The Lexical Features of Types of Linguistic Collocations: A Study with Special Reference to Selected Public Administration Circulars in Sinhala and Their English Translations

Malsha, R.P.B.[1], Manoj Ariyaratne[2,*] and Chulananda Samaranayaka[3]

[1, 2] Department of Languages, Faculty of Social Sciences and Languages, Sabaragamuwa University of Sri Lanka, 70140, Belihuloya.
[3] The Ministry of Public Administration, Home Affairs, Provincial Councils, and Local Government.

Abstract

Language is designed to fulfil the communication requirements of communities. Each language has specific and unique characteristics in comparison to any other language. Collocation is also a vital concept that assists in enhancing the quality and formality of both written and spoken languages. The concept of collocation means a set of words that naturally or unintentionally co-occur together to deliver a particular meaning. The attention of many researchers and scholars has been drawn to this concept due to the absence of a recognized or fully accurate definition and classification in this regard. This study aimed at identifying the lexical features of the types of linguistic collocations used in Public Administration Circulars in Sinhala. The study focused on four linguistic collocations suggested by Li et al. (2005). Accordingly, lexical features such as word length, word frequency, high-frequency words, and low-frequency words were identified. The data were collected for the research after extensive analysis of circulars and then the specific sets of words and terms were extracted. Content analysis was used to analyze data. The results of the study indicate that the most frequently used type of linguistic collocation in Public Administration Circulars are fixed collocations. The highest word length of fixed and strong collocations belongs to the 3-letter word group and in loose collocations, the 4-letter word group represents the highest value. Furthermore, the word frequency of fully fixed collocations is equal for each word. The maximum word frequency of fixed collocations and strong collocations is 13, whereas the highest word frequency of loose collocations is 4.

Keywords: High-frequency words, Linguistic collocations, Low-frequency words, Word frequency, Word length

INTRODUCTION

Language is the preliminary resource of communication. Regardless of the medium: written, spoken, or sign language, any language plays an important role in fulfilling the purpose of communication. According to the definition of Oxford Dictionary, language is "the principal method of human communication, consisting of words used in a structured and conventional way and conveyed by speech, writing or gesture". Each language has unique features that represent the culture, social, economic, and political characteristics of the community it belongs to. Lexical features which include word length, word frequency, high-frequency words, and others are a common phenomenon in a language. English and Sinhala do not find any exceptions.

Collocations are part of a language that the language users unconsciously use to communicate their ideas, feelings, and expressions. Researchers and scholars such as Firth, Baker, and Newmark have suggested definitions of collocation. Firth, J. R was the first person who theoretically introduced the concept of collocation (as cited in Krishnamurthy, 2006). The general idea of the definitions suggested by most of these researchers and scholars is that collocations are a set of words that occur together and the formation of collocations in a language is a natural process. However, native speakers of a language use the language-related collocations without making additional effort to identify them whereas non-native speakers deal with certain issues due to a lack of understanding or knowledge regarding collocations.

Hence, many non-native language users often misapprehend collocations with idioms/ fixed expressions. Even so, collocations and idioms are two different concepts, the literal meaning of the words used in idioms does not

* **Corresponding author:** Tel.: +94716824738; Email: manojsab@gmail.com
https://orcid.org/0009-0008-4576-1889

convey the original meaning implied by the idiom whereas the meaning of collocations can be understood through the literal meaning abstracted from the individual words of the collocation. For instance, the collocation, 'pay a visit' is generally understood even by a non-native speaker whereas it is difficult to understand the meaning of the idiom, 'over the moon' as the individual meaning of the words does not define the original implication of this expression.

The knowledge of collocations of any language is paramount as it indicates language proficiency in both written and spoken language and improves the writing style and formality of a text. Apart from that, there is a shortage of a general classification on collocations as the researchers and scholars introduce different views and approaches in this regard. Li, W., Lu, Q., and Xu, R. (2005) have introduced four types of collocations: idiomatic collocations (fully fixed collocations), fixed collocations, strong collocations, and loose collocations based on elements such as compositionality, modifiability, substitutability, and internal association. Furthermore, Hafiz (2002, as cited in Abdelhamid, 2021) has suggested twelve types of collocations based on grammatical and lexical patterns of collocations. Verb + noun, Verb + prepositional noun phrase, Verb + prepositional noun phrase, and Verb + noun phrase are a few types among them. Likewise, several researchers who were interested in the field of collocations have classified collocations based on various features or aspects.

Translation always links languages decreasing cultural, geographical, or linguistic differences. Translation can be generally defined as the process of converting the meaning of a text in one language into another language. According to Catford (1965), "translation is the replacement of textual material in one language (SL) by equivalent textual material in another language (TL)". As translation is always bound with language and communication, the translation process of transferring the meaning of collocations in different languages is worth studying. According to scholars and researchers such as Mona Baker and Peter Newmark, translating collocations from one particular SL to a TL is difficult as a collocation in one language would not be a collocation in another language. Besides, certain words that co-occur in one language will vary in another language. Recently, many researchers are interested in studying different aspects of collocations and the relationship between collocations and translation.

Although several studies have been conducted on collocations by researchers focusing on languages such as English and Arabic, studies on collocations with reference to the Sinhala language have rarely been found. Therefore, this study attempts to identify the lexical features of types of linguistic collocations used in Public Administration Circulars in Sinhala. The following research questions will be addressed at the end of this study: what are the Lexical Features of Types of Linguistic Collocations and what are the linguistic collocation types that are most frequently used in Public Administration Circulars?

The specific objectives of this research are to investigate the lexical features of types of linguistic collocations and identify the linguistic collocation types that are most frequently used in Public Administration Circulars. As the research methodology, a qualitative research approach is used for the present study. 12 Public Administration Circulars will be used as primary data and relative books, e- journals, and other online resources will be used as the secondary data in order to achieve the objectives of the study. All the data

used for this study will be in word form and the linguistic collocations in 12 Public Administration Circulars in Sinhala and the English translation of the same will be the unit of analysis of the research. The data will be collected by reading the selected Public Administration Circulars thoroughly. The collected data will be analyzed using the content analysis method.

Studies on collocations, types of collocations, and translating collocations from one language to another have been conducted by researchers with the purpose of filling the gaps in knowledge. However, a recognized or commonly accepted classification of collocations or a fixed and accurate definition has not yet been introduced. This shortage of knowledge evokes the interest in collocations found in Sinhala language and the unit of analysis of this study. Additionally, there is a shortage of studies to identify the lexical features of types of linguistic collocations. Therefore, this research will be the first attempt to identify the lexical features of types of linguistic collocations used in Public Administration Circulars in Sinhala and their English translations.

The current study is limited to one type of source text: Public Administration Circulars and among all circulars 12 selected Public Administration Circulars will be studied with the purpose of answering the research problem and questions. Additionally, this research is limited to Sinhala and English languages. Although there are numerous aspects that are linked together with collocations, the present study will only be conducted to identify the lexical features of types of collocations. Moreover, the results of the study cannot be generalized beyond the selected unit of analysis. Therefore, future researchers can study the same topic with relevance to a different text type or identify other linguistic features of types of collocations. Moreover, as the present study is related to collocations in Sinhala, future researchers have the opportunity to conduct research with reference to another language pair such as Tamil and English or Tamil and Sinhala.

THEORETICAL BACKGROUND

Language and Linguistics: Language which is the key to communication between members of different social groups and communities, assists all living creatures to express their feelings, attitudes, and emotions using different gestures and sounds. Different scholars and researchers have made efforts to interpret the meaning of the concept, 'language' based on its characteristics, functions, and many other different aspects of language. Chomsky's stated that language is the inherent capability of native speakers to understand and form grammatical sentences. A language is formed of (finite or infinite) sentences, each finite length is constructed out of a few set of elements (Chomsky, 2000).

