy textbook, which is a large number of words just in one book for the students to refer to the dictionaries to acquire the meanings of those words. (2003, 2004). Accordingly, with such a backdrop, it would be difficult for non-native speakers of English to comprehend the meaning of words as this technical vocabulary is specific to certain contexts.

To overcome this issue, Chung, Nation, and other researchers conducted a few types of research to estimate the adequate number of specialized vocabularies that would be needed for higher studies. Nation estimated that the specialized vocabulary for a particular-field ranges from 1000 to 5000 words based on the research done by Coxhead and Nation in 2001 which represented 2000 most common words from General Service List which have appeared in approximately 80% of academic texts based on introductory level undergraduate textbooks. (2008). However, as these researches were based on preliminary level undergraduate texts, they may not have drawn sufficient data. Hence, some researchers have advocated the necessity of paying special attention to specialized vocabulary teaching for the learners who follow ESP courses in which those classes are mainly aimed at improving the vocabulary they seek for their disciplines. Moreover, Hyland and Tse recommended preparing “corpus-informed lists” from the texts that the learners will use to study and refer to in their own academic disciplines. Thus, they can practice that specialized vocabulary along with their peers and teachers in the ESP classes. (2007).

In summary, the number of researches dedicated to vocabulary teaching in the ESP context is limited even though researchers and practitioners have identified the usefulness of corpora in language classrooms based on corpus-based lists. However, one controversial issue still persists which is whether to teach vocabulary implicitly or explicitly for a better understanding of the learners. Schmitt and Nation stated it is better to use a combination of both approaches as some of the vocabulary is best explained through repeated exposure to certain vocabulary activities whereas other characteristics like grammar features should be explained explicitly. To follow both the approaches for ESP learners,

Corpus-based language learning should be implemented and as a result, it allows to meet the needs of learners by stimulating their motivation. (Sun, 2007).

METHODOLOGY

The study is a corpus-based study which is done using the Ant Conc software to create a keyword list for the Agriculture Undergraduates. The purpose of creating the keyword list is to show the possibility of using Corpus– based approach in the ESP classroom for the vocabulary improvement of the Agriculture Undergraduates. The corpus created using the software was named as Corpus for Agriculture Undergraduates (CAU) and the British English 2006 (BE06) corpus compiled by Paul Baker was used as the reference corpus of the study. 100 vocabulary items were selected from the list of keywords and they were categorized thematically. Thematic categorization of keywords was done as it helps teachers to design vocabulary lessons easily. Finally, two vocabulary related sample lessons were created to show that the Corpus-based approach is applicable and effective in the ESP classrooms in state universities of Sri Lanka.

Corpus Analysis Tool

AntConc 4.0 version of AntConc software which is available for free download in the Internet is used as the corpus analysis toolkit of the study. This

software is a very useful toolkit which is designed by Laurence Anthony for finding frequency patterns of word sequences and keywords of a corpus. This is a multi – purpose corpus analysis toolkit designed specifically for use in the classrooms.

Main corpus and Reference corpus of the study

Corpus for Agriculture Undergraduates abbreviated as CAU is the main corpus of the study. It is created with the reading materials provided for the First-Year undergraduates of the Faculty of Agriculture, Rajarata University of Sri Lanka. The reference corpus of the study is British English 2006 (BE06) Corpus compiled by Paul Baker. The reference corpus was not downloaded from internet as it was already installed in the software itself.

Creating a list of keywords and thematic categorization of keywords

A list of keywords was created using the AntConc software and it was done by comparing the main corpus of the study and the reference corpus of the study with each other. The software generated a list of keywords which contains 1388 keywords. Out of them, 100 keywords were randomly selected by giving priority to the content words for the thematic categorization of keywords. Accordingly, keywords were categorized under 08 themes as shown in the following table.