A language is used with the purpose of covering several functions of communication. According to Alshami (2019), there are two main categories of functions of language as micro functions and macro functions. According to him, micro functions cover "particular individual uses" and macro functions are "related to larger, more general purposes underlying language". The micro and macro functions suggested by Alshami are listed below.

Micro Functions:

- Physiological function
- Phatic function

- Recording function
- Identifying function
- Reasoning function
- Communicating function
- Pleasure function

Macro Functions

- Ideational function
- Interpersonal function
- Poetic function
- Textual function

(Alshami, 2019).

Linguistics is the systematic or scientific study of language. Studying a particular language with the purpose of providing an accurate and complete description for them and studying the nature of language to "find out how language works" are the main goals of linguistics. (Handke, 2014). Handke (2014) states that there are three main branches of linguistics such as sound, structure, and meaning. Phonetics and phonology are studied under sounds, morphology, and syntax belong to the branch, structure and the third branch refers to semantics and pragmatics. In addition to the above-mentioned branches some other hybrid branches of linguistics suggested by Handke (2014) are: psycholinguistics, computational linguistics, sociolinguistics, and corpus linguistics. The definitions mentioned above differentiate the similarities and dissimilarities between the two concepts: language and linguistics.

Collocations: The word collocation has been derived from the Latin verb '*collocare*' and its translated meaning according to Online Etymology Dictionary is place/ arrange. Collocations of a language are generally defined as two-word combinations that precede together. However, collocations have not yet been defined accurately as the studies on characteristics and taxonomies of collocations are under debate. Baker (1992) states that words are combined with other words "to form stretches of language" as words rarely occur individually. This combination of words does not commence randomly but is based on restrictions to convey a particular meaning. Moreover, Baker (1992) explains that collocations are combined in accordance with a grammatically acceptable order and "patterns of collocations are largely arbitrary and independent of meaning".

Collocations are studied under different types suggested by various researchers and theorists as a standard classification of collations has not been identified. Wermter and Hahn (2004) divide collocations into three classes depending on the different degrees of "semantic compositionality of the basic lexical entities involved". (as cited in Anagmpstou et Weir, 2011). The three classes of collocations are idiomatic phrases, support verb constructions/ narrow collocations, and fixed phrases. In idiomatic phrases, the meaning of the collocation is metaphorical or figurative. The second collocation class, support verb constructions/ narrow collocations is defined as the expressions "in which at least one component contributes to the overall meaning in a semantically transparent way and thus constitutes its semantic core" and fixed phrases are the expressions of which the entire lexical meanings of the elements are contributed in the overall meaning in a semantically transparent manner (as cited in Anagmpstou et Weir, 2011). Wanyin Li et al. introduce four (04) types of collocation

based on the properties such as compositionality, modifiability, substitutability, and internal association. The four types are as follows.

Type 0 collocations: "These are fully fixed collocations which include some idioms, proverbs, and sayings and etc." (Wanyin Li, 2005)

Type 1 collocations: "These are fixed collocations in which the appearance of one word implies the co-occurrence of another one". (Wanyin Li, 2005)

Type 2 collocations: "These are strong collocations which allow very limited substitution of components. These collocations are classified with type 3 collocations when substitution can occur at only one end, not both ends". (Wanyin Li, 2005)

Type 3 collocations: "These are loose collocations which allow more substitutions of components; however, a limitation is still required to restrict the substitution". (Wanyin Li, 2005)

Miglė Stulpinaitė. Jolita Horbačauskienė. Ramunė Kasperavičienė, (2016) names these four types as idiomatic collocations (fully fixed collocations), fixed collocations, strong collocations, and loose collocations (normal collocations). However, collocations are often misunderstood by certain students or learners as idioms or other fixed expressions. Collocations are sometimes misunderstood as they are similar to idioms. Pointing out Benson et al.'s (1986) ideas, Stulpinaitė et al., (2016) state that:

"The main difference between these linguistic phenomena is that the meaning of idioms is pre-constructed and not semantically literal; therefore, the words have not retained their conventional literal meanings and cannot be analyzed as such, whereas collocations are pre-constructed and semantically literal (non-metaphorical), which means that the words have retained their conventional literal meanings".

Stulpinaitė et al., (2016) describe that knowledge of collocations is vital for any language learner as collocations make the language sound more natural, expressive, and precise and avoid repetition of any idea while improving writing skills and preserve the formality of a text.

Translation: The field of translation is a fundamental requirement bound with languages as the languages used around the world differ from one to another based on various reasons such as geographical, cultural or political differences and etc. Therefore, translation is always identified as a useful method of communication that contributes to overcoming barriers and obstacles in communication between a number of language communities. Scholars and researchers have made various efforts to define the concept of "translation" based on features, functions, and many other important aspects of translation. Peter Newmark (1988) defines translation simply means the process of rendering the sense of a text from a source language to a target language. He further explains that evaluating the quality of a translation as perfect or correct translation is difficult due to the reason that a translator always works on improving the knowledge and the quality of a particular translation as much as possible.

Newmark (1988) divides translation into two main subcategories as Literary Translation and Non-Literary Translation. De Silva (2018) describes that literary translation is the translation of literary texts such as novels, poems, and dramas and aims at entertaining the reader

through creatively organized texts whereas non-literary translation is explained as the transference of meaning of structured and organized texts with real information related to different subject areas and professions, from one language to another language. The non-literary translation is studied under two subcategories by Gotti and Sarcevic. According to these two researchers, general translation is a translation of text materials that are not related to any specific subject area or type and specialized translation is the translation of text materials that belong to specialized areas or types. Specialized translation targets a specific audience whereas general translation does not pay much attention to special terminologies. Gotti and Sarcevic further state that specialized translation covers all the specialist subjects such as science, technology, law, and marketing.

Administrative translation is also known as another specialized area that belongs to the non-literary translation category. The translation of administrative texts used by different government bodies and organizations can be generally defined as administrative translation. such as Circulars, Memorandums, Gazettes, and Parliament Notes are some examples of administrative texts that are used in the public administration sector.

Public Administration Circulars on which the current study focuses are designed to inform the public servants or relevant public authorities regarding the revisions, inclusions, or removals made to the Establishment Code.

Translating Collocations: Collocations vary from one language to another as languages are always developed with unique features. Accordingly, collocations can be language-specific, culture-specific, or universal. Translating collocations from a source language (ST) to a target language (TT) may affect to create certain problems as it is difficult to translate collocations without any changes (Miglė Stulpinaitė. Jolita Horbačauskienė. Ramunė Kasperavičienė, 2016). Moreover, the translation issues such as cultural untranslatability and non-equivalence exist between languages and can also affect in a negative manner when translating collocations.

According to Newmark (1998), translating "English Collocations which appear arbitrarily to juxtapose nouns with verb-nouns" is difficult as "they indicate the two most significant meaning components, but have varied and sometimes mysterious case-relations". He further explains that Non-British collocations are less challenging to translate because of the 'less arbitrary' nature. To translate collocations translation strategies such as transference, literal translation, and descriptive and functional equivalent can be applied in a way that suits the target audience.

Not only Newmark but also Baker (1992) illustrates that collocational patterning differs from the source language (SL) to the target language (TT) and such differences make problems in translating collocations. The engrossing effect of source text patterning, misinterpreting the meaning of a source-language collocation, the tension between accuracy and naturalness, culture-specific collocations, and marked collocations in the source text are indicated as the translation problems encountered while translating non-literary texts.

In brief, although the arbitrary nature of collocations makes translating collocations a difficult task, overcoming such challenges using various strategies and having a clear and specific understanding of the usage of collocations and the types of collocations are useful in improving language skills.

Empirical Literature

A study under the title "Issues in Translation of Linguistic Collocations" has been conducted by Miglė Stulpinaitė, Jolita Horbačauskienė, and Ramunė Kasperavičienė in 2016 in Lithuania. The aim of the study was to analyse the characteristics of linguistic collocations in English while identifying the possible translation transformations in the Lithuanian language. The study was based on the source texts obtained in English and the translations of the same texts in Lithuania. the classification of collocations which has been introduced by Li et al. (2005) has been applied in this study. Strong collocations, fully fixed collocations, loose collocations, and fixed collocations are the four types that have been used by researchers. The results of the study show that strong collocations are the most frequently used collocation type among the data it is 53% as a percentage. 32% of data were loose collocations and there were only 8% of fixed collocations. Fully fixed collocation type was the least number of collocations used among the data and it is 7% out of all the collected data. To identify the translation shifts in this study, translation shifts suggested by Catford have been applied. However, as more than one-third of the results belong to category shift, the researcher focuses on its four subcategories in this particular study: rank shift, class shift, intra-system shift, and structure shift. According to the results, the most frequent form of translation shift was rank shift and structure shift was the least applied translation shift.

In 2021, a study has been conducted under the title "Translating Collocations from English into Arabic and vice versa: An Empirical Study" by Ibrahim Ahmed Abaker Abdelhamid with the aim of aims "investigating the challenges that Sudanese EFL university students encounter when rendering English collocations into their Arabic equivalences and vice versa as well as the reasons behind these challenges". As the results, the researcher has found that the students have difficulties in translating collocations from English into Arabic as the students are unaware of the linguistic and cultural differences between the two languages and frequently use the literal translation strategy when translating collocations.

"The Translation of Lexical Collocations in Undergraduate Students' Theses' Abstract: Students versus Google Translate" is a study conducted by Clara Herlina Karjo and Ecclesia Metta in 2019. The main objective of the study was to compare the translations of lexical collocations in the abstract sections obtained from students' theses and find out the errors made in translating lexical collocations by Google Translate or student translators. The primary concern of the research was on English and Indonesian languages. The results of the study revealed that Google Translate makes more errors in forming lexical collocations than students. Additionally, Google Translate needs to be improved in order to accurately translate the lexical collocations however, some errors in the translations have occurred due to the inaccuracy of the source text.

Methodological Literature

Considering the above-mentioned studies, different types of research methodologies have been applied by the researchers to achieve the objectives of the above-mentioned research. The study conducted by Miglė Stulpinaitė, Jolita Horbačauskienė, and Ramunė Kasperavičienė, "Issues in Translation of Linguistic Collocations" has used a qualitative research approach. Data for the study has been randomly collected observing 250

English collocations which were included in popular science magazines in English and their translated versions in Lithuanian.

The study, "Translating Collocations from English into Arabic and vice versa: An Empirical Study" by Ibrahim Ahmed Abaker Abdelhamid was designed using a diagnostic test as the data collection method to investigate the awareness of EFL students in rendering English collocations into Arabic and Arabic collocations into English. The test for this study includes two parts and the first part is translating eight sentences with collocations from English to Arabic whereas eight sentences with collocations in Arabic are provided to be translated into English in the second part. In addition, the descriptive-analytical method has been used as the data analysis method. A panel of four university professors was selected by the researcher to ensure the validity of the test. 26 Sudanese EFL students, between 20-30 years old, studying at Nahda College in Sudan, have been selected for achieving the purposes of the study Statistics such as frequencies, percentages, mean, and standard deviation has been used for data analysis.

For the study conducted by Clara Herlina Karjo and Ecclesia Metta, data were collected from twenty theses of English Literature students at Binus University in English and Indonesian. Accordingly, four sets of data have been analyzed: a set of: a set consists of 20 Students' Original Abstracts in English (STE), a set consisting of 20 Google Translations of STE in Indonesian, (GTI), a set comprises of Students' Original Abstracts in Indonesian (STI) and another set comprises of Google Translation of STI in English (GTE). The samples have been selected based on Hill's (2000) taxonomy of collocation, consisting of eight categories. Besides, the collocations in the students' texts and their translations were compared and examined thoroughly to present how students and Google form the collocations and what error types have been made while translating the collocations. Errors in translation were measured based on the formation and meaning equivalence between the source and target language.

Knowledge Gaps

Collocations in a language are an area that many researchers were interested in finding a clear definition, standard, or recognized classification of collocations. Certain researchers were inspired to conduct research in this regard mainly focusing on the issues/ problems arising or strategies applied in the process of translating collocations. Most of these studies have used English collocations or collocations in other languages such as Arabic and Chinese as the main concern of the study. In addition, a few types of research were conducted exploring lexical features of various language outputs such as journal articles and teacher talks.

However, there is a very limited number of research that focus on identifying the lexical features of collocations and there is no study conducted based on studying the collocations in Sinhala language as in Sinhala language, collocations are not studied as a separate field as in English language. Therefore, this study fills this knowledge gap and could be beneficial for future researchers who are interested in Collocations/ Collocations in the Sinhala language, future translators, and learners of translation studies.

METHODOLOGY

The main objective of this study is to identify the lexical features of the types of linguistic collocations used in Public Administration Circulars in Sinhala. The present qualitative study is explanatory in nature and a deductive approach is employed. Primary and secondary data were collected for this research. The researcher used sets of words (collocations) and terms that were abstracted from selected Public Administration Circulars published by the Ministry of Public Administration, Home Affairs, Provincial Councils, and Local Government as primary data and some of the secondary data used for the study were journal articles, relative books, and online resources. All the collected data are in word form. The content of both the source and target documents was extensively analyzed and collected the contextual data for the study using the purposive sampling technique.

As mentioned above after extensive analysis of circulars and then the specific sets of words and terms were extracted.

Primary Data: 12 selected Public Administration Circulars

Table 1: Number of documents and collocations

Number	Name of the document	Number of collocations
1	Public Administration Circular 04/2023	05
2	Public Administration Circular 08/2022	08
3	Public Administration Circular 21/2022	54
4	Public Administration Circular 22/2023	62
5	Public Administration Circular 23/2022	20
6	Public Administration Circular 25/2022	35
7	Public Administration Circular 29/2022	06
8	Public Administration Circular 31/2022	02
9	Public Administration Circular 05/2020	03
10	Public Administration Circular 18/2020	03
11	Public Administration Circular 03/2016	110
12	Public Administration Circular 35/92 (VI)	02

Secondary Data: Textbooks, Journal Articles/ Research Articles, Websites and Lecture Notes

First, the collected data were organized using four collocation types introduced by Li et al. (2005). The four types are as follows:

Type 0 collocations: This type includes fully fixed collocations such as some idioms, proverbs, sayings and etc.

Type 1 collocations: These collocations are fixed collocation in which "the appearance of one word implies the co-occurrence of another one".

Type 2 collocations: These are also known as strong collocation which "allow very limited substitution of components. These collocations are classified with type 3 collocations when substitution can occur at only one end, not both ends".

Type 3 collocations: Collocations belonging to this type allow more substitutions of components when compared to the above types. "However, a limitation is still required to restrict the substitution".

(Li et al., 2005).

The lexical features of each type of linguistic collocation were deeply studied. As the final step, the researchers identified what linguistic collocation types are most frequently used in Public Administration Circulars. Accordingly, attempts were made to identify word length, word frequency, high-frequency words and low-frequency words of each type of linguistic collocations mentioned above. The word length of collocations was calculated manually and the word frequency of collocations was measured using an online word frequency calculator (Word Frequency Counter - Count Word Occurrences - Online - Browserling Web Developer Tools). The collected data was analyzed using the content analysis method.