S/N	Theme	Keywords
01	Words related with machinery and farm	Farmer, Fertilizer, Laboratory, Manure, Mill, Straws, Stovers, Poultry, Forages, Hay, Residual
02	Words related with plants	Foliage, Mucilage, Orange, Oryza, Panicum, Pipal, Vigna, Sago, Sepium, Soybeans, Tea, Fruit, Mustard, Jowar, Sugar, Oat, Seaweeds, Sesbania, Tamarind, Cereals, Pod, Pith, Cassava, Sunflower, Corn, Kernels, Husks, Pulse, Stems, Sugarcane, Canola, Bamboo, Cowpea, Tubers, Citrus
03	Words related with animals	Moth, Skin, Sheep, Embryo, Buffoloes, Coat, Goats, Fur
04	Words related with weather and climate	Oxygen, Winter, Rainfall, Summer, Temperate, Warm
05	Words related with geographical locations	Environment, Lunuwila, Maharashtra, Pastures, Peradeniya, Asia, Plains, Sudan, Uttar, African, Tropics, Grazing
06	Words related with processes	Ingredients, Replicates, Steam, Lactating, Lactation, Harvesting, Yielding, Processed, Experimental, Fermentation, Flowering, Crossbred, Sowing, Preparation
07	Words related with calculations	Percent, Unit, Size
08	Words related with chemical reactions and nutrients	Metabolizable, Glucans, Glucosinolate, Mimosine, Yeast, Sulphur, Microbial, Ammonia, Carbohydrates, Enzymes, Fiber

Table 01: Thematic categorization of keywords

DISCUSSION

Corpus – based approach can be used in the second language teaching classrooms in two ways: students use the corpus analysis software directly or the presentation of teacher-prepared corpus – based activities in handouts or presentations. In the direct use of software, learners learn the behavior of the vocabulary terms themselves but, they may find certain difficulties in using the software directly because of the intimidate amount of concordance examples which can be meaningless. In the indirect use of software, the teacher accesses both keywords and concordance lines and selects some examples which are understandable to learners which is more effective than the use of software by the

students themselves. For this study, the second method or the indirect technique was implemented. The use of corpus-based activities in ESP classrooms is essential for the learners to learn vocabulary related to their field of study easily and to get acquainted with real reliable language use. Observing the concordance lines of selected keywords, the students can easily analyze language patterns and collocations of them. Conrad (2000) says that there are three major impacts of corpus-based approach on teaching grammar too in second language teaching classrooms. These three significant impacts are: “Detailed descriptions of grammar rules will be replaced by register-specific descriptions; Grammar instruction will be closely integrated with the teaching of vocabulary.

Lexico-grammar patterns will be central to language description and language learning; Emphasis on structural accuracy will be shifted to the appropriate use of alternative grammatical constructions” (Conrad, 2000). Accordingly, using corpus-based approach in vocabulary teaching in ESP classrooms is effective for the improvement of other language skills

of learns too. As the purpose of the study is to investigate the possibility of using Corpus-based approach in ESP classrooms of state universities in Sri Lanka, two sample vocabulary lessons have been designed by the researchers using the list of keywords created using the AntConc toolkit to show the possibility of using this approach in ESP classrooms in the state universities of Sri Lanka

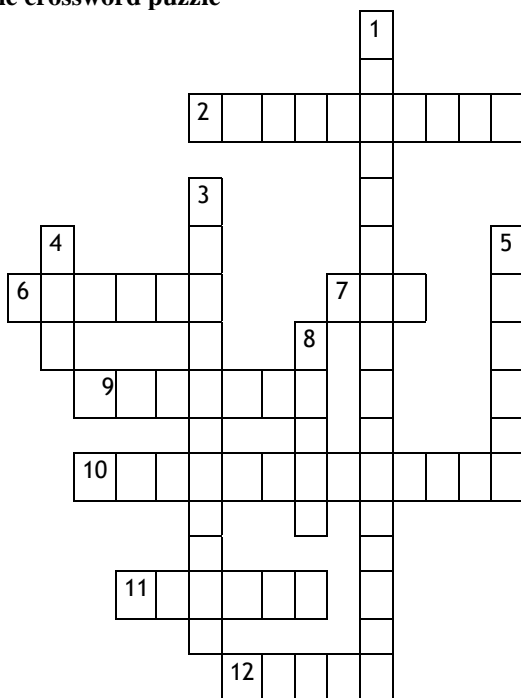
Activity 1

Match the words in “Column A” with their meaning given in “Column B”.