RESULTS AND DISCUSSION

The main objective of this study is to identify the lexical features of the types of linguistic collocations used in Public Administration Circulars in Sinhala. Meanwhile, the present study investigates the lexical features of types of linguistic collocations and identifies which linguistic collocation types are most frequently used in Public Administration Circulars based on the collected data. **310** collocations were collected from 12 Public Administration Circulars for the present study and the collected data were categorized under four linguistic collocation types suggested by Wanyin Li et al. (2005). The four types of collocations are as follows:

Type 0 - Idiomatic (Fully Fixed) Collocations

Type 1 - Fixed Collocations

Type 2 - Strong Collocations

Type 3 - Loose Collocations

The collected data were categorized under the above-mentioned four types of linguistic collocations based on the following qualities:

Idiomatic (Fully Fixed) Collocations

- Consist of a fixed form
- Constituents of the collocations cannot be replaced with other words.
- The meaning of the collocation cannot be determined by its individual parts.
- Synonyms cannot be substituted.
- Cannot alter the word order of the collocation

Examples:

1.   යතුරු ලේබක   - Typist

2.   පාලන විෂය   - control

The above-mentioned examples indicate that the meaning of the collocations in Sinhala cannot be determined by their individual constituents. In the example, "*යතුරු ලේබක*" is a collocation formed with two words and its English

translation is 'Typist'. The constituent '*යතුරු*' means 'keys' in English and the other constituent '*ලේබක*' means 'writer'. Accordingly, the individual meaning of the two constituents does not convey its original meaning and the word order cannot be replaced or there is no other synonym in Sinhala language to be substituted in order to convey the original meaning. Therefore, it has a fixed form.

Fixed Collocations

- Includes a fixed form
- Constituents cannot be replaced or substituted
- The constituents repeatedly co-occur with each other.
- Cannot be replaced with a synonym to maintain the same meaning.
- The order cannot be altered.

Examples:

1.   විශ්‍රාම වැටුප්   - pension

2.   කාර්යක්ෂමතා කඩඉම   - Efficiency Bar

The first example of this type is "*විශ්‍රාම වැටුප්,*" which in English means 'pension'. In circulars or in formal contexts, the constituent "*විශ්‍රාම*" is often followed by the word "*වැටුප්*" as the second constituent always collocates with the first constituent to indicate the meaning, 'pension'. There are certain words in Sinhala language that repeatedly co-occur with another word in order to exude the meaning or idea of an individual word in another language. For instance, the collocation "*කාර්යක්ෂමතා කඩඉම*" is formed with the combination of both its constituents to indicate the meaning of 'efficiency bar examination'. Hence, the order of the constituents in these fixed collocations cannot be replaced with any similar word in order to determine the same meaning.

Strong Collocations

- A restricted modification of constituents is applicable.
- Such modifications are not allowed to change the order of constituents.
- Constituents can be replaced with a limited number of synonyms.

Examples:

1.   ආකෘති උපයෝගී කරගැනීම   - formats shall be made applicable

2.   විනය ක්‍රියාමාර්ග   - disciplinary actions

The two examples mentioned above are examples for strong collocations. The first example of which the translated meaning in English is 'make formats applicable' includes three constituents. In the Sinhala language, the word '*ආකෘති*' is a noun where the other word, '*උපයෝගී කරගැනීම*' is a verb. Unlike in English, in Sinhala, most of the verbs are formed individually. A verb is usually bound with words such as '*කිරීම*' (to do), '*දීම*' (to give), '*වීම*' (happen), '*ගැනීම*' (take). The general meaning of all these verbs is 'doing something'. Accordingly, the word division methods vary in the Sinhala language in comparison to the word division methods used

in the English language. In this example, the verb 'උපයෝගී කරගැනීම' can be replaced with synonyms such as 'යොදා ගැනීම' and 'භාවිත කිරීම' which have the similar meaning of 'making something applicable'. For the second example, the collocation 'විනය ක්‍රියාමාර්ග' can be replaced with synonyms such as 'විනය කාරණා' without changing the meaning. Although, certain constituents of these collocations can be replaced or substituted with a limited number of synonyms, the order of the constituents is preserved, as shown in the examples.

Loose Collocations

- Substitution of constituents is allowed.
- Can be replaced with synonyms and word combinations that have an equivalent meaning.
- Orders can be changed.

Examples :

1. වැරදි ලෙස නිර්වචනය කරනවා - misinterpretation

2. පරිසරය සුරැකීම - protection of environment

The collocation type named 'Loose Collocations' has a minimum number of restrictions on replacing constituents in comparison to the other three collocation types mentioned above. The first example suggested under this type, 'වැරදි ලෙස නිර්වචනය කරනවා' which means misinterpretation, can be replaced with 'සාවද්‍ය ලෙස නිර්වචනය කිරීම', 'වැරදි ලෙස අර්ථ දැක්වීම', 'සාවද්‍ය ලෙස අර්ථ දැක්වීම' without changing the meaning. The order of the constituents in the given suggestions for the first example has also been altered while preserving the sense of the collocation. The second example, 'පරිසරය සුරැකීම' (protection of the environment), belongs to the type of loose collocations as the said collocation can be replaced as 'පරිසරය ආරක්ෂා කිරීම, ස්වභාවධර්මය රැකගැනීම' preserving the original meaning of 'protecting the environment'. As shown in the above examples, the quality of allowing to collocate with the number of words.

Based on the criteria given above, table 2 shows that there were 2 fully fixed collocations, 163 fixed collocations, 125 strong collocations, and 20 loose collocations among the 310 selected collocations.

Table 2: Types of Linguistic Collocations

Type of Collocation	No. of Collocations	As a percentage
Fully Fixed (Idiomatic) Collocations	2	0.6%
Fixed Collocations	163	52.6%
Strong Collocations	125	40.3%
Loose Collocations	20	6.4%
TOTAL	310	

As a percentage, 0.6% of collocations were fully fixed collocations, 52.6% were fixed collocations, 40.4% of selected collocations were strong collocations and 6.4% were loose collocations. Accordingly, fixed collocations are the most frequently used collocations in the selected Public Administration Circulars in Sinhala and fully fixed collocations or idiomatic collocations are the least frequently used collocation type in the selected circulars.

Collocations are known as sets of words that co-occur in a constant pattern. Such word combinations are used by native speakers of a language unconsciously, as one word follows the other naturally without making an additional effort to form them. Additionally, collocations are formed with the combination of lexical items such as nouns, verbs, adjectives and etc. According to the present study, the following word combination patterns were identified:

- Adjectives + Nouns

Examples :

1. විශේෂ දීමනාව - special allowance

2. සාධාරණ සේවා කාලය - reasonable period

3. පරිපාලනමය අවශ්‍යතා - administrative requirements

4. ප්‍රියමනාප සේවා ස්ථානය - popular service station

5. මූල්‍යමය ප්‍රතිලාභ - financial benefits

- Nouns + Nouns

Examples:

1. රාජකාරි ස්ථානය - service station

2. සම්මන්ත්‍රණ ශාලාව - conference hall

3. වැටුප් හා දීමනා - salaries and allowances

4. වැටුප් ව්‍යූහය - salary structure

5. ආයතන සංග්‍රහය - Establishments Code

- Verb + Verb

Examples:

1. දන්වා එවීම - inform

- Nouns + Verb

Examples:

1. වැටුප් පරිවර්තනය - conversion of salary

2. ආකෘති උපයෝගී කරගැනීම — - formats shall be made applicable

3. අවධානය යොමු කරවීම — - attention is drawn

4. වැටුප් පියවර ගණනය කරනවා — - calculate the salary point

5. දීමනා ගෙවීම — - pay allowance

- Verbs + Nouns

Examples:

1. සේවය හැරයාමේ නිවේදනය — - the notice on vacation of post

- Adverb + Verb

Examples:

1. සීමාවන සේ පවත්වාගෙන යාම — - to continue confining to…..

2. නිවැරදි බවට සහතික කරනවා — - certify as correct

3. නියමිත පරිදිම සම්පූර්ණ කිරීම — - duly complete

4. අදියර වශයෙන් කියාත්මක වන — - implemented on phase basis

   සෑපුවම දන්වනු ලබන — - directly inform

- Adjective + Adjective

Examples:

1. සාවද්‍ය හෝ අසම්පූර්ණ — - false or incomplete

2. <u>සකිය හා සතුටුදායක</u> සේවා කාලය — - active and satisfactory period of service

- Adverb + Adverb

Examples

:

1. සත්‍ය හා නිවැරදි බව — - true and accurate

- Noun + Preposition

Examples:

1. නිර්දේශයක් රහිතව — - without a recommendation

2. නිර්දේශයක් සහිතව — - with recommendation

3. පූර්ව අනුමැතීන් මත — - permitted

- Verb + Preposition

Examples:

1. එකඟතාව ලබා ගැනීමට යටත්ව — - <u>subject to</u> the concurrence of

In addition to that, at certain instances, conjunctions such as "*සහ*" (and), "*හෝ*" (or) and "*පිළිබඳ*" (means 'about') were used in collocations in Sinhala.

Example:

ආගමන <u>හා</u> විගමන — - Immigration and Emigration

අල්ලස් හෝ දූෂණ චෝදනා — - Allegation of Bribery or Corruption

අඛණ්ඩව හෝ බණ්ඩව — - Uninterrupted periods or periods with service break

Official or formal documents, such as circulars, are usually translated into English and Tamil in Sri Lanka. The collocations used in these documents are also translated into the target languages using collocations or paraphrasing the idea depicted in collocations. In this regard, collocations cannot be translated into the target languages based on word-for-word translation strategies but rather on sense-for-sense translation techniques. When a source text is translated into a target language, a number of changes occur within the translation process in order to convey an equivalent meaning to the target audience. Such changes are also identified in translating collocations from Sinhala to the English language.

In consequence, in certain instances, the order of a collocation in Sinhala shifts when translating the particular collocation into English. In the Sinhala language, the sentence structure is subject + object + verb. In English sentences, this structure or order changes as subject + verb + object. In addition, the majority of verbs are generally followed by adverbs, and most of the nouns are followed by adjectives. However, this general structure changes at times when translating certain collocations in Sinhala into English. For instance, the noun and verb collocation '*උපදෙස් නිකුත් කිරීම*' of the source text (ST) is translated into English as 'issue instructions'. According to this example, the noun and verb collocations of the Sinhala text became verb and noun collocations in the target text (TT). The adverb and verb collocation '*කල්පිතව ගණනය කරනවා*' of the ST was changed into a verb, preposition and adverb collocation, 'calculate on the hypothetical basis', in the TT. Apart from that, certain collocations that were formed with the combination of noun + noun + noun + noun such as '*පවුල් සෞඛ්‍ය සේවා නිලධාරීන්*' were translated into English using an individual word, such as 'midwife'. Furthermore, verb and verb collocations such as '*දන්වා එවීම*' was changed into the verb 'inform' in the translation process. Even so, such changes do not affect the original meaning that exists in the collocations of the ST.

Lexical features of types of linguistic collocations

In this study, the collocations, which were gathered as data, were categorized under four types of linguistic collocations namely: idiomatic or fully fixed collocations, fixed collocations, strong collocations, and loose collocations. As mentioned above, one of the objectives of this study is to investigate the lexical features of different types of linguistic collocations. Lexical features are a common phenomenon in a language and word length, word frequency and high

frequency words are some of the lexical features that are included in this phenomenon. Accordingly, the researcher attempted to identify word length, word frequency, high-frequency words, and low-frequency words that belong to each type of linguistic collocation mentioned above.

Word length: The word length in English is usually measured by calculating the letters and the same method is used in this study to calculate the word length in collocations selected from Public Administration Circulars in Sinhala. The lexical feature, word length, indicated the complexity and difficulty of particular texts or lexical units.

Table 3: Word Length

	Idiomatic or Fully Fixed Collocations	Fixed Collocations	Strong Collocations	Loose Collocations
1 –letter words	-	18	8	-
2 –letter words	-	60	47	8
3 –letter words	4	144	133	20
4 –letter words	-	117	94	21
5 –letter words	-	54	55	9
6 –letter words	-	31	25	5
7 –letter words	-	12	11	1
8 –letter words	-	4	1	-
9 –letter words	-	1	-	-

Among the selected data, only two (2) collocations belonged to the collocation type named idiomatic or fully fixed collocations. As shown in Table 2, the proportion of the 3–letter words group is the highest and its numerical value is 4. In fixed collocations, 3 –letter words group represents the highest word length and there are 117 4–letter words. The highest word length of strong collocations is 133 and it is the proportion of 3-letter words. In addition, the proportion of 4- letter words is the highest value in the loose collocations and the mean value is 1, which represents the proportion of 7–letter words. According to the data given in the table, the proportion of 3–letter words and 4–letter words shows the highest values in each type. Furthermore, among the collected data, fully fixed or idiomatic collocations, fixed collocations and loose collocations have no examples that exceed 9–letter words.

Word frequency: Word frequency, which assists in identifying key words and common words in a corpus, can be defined as the frequency or the number of times a word occurs in a corpus. Word frequency is significant as it makes vocabulary and the content of a text more understandable.

Table 4: Word Frequency of Types of Collocations

Idiomatic or Fully Fixed Collocations		Fixed Collocations		Strong Collocations		Loose Collocations	
	frequency		frequency		frequency		frequency
පාලන	1	සේවා	13	කිරීම	23	කරනවා	4
විෂය	1	සේවය	12	වැටුප්	12	ක්‍රියාත්මක	3
යතුරු	1	හා	11	කරනවා	8	කිරීම	3
ලේබක	1	ලේකම්	9	පරිදි	6	වශයෙන්	2
		දීමනාව	7	ගැනීම	6	වන	2
		නිවාඩු	6	කාලය	6	වැටුප	2
		කිරීම	6	දීමනා	5	ලබන	2
		වැටුප්	5	හා	5	ගණනය	2
		රාජ්‍ය	5	සේවා	5	අදියර	1
		හෝ	4	ලබා	4	ආරම්භක	1
		විශේෂ	4	සම්පූර්ණ	4	මාසික	1
		නිලධාරී	4	සංශෝධනය	4	ඉල්ලා	1
		සහකාර	4	ස්ථාන	4	සිටිනු	1
		පොලිස්	4	මාරු	4	සහන	1
		ස්ථාන	4	ඉදිරිපත්	3	උපදෙස්	1
		මත	3	උපදෙස්	3	නිකුත්	1
		ආයතන	3	කරුණු	3	උපරිම	1
		පරීක්ෂක	3	මුදල්	3	කාල	1