Column A	Column B
Lactating	The act of scattering seeds on land for growing
Harvesting	A process in which an agent causes an organic substance to break down into simpler substances
Fermentation	Production and secretion of milk by the mammary glands
Crossbred	A food to which chemicals have been added that make it last longer
Processed	The activity of picking and collecting crops, animals or fish to eat
Sowing	To cause a plant or animal of a different type in order to produce a new variety

Activity 2

Fill the crossword puzzle



Across

2. A person who studies and experiments with the selection and growth of crop plants and pastures.

6. A person who owns or manages a farm.

7. An animal kept to produce milk or beef.

9. A powerful motor vehicle with large rear wheels, used chiefly on farms for hauling equipment and trailers.

10. A person who studies the origin, management and use of soils.

11. A breed of sheep.

12. A pale yellow oval citrus fruit with thick skin and fragrant, acidic juice.

Down

1. Pollination of a flower or plant with pollen from another flower or plant.

3. The science or practice of farming.

4. Grass that has been mown and dried, often use as sheep food.

5. A tapering orange-colored root eaten as a vegetable.

8. Wheat or any other cultivated cereal used as food.

Accordingly, the Corpus-based approach can be used in the ESP classrooms of the state universities of Sri Lanka successfully even though there are certain limitations and challenges that should be carefully dealt with by the teachers and learners.

CONCLUSION

There are so many advantages in Corpus - based approach for the second language learners in learning second language and although using this approach is challenging, it has a big potential in ESP classrooms. There were certain limitations were identified in the methodology used in the study. The present study has some practical limitations since the selected corpus is very limited as the corpus is specifically created for the ESP learners. Hence, having a limited corpus results in learning a few numbers of words which may not be enough for the students. Also, it is stated that corpus can provide natural contexts for the learners for better understanding, yet it is still not completely possible to represent the textual context in a more authentic way in this study which is related to agriculture and that is another

limitation of this study. Corpus-based approach in second language teaching classrooms is an area where future research is essential. Further studies in this field could explore the impact of using corpus-based activities on the autonomy of the students and more empirical studies are required to explore the long-term effects of learning vocabulary and grammar through corpus-based activities. It is also suggested that further studies are required to investigate the students' attitudes towards corpus-based activities used in the classroom. It is also significant to examine the effect of using corpus-based approach on students' self-confidence or self-efficacy. Accordingly, even though there is a high possibility of using corpus-based approach in ESP classrooms in state universities in Sri Lanka, further research is essential to develop this approach and identify the problems and challenges of this approach.

REFERENCES

Bonelli, E. T. (2001), "Corpus linguistics at work", J. Benjamins, Amsterdam; Philadelphia.