		පටිපාටිය	3	රාජ්‍ය	3	සීමාවන්	1
		කාල	3	භාෂා	3	වීම	1
		කාලය	3	විනය	3	කල්පිතව	1
		සහ	3	නොවන	2	ජනතාව	1
		දෙපාර්තමේන්තුව	3	ශ්‍රේණියට	2	දැනුවත්	1
		පළාත්	3	උසස්වීම්	2	ණය/	1
		මාරු	3	කටයුතු	2	අත්තිකාරම්	1
		පවුල්	2	සුදුසුකම්	2	ලබාදීම	1
		වෛද්‍ය	2	ඇති	2	තාවකාලික	1
		නිල	2	අභියාචනා	2	පියවරක්	1
		අමාත්‍යාංශ	2	ආකෘති	2	තොරතුරු	1
		දෙපාර්තමේන්තු	2	සාමාජිකත්වය	2	රැස්	1
		ගැනීම	2	ඉල්ලීම්	2	දෛනික	1
		අධ්‍යක්ෂ	2	උප	2	චේතන	1
		ජනරාල්	2	තෘතිය	2	පරිසරය	1
		ප්‍රධානියා	2	මට්ටම	2	සුරැකීම	1
		අත්තිකාරම්	2	යටත්	2	පූර්ණ	1
		කිරීමේ	2	වැටුප	2	මූලික	1
		පත්	2	දායක	2	මූල්‍යමය	1
		කරනවා	2	අයකිරීම	2	අගය	1
		කළමනාකරණ	2	දීර්ඝ	2	රාජකාරියක්	1
		කාර්ය	2	නියමිත	2	නිර්වේතනික	1
		නිලධාරීන්	2	නිර්දේශයක්	2	තත්ත්වයෙන්	1
		කැමැත්ත	2	පත්වීම	2	ඉටු	1
		කොමිෂන්	2	පර්යේෂණ	2	වැටුප්	1
		සභාව	2	බඳවා	2	පියවර	1
		ශිල්පී	2	රෝග	2	වැරදි	1
		වියදම්	2	ශ්‍රේණිය	2	ලෙස	1
		තනතුර	2	රජයේ	2	නිර්වචනය	1
		නිවැරදි	2	රාජකාරි	2	සාවද්‍ය	1
		පළපුරුද්ද	2	ක්‍රියාත්මක	2	වර්ගීකරණය	1
		යටතේ	2	ස්ථානමාරු	2	සම්පූර්ණව	1
		නීති	2	ක්‍රියාවලිය	2	දන්වනු	1
		ලිපිය	2	වැඩ	2		
		වෘත්තීය	2	සකස්	2		
		සෞඛ්‍ය	2	තීරණ	2		
		උපාධිය	2	සාධාරණ	2		
		කටයුතු	2	අගතියක්	1		
		ප්‍රාදේශීය	2	අතිකාල	1		
		ස්ථානය	2	අඩි	1		
		මහජන	2	අධ්‍යාපන	1		
		වාර්තා	2	නභා	1		
		අමාත්‍යාංශය	2	සිටුවීම	1		
		කාර්යසාධනය	2	අනිවාර්යෙන්	1		

		පුරුෂ	2	වාර්තා	1		
		විශ්‍රාම	2	(සේවය	1		
		බව	2	සඳහා)	1		
		ශ්‍රී	2	අනිවාර්යයෙන්ම	1		
		ලංකා	2	අභිමතය	1		
		සාමාන්‍ය	2	අරමුණු	1		
		අබණ්ඩව	1	ළහා	1		
		බණ්ඩව	1	කර	1		
		අඩ	1	උපයෝගී	1		
		අඩු	1	කරගැනීම	1		
		අදායම්ලාභී	1	ආදර්ශ	1		
		අති	1	පත්‍රය	1		
		ගැසට්	1	ආධුනික	1		
		පත්‍රය	1	පුහුණුලභින්ගේ	1		
		අතිරේක	1	ආශ්‍රය	1		
		අතුරු	1	ඉතා	1		
		අත්සන	1	වගකීමෙන්	1		
		මුද්‍රාව	1	යුතුව	1		
		අධිකරණ	1	යොමුකිරීම	1		
		අධ්‍යයන	1	සලකා	1		
		අධ්‍යාපන	1	බලනවා	1		
		පරිපාලන	1	ඉහළ	1		
		ව්‍යවස්ථාව	1	කාර්යාලය	1		
		අනුකූල	1	වගන්තිය	1		
		නොවන	1	අනුගමනය	1		
		පරිදි	1	උපරිම	1		
		අනුයුක්තය	1	උපරිමයෙන්	1		
		අවසන්	1	ඔබ්බට	1		
		අන්තර්	1	එකඟතාවය	1		
		අමාත්‍ය	1	ගැනීමට	1		
		මණ්ඩලය	1	යටත්ව	1		
		අමාත්‍යාංශ	1	එරෙහිව	1		
		අල්ලස්	1	දැක්වීම	1		
		දූෂණ	1	කමිටු	1		
		චෝදනා	1	නාමිකව	1		
		අවධානය	1	පිහිටුවීම	1		
		යොමු	1	තහවුරු	1		
		කරවීම	1	දැක්වීම්	1		
		අවශ්‍යතාවය	1	සමාලෝචනය	1		
		රඳවා	1	කළමනාකරණ	1		
		ආගමන	1	ක්‍රමානුකූලව	1		
		විගමන	1	දෙනවා	1		
		සංග්‍රහය	1	වකුලේබ	1		
		ආශ්‍රිත	1	නිකුත්	1		
		උත්සව	1	ජ්‍යෙෂ්ඨත්වය	1		

		උප	1	තීරණය	1		
		උප්පැන්න සහතිකය	1	තාක්ෂණ සේවය	1		
		උසස්	1	තාවකාලිකව	1		
		උසස්වීම්	1	අනියුක්ත	1		
		ලිපි	1	ජොෂ්ඨ	1		
		එකහතාවය	1	හෝ	1		
		ඇතිව	1	වෘත්තිය	1		
		කමිටුවක්	1	දඩුවමක	1		
		පටිපාටික	1	බලපෑමට	1		
		රීති	1	වීම/	1		
		කාර්යක්ෂමතා	1	භාජනය	1		
		කඩිඉම	1	වීම	1		
		කාර්යසාධන	1	දළ	1		
		වාර්තාව	1	ගෙවීම	1		
		කාර්යාල	1	ප්‍රතිපූර්ණය	1		
		සහයක	1	කාලයක්	1		
		කාර්යාලගත	1	ධුරය	1		
		පරාසයන්	1	දැරීම	1		
		සටහන	1	නැවත	1		
		සීමාව	1	සේවයේ	1		
		ඉකුත්	1	යෙදවීම	1		
		වීම	1	නිත්‍යානුකූල	1		
		පළ	1	ගිවිසුමකට	1		
		විමසීම	1	බැඳී	1		
		බන්දු	1	නියමය	1		
		රේඛන	1	නියමයන්ට	1		
		ගමන්	1	අනුගතව	1		
		ගැළපීමේ	1	පරිදිම	1		
		ගිවිසුම්ගත	1	නියාමන	1		
		බලධරයා	1	රහිතව	1		
		ජාතික	1	සහිතව	1		
		ජීවන	1	නිවැරදි	1		
		ජොෂ්ඨ	1	බවට	1		
		ජොෂ්ඨතාව	1	සහතික	1		
		කුසලතාව	1	නිවැරදිව	1		
		පන්තිය	1	සම්පූර්ණ	1		
		තනතුරු	1	(ඇමුණුමක්)	1		
		තාමය	1	නිසි	1		
		තාක්ෂණ	1	ඉල්ලුම්	1		
		ත්‍රිවිධ	1	අවසන්	1		
		හමුදා	1	ස්ථිර	1		
		සාමාජිකයන්	1	පත්වීම	1		
		දන්වා	1	ලිපි	1		
		එවීම	1	පරිපාලනමය	1		
		දිස්ත්‍රික්	1	අවශ්‍යතා	1		

		දීප	1	පරිබාහිර	1		
		ව්‍යාප්ත	1	වැටුපක්	1		
		දේශීය	1	පරිවාස	1		
		ආදායම්	1	ග්‍රන්�ථ	1		
		ද්විතීයන	1	පත්‍රිකා	1		
		පදනම	1	පහත	1		
		නාමිකව	1	කරනු	1		
		පත්කිරීම	1	ලැබේ	1		
		නිත්‍ය	1	පූර්ණ	1		
		නියෝගයකින්	1	පූර්ව	1		
		තීරණයකින්	1	අනුමැතින්	1		
		නිරීක්ෂණ	1	මත	1		
		නිර්දේශ	1	පොදු	1		
		වෙබ්	1	පෞද්ගලික	1		
		අඩවිය	1	කාර්ය	1		
		දිනයක්	1	මණ්ඩලයන්	1		
		ප්‍රකාශයට	1	ප්‍රමාණවත්	1		
		සම්පූර්ණ	1	ප්‍රාථමික	1		
		තොරතුරු	1	ද්විතිය	1		
		නිශ්චිත	1	ගැනීමේ	1		
		නිසි	1	පරිපාටි	1		
		බලය	1	බෝ	1		
		කැඳවනවා	1	පාලනය	1		
		කෙටුම්පත්	1	මාණ්ඩලික	1		
		සම්පාදක	1	මාසික	1		
		පත්වීම්	1	ආපසු	1		
		පරිපූරක	1	මුලායම	1		
		පර්යේෂණ	1	ප්‍රතිලාභ	1		
		නිලධාරිනී	1	මුළු	1		
		පශ්චාත්	1	යටිතල	1		
		ප්‍රධාන	1	පහසුකම්	1		
		පාර්ලිමේන්තු	1	සංවර්ධනය	1		
		පැමිණිල්ල	1	අත්‍යවශ්‍ය	1		
		මෙහෙයවනවා	1	සේවාවන්	1		
		පැමිණීමේ	1	කාර්යාල	1		
		පිළිගත්	1	ස්ථානය	1		
		විශ්වවිද්‍යාලයක්	1	නිවාඩු	1		
		පුරවාසන්න	1	දුම්රිය	1		
		වසර	1	බලපත්‍ර	1		
		කොස්තාපල්	1	ප්‍රතිපත්ති	1		
		සැරයන්	1	ප්‍රතිපත්තිය	1		
		නියාමන	1	ප්‍රවීණතාවය	1		
		සේවාවන්	1	රාමලාන්	1		
		කොටියසය	1	(රාමසාන්)	1		
		ප්‍රියමනාප	1	සමය	1		
		බඳවා	1	වසංගත	1		

		ගැනීමේ	1	මර්ධන	1	
		මධ්‍යම	1	වැඩසටහන්	1	
		රජය	1	වාර්ෂික	1	
		සම්බන්ධතා	1	සඳහා	1	
		මහා	1	අනිවාර්යයෙන්	1	
		භාණ්ඩාගාරය	1	වන	1	
		මාධ්‍ය	1	නිලධාරීන්	1	
		මෝටර්	1	ක්ෂේත්‍රය	1	
		රථ	1	සහ	1	
		ප්‍රවාහන	1	තලය	1	
		යටිතල	1	පරිමාණය	1	
		පහසුකම්	1	පරිවර්තනය	1	
		යාඥා	1	පියවර	1	
		ආගමික	1	පිළිබඳ	1	
		වතාවත්වල	1	විස්තර	1	
		රජයේ	1	ප්‍රතිපූරණය	1	
		අධිනීතිඥ	1	වර්ධක	1	
		රාජකාරි	1	උපයාගැනීම	1	
		සඳහා	1	වර්ධකය	1	
		සංස්ථා	1	ව්‍යුහය	1	
		ව්‍යවස්ථාපිත	1	පිළිවෙලක්	1	
		මණ්ඩල	1	මුර	1	
		රැකවරණ	1	වැඩිදුර	1	
		නිලධරයන්	1	වැඩිපුර	1	
		රැකියාව	1	ගෙවීම	1	
		රැස්වීම්	1	විධිමත්	1	
		වාරය	1	විසඳුම්	1	
		රේබිය	1	ලබාදීම	1	
		ලකුණු	1	විධිවිධාන	1	
		ලිබිත	1	අනුව	1	
		එකහත්‍ය	1	ක්‍රියාමාර්ග	1	
		වගන්තියෙන්	1	නියෝගය	1	
		ආවරණය	1	හේතු	1	
		වී	1	විශේෂිත	1	
		නැති	1	විශ්‍රාම	1	
		වගබලා	1	සමස්ත	1	
		වනජීවී	1	ගුණාත්මක භාවය	1	
		සංරක්ෂණ	1	සහන	1	
		වාර්ෂික	1	කාල	1	
		වාස්තු	1	පරාසයන්	1	
		විද්‍යාඥ	1	සීමාවන	1	
		රහිත	1	සේ	1	
		වැන්දඹු/	1	පවත්වාගෙන	1	
		වැන්දඹු	1	යාම	1	
		අනත්දරු	1	සීමිත	1	
		විකිරණ	1	සුදුසු	1	

		විදේශ	1	සේවයෙන්	1	
		විපක්ෂ	1	මුදා	1	
		නායක	1	හරිනවා	1	
		විරුද්ධව	1	යොදවා	1	
		කරුණු	1	බලාත්මක	1	
		දැක්වීම	1	කියා	1	
		විවාහක/	1	පටිපාටීන්	1	
		අවිවාහක	1	හැකි	1	
		දින	1	සෑම	1	
		විෂයභාර	1	උත්සාහය ක්ම	1	
		සුදුසුකම්	1			
		ශාස්තුපති	1			
		ගුරු	1			
		විදුහල්පති	1			
		සංයුක්ත	1			
		සක්‍රිය	1			
		සතුටුදායක	1			
		සත්‍ය	1			
		සමාජ	1			
		විරෝධි	1			
		ක්‍රියා	1			
		නිරාවරණ	1			
		වැඩසටහන්	1			
		සම්බන්ධිකරණ	1			
		සම්මන්තුණ	1			
		ශාලාව	1			
		සහතික	1			
		කළ	1			
		පිටපත්	1			
		සානුකම්පිතව	1			
		සලකා	1			
		බලා	1			
		තත්වය	1			
		සාවද්‍ය	1			
		අසම්පූර්ණ	1			
		සීමා	1			
		නිර්ණය	1			
		සුහද	1			
		හැර	1			
		යාම	1			
		සේවයෙන්	1			
		මුදා	1			
		නොහැරීම	1			
		සම්පූර්ණ	1			
		ගණය	1			
		ව්‍යවස්ථා	1			

		ස්ථානයට	1				
		සොලිසිටර්	1				
		ස්ත්‍රී/	1				
		භාවය	1				
		වකුය	1				
		දැන්වීම්	1				
		නිකුත්	1				
		ස්ථිර	1				
		ස්වදේශ	1				
		හෙද	1				

According to Table 3, the word frequency of idiomatic collocations is equal. The highest word frequency of fixed collocations is 13, which indicates that the constituent '*සේවා*' (services) was collocated with most of the collocations under this type. In addition, the constituent '*සේවය*' which carries a similar meaning, "service", is the singular form of the noun '*සේවා*' and represents the second highest frequency in fixed collocations. In the linguistic collocation type named strong collocations, the verb '*කිරීම*' was a constituent of 13 collocations and therefore, it appears as having the highest frequency under this category. However, the constituent '*කිරීම (do)* often combines with another word as it is generally used to form verbs in Sinhala language. The highest word frequency of loose collocations is 4 of the constituent '*කරනවා*' and the constituent '*අදියර*' represents the minimum value of word frequency as 1 in loose collocations.