- Breyer, Y. (2008), "Learning and teaching with corpora: Reflections by student teachers", *Computer Assisted Language Learning*, Vol. 22, No. 2, pp. 153-172.
- Chung, T. M., & Nation, P. (2003). "Technical vocabulary in specialized texts; Reading in a Foreign Language", Vol. 15, pp.103-116. <http://nflrc.hawaii.edu/rfl/October2003/chung/chung.html>
- Chung, T. M., & Nation, P. (2004), "Identifying technical vocabulary", *System*, Vol. 32, pp. 251- 263. <http://dx.doi.org/10.1016/j.system.2003.11.008>
- Coxhead, A., & Nation, P. (2001), "The specialized vocabulary of English for Academic purposes", in J. Flowerdew, & M. Peacock (Ed.), *Research perspectives on English for academic purposes*, pp. 252-267.
- Dechert, H.W. (1983), *How a story is done in a second language. Strategies in Interlanguage Communication*, Longman, London.
- Donesch-Jezo, E. (2013), "Using language corpus in teaching foreign language vocabulary", *International Multidisciplinary e-Journal*, Vol. 2, No. 1, pp. 11-25. https://www.researchgate.net/publication/308915434_Using_Language_Corpus_in_Teaching_Foreign_Language_Vocabulary
- Douglas, B., Susan, C., Randi, R. (2000), *Corpus Linguistics, Foreign Language Teaching and Research Press, Beijing*.
- Ellis, R. (1994), *The Study of Second Language Acquisition*, Oxford University Press, Oxford.
- Fraser, S. (2004), *The lexical characteristics of technical texts*. In K. Bradford-Watts, C.
- Giavrimis, P. (2011), "Information and Communication Technologies and Development of Learners' Critical Thinking: Primary School Teachers' Attitudes", *ERIC*, Vol. 4, No. 3, pp.150-160. <https://doi.org/10.5539/ies.v4n3p150>
- Gui, S. (2003), "Memory and English learning", *Foreign Language World*. No.3, pp. 2-8.
- Hou, H. (2014), "Teaching Specialized Vocabulary by Integrating a Corpus-Based Approach: Implications for ESP Course Design at the University Level". *English Language Teaching*, Vol. 7, No. 5, pp. 26-37. [10.5539/elt.v7n5p26](https://doi.org/10.5539/elt.v7n5p26).
- Hunston, S. (2018), *Introduction to a corpus in use*, Cambridge University Press
- Hyland, K., & Tse, P. (2007), "Is there an "academic vocabulary"?", *TESOL Quarterly*, Vol. 41, pp. 235-Q253.
- Kaewphanngam, C. (2002), *A corpus analysis of psychology texts as a basis for the development of teaching materials in English for academic purposes*. Mahidol University, Bangkok.
- Kennedy, G. (2014), *An Introduction to Corpus Linguistics*, Routledge Taylor & Francis Group, London and New York.
- Kübler, S., & Zinsmeister, H. (2015), *Corpus Linguistics and Linguistically Annotated Corpora*, Bloomsbury Publishing, London.
- Liang, S. (2005), "Research on corpora and vocabulary teaching strategies". *CAFLE*, No.5, pp. 33-36. DOI: [10.2991/seiem-18.2019.54](https://doi.org/10.2991/seiem-18.2019.54).
- Lu, H. (2018), "A Study on Corpus-based EFL Vocabulary Teaching". *Advances in Social Science, Education and Humanities Research*, Vol. 286, pp. 211-213. <https://doi.org/10.2991/seiem18.2019.54>.
- Maddalena, S.R. (2001), "An investigation into how corpus analysis may be used in the second language classroom to solve some of the problems surrounding non-native speakers' understanding of seemingly synonymous words", *ERIC*, Vol. 143, pp. 1-14.
- Nation, P. (2001), *Learning vocabulary in another language*. CUP, Cambridge
- Neveen, A. S., & Salma, W. (2010), "Corpus for Classrooms: Ideas for Material Design", *ELR Journal*.

- <https://dbe.metu.edu.tr/convention/proceedingsweb/Corpus.pdf>.
- Rana, A.M., & Amin, Z. (2020). "Teaching vocabulary through corpus to the students of graduation", *International Journal of Management*, Vol. 11, No. 12, pp. 1190-1204. DOI: 10.34218/IJM.11.12.2020.111
- Rizzo, C. R. (2010), "Getting on with corpus compilation: From theory to practice", *ESP World Issue*, Vol. 9, No. 1, pp. 1-23.
- Sinclair, J. M., & Renouf, A. (Eds.). (1988). *A lexical syllabus for language learning*. In R. Carter & M. McCarthy (Eds.), "Vocabulary and language teaching" (pp. 140-158). Harlow: Longman.
- Sun, Y. C. (2007), "Learner perceptions of a concordancing tool for academic writing", *Computer Assisted Language Learning*, Vol. 20, No. 4, pp. 323-343. <http://dx.doi.org/10.1080/09588220701745791>.
- Sutarsyah, C, Nation, P., & Kennedy, G. (1994), "How useful is EAP vocabulary for ESP? A Corpus Based Study", *RELC Journal*, Vol. 25, No. 2, pp. 34-50. <http://dx.doi.org/10.1177/003368829402500203>
- Teubert, W., & Cermakova, A. (2007), *Corpus linguistics: A short introduction*, NY: Continuum.
- Wang, J. (2012), "Corpus-based English vocabulary teaching: theory and application", *Foreign Language Research*, No.4, pp.127-130. <https://doi.org/10.1177/0033688220964162>.
- Waring, R., & Nation, P. (2004), "Second language reading and incidental vocabulary learning", *Angles on the English-Speaking World*, Vol. 4, pp. 97-110.
- Willis, J., & Willis, D. (1989). "Collins COBUILD English course." London: Collins COBUILD.