High-frequency words: The words that are most frequently used or appear in a language are defined as high-frequency words. The words 'and', 'or', 'the', and 'be' verbs are examples of high-frequency words in the English language. Words such as '*හා/ සහ* (and), '*හෝ*' (or), '*කරනවා/ කිරීම*' (do), '*වීම*' (happen/do), '*දීම*' (give) and '*ගැනීම*' (give) are a few examples for high-frequency words in the Sinhala language. '*කරනවා/ කිරීම*', '*වීම*', '*දීම*' and '*ගැනීම*' are certain words that usually co-occur with a verb to complete the idea indicated by a verb. This is a significant and unique feature of the Sinhala language compared to the English language. Such high-frequency words make a text easily understood by a reader.

Low-frequency words: Although high-frequency words appear in a text or a corpus frequently, low-frequency words rarely appear in a text or a corpus. Subject-specific terms and technical terms are such words that are not most frequently used in general contexts. Accordingly, Public Administration Circulars also include low-frequency words such as 'reimburse' and 'regulatory'. '*ප්‍රතිපූර්ණය කිරීම*', '*නියාමන*', '*සංශෝධනය*' and '*කෙටුම්පත*' are examples abstracted from the collocations selected from Public Administration Circulars for low-frequency words in Sinhala as such words are included in formal or legal documents according to the relevant contents. Although such low-frequency words are used in specific contexts based on the requirement, using low-frequency words in general contexts may make it difficult for the target audience to understand a text.

Finally, the results mentioned above show that Public Administration Circulars in Sinhala include lexical features such as word length, word frequency, high-frequency words, and low-frequency words. They are beneficial in identifying the main ideas, specific focus, complexity, or difficulty of a text or a corpus and preserving the formality of texts. Additionally, fixed collocations and strong collocations are the types of linguistic collocations that are most frequently used in the selected Public Administration Circulars, and fully fixed collocations or idiomatic collocations are occasionally found. However, the results may vary based on the number of collected data.

CONCLUSION

This study was conducted with the aim of identifying the lexical features of the types of linguistic collocations used in Public Administration Circulars in Sinhala. Based on the discussion in the previous chapter, there are some points that can be drawn from this research. Collocations are described as a set of words that are always bound and co-occur together. The results show that the collocations collected from the 12 Public Administration Circulars can be categorized under the four types of linguistic collocations introduced by Li et al. (2005). The four types of linguistic collocations are idiomatic (fully fixed) collocations, fixed collocations, strong collocations, and loose collocations. Among the said types of linguistic collocations, there are three types, namely idiomatic (fully fixed) collocations, strong collocations, and fixed collocations, that depend on restrictions or limitations whereas the other type, loose collocations, are not formulated based on such limitations.

The fully fixed or idiomatic collocation type has the highest number of limitations when compared to the other types. Individual constituents or the parts of the collocation cannot determine the implied meaning of fully fixed collocations and such collocations cannot be replaced or substituted as they are immutable collocations. Fixed collocations are formed with rigid combinations and one constituent of the collocation often follows the other constituent naturally. Native speakers or second-language speakers of any language use this type of collocation unintentionally due to the fixed form they have. In comparison to the two types of linguistic collocations mentioned above, strong collocations can be modified or one constituent can be substituted by another word with a similar meaning. However, such collocations have the limitation that the order of this type of collocation is unchangeable. The linguistic collocation type, loose collocations, is allowed to be modified, replaced, or substituted and the order can also be changed. However, a minimum level of restrictions still remains.

According to the results of the study, the most common type of linguistic collocations in the Public Administration Circulars in Sinhala are fixed collocations (52.6%), and strong

collocations account for the second largest linguistic collocation type, 40.3%, out of the total. Compared to the above two types of linguistic collocations, loose collocations account for 6.4% of the total number of collected data and fully fixed collocations are the least frequently used linguistic collocation type. Studying the lexical features of each of these linguistic collocation types, word length, word frequency, high-frequency words and low-frequency words were identified. As administrative texts such as Public Administrative Circulars use subject-specific terms and longer words, the word length of each type of linguistic collocations varies from one to another. When calculating the word frequencies, it was identified that, unlike in the English language, the verbs in the Sinhala language are always formed with combinations of words such as '*කිරීම*' (to do), '*දීම*' (to give), '*වීම*'(happen), '*ගැනීම*' (take). The usage of such words varies according to the meaning that needs to be implied in a particular context. This feature of the Sinhala language indirectly affects lexical features such as word length and word frequency. Apart from that, lexical features such as high-frequency words and low-frequency words do not find any exceptions in the English and Sinhala languages.

In addition, the findings of the study indicated that certain collocations in Sinhala are not referred to as collocations in English. Similarly, the structures or the order of collocations may differ between the two languages.

Depending on the findings of the research, the following recommendations are often effective:

- It would be preferable to teach collocations to students who are interested in Sinhala, as a separate part.

- The institutions and universities related to translation studies in Sri Lanka should encourage their undergraduates or learners to produce more studies, and research addressing the sector of translating administrative texts.

- It is further recommended to use collocations and to be familiar with the types of collocations as much as possible in order to enhance writing skills while preserving the complexity and formality of administrative texts.

The term 'collocation' is a significant linguistic phenomenon in any language and there are only a limited number of studies conducted regarding this concept. The immense importance carried by collocations in a language has not yet been identified and motivated. The present study focused only on identifying the lexical features of different types of linguistic collocations. The four linguistic collocation types introduced by Wanyin Li et al. (2005) were the major concern of this study and the characteristics and taxonomies of collocations are still under debate. Hence, new research can be conducted in the future regarding the characteristics and taxonomies of collocations relevant to the Sri Lankan context. Additionally, researchers who are interested in conducting studies based on collocations can also conduct studies based on several other objectives. In the future, studying the challenges or problems encountered while translating collocations related to Sinhala and English, as well as the translation strategies used in translating collocations in Sinhala to English, will be more beneficial.

REFERENCES

Abdelhamid, I. A. (2021). Translating Collocations from English into Arabic and vice versa: An Empirical Study. International Journal of Linguistics and Translation Studies, 2(3). 1-11

Alshami, I. (2019). Language and linguistics. Retrieved from https://www.researchgate.net/publication/331454003_Language_and_linguistics

Baker, M. (1992). In Other Words: A coursebook on translation. Routledge. Retrieved from https://www.academia.edu/30517772/Mona_Baker_In_Other_Words_Coursebook?email_work_card=title

Catford, J. (1965). A Linguistic Theory of Translation. Oxford University Press.

Chomsky, N. (2000). The Architecture of Language. Oxford University Press.

Karjo C.H., & Metta, E. (2019). The Translation of Lexical Collocations in Undergraduate Students' Theses' Abstract: Students versus Google Translate. Lingua Cultura, 289-296. doi:10.21512/lc.v13i4.6067

Handke, J. (2014). The Virtual Linguistics Campus. July 31,2014. *What is Linguistics (not)?* [Video] YouTube. https://www.youtube.com/watch?v=bzz1pFWAtMo&t=6s

Stulpinaitė. M., Horbačauskienė. J. & Kasperavičienė. R. (2016). Issues in Translation of Linguistic Collocations. Research Journal: Studies about Languages, 31-41.

Anagnostou, N.K., & Weir, G.R. (2011). From corpus-based collocation frequencies to readability measure. Retrieved from https://www.semanticscholar.org/paper/From-corpus-based-collocation-frequencies-to-Anagnostou-Weir/be513abe8f3aef31c372af18505c227db71984e0

Newmark, P. (1988). A Textbook of Translation. Prentice HaH International vUIO Ltd.

Silva, J. D. (2018). භාෂා පරිවර්තන ප්‍රවේශය (bhasha pariwarthana praweshaya). Colombo: S. Godage & Brothers (Pvt) Ltd.

Wanyin Li, Q. L. (2005, March). Similarity-Based Chinese Synonym Collocation Extraction. Computational Linguistics and Chinese Language Processing, 10, 123-144.