

**The Semantic Prosody of Natural Phenomena in the Qur'an: A Corpus-
Based Study**

Hala Jamal Ali Alshahrani

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Languages, Cultures and Societies

January 2020

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Hala Jamal Ali Alshahrani to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

© 2020 The University of Leeds and Hala Jamal Ali Alshahrani

Dedication

To my loving husband, Turki, for supporting me to pursue my dream.

*To my beautiful children: Dana, Thamer, and Nawaf, who have been there for me
through the best and worst of times.*

Acknowledgements

First and foremost, I would like to thank Allah Almighty for giving me the strength, knowledge, ability and opportunity to undertake and accomplish this work. Without His blessings, this achievement would not have been possible.

Throughout the writing of this dissertation, I have received a great deal of support and assistance. I would like to express my sincere gratitude to Dr Claire Brierley and Professor James Dickins, my research supervisors, for their patient guidance and enthusiastic encouragement. Their expertise was invaluable in the formulating of both the topic and methodology of this interdisciplinary research.

I would like to further express my deepest gratitude and appreciation to Dr Claire for the very useful Python tutorials, which provided me with the tools that I needed to choose the right direction and complete my dissertation. Your support and useful critiques throughout the development of every single aspect of this work have been very much appreciated.

I am thankful to Mrs Karen Priestley, Doctoral College Officer, who has offered her generous support and kind guidance throughout the PhD journey. She has greatly facilitated the administrative work related to this research.

I would like to thank my parents for their unconditional love, wise counsel, and sympathetic ear. You are always there for me.

Finally, there are my friends, who were of great support in providing a happy distraction to rest my mind outside of my research. I would like to acknowledge my friends, Drs: Jawhara, Khuzami, Amira, and Abeer. You supported me greatly and were always willing to help me.

Abstract

This thesis explores the Semantic Prosody (SP) of natural phenomena in the Qur'an and five of its prominent English translations [Pickthall (1930), Yusuf Ali (1939/ revised edition 1987), Arberry (1957), Saheeh International (1997), and Abdel Haleem (2004)]. SP, scarcely explored in Qur'anic research, is defined as 'a form of meaning established through the proximity of a consistent series of collocates' (Louw 2000, p.50). Theoretically, it is both an *evaluative prosody* (i.e., lexical items collocating with semantic word classes that are positive, negative, or neutral) and a *discourse prosody* (i.e., having a communicative purpose).

Given the stylistic uniqueness of the Qur'an and considering that SP can be examined empirically via corpora, the present study explores the SP of 154 words associated with nature referenced throughout the Qur'an using Corpus Linguistics techniques. Firstly, the Python-based Natural Language Toolkit was used for the following: to define nature terms via WordNet; to disambiguate their variant forms with Stemmers, and to compute their frequencies. Once frequencies were found, a quantitative analysis using Evert's (2008) five-step statistical analysis was implemented on the 30 most frequent terms to investigate their collocations and SPs. Following this, a qualitative analysis was conducted as per the Extended Lexical Unit via concordance to analyse collocations and the Lexical-Functional Grammar to find the variation of meanings produced by lexico-grammatical patterns. Finally, the resulting datasets were aligned to evaluate their congruency with the Qur'an.

Findings of this research confirm that words referring to nature in the Qur'an do have semantic prosody. For example, astronomical bodies are primed to occur in predominantly positive collocations referring to glorifying God, while weather phenomena in negative ones refer to Day of Judgment calamities. In addition, results show that Abdel-Haleem's translation can be considered the most congruent.

This research develops an approach to explore themes (e.g., nature) via SP analysis in texts and their translations and provides several linguistic resources that can be used for future corpus-based studies on the language of the Qur'an.

Table of Contents

Dedication	iii
Acknowledgements	v
Abstract.....	vii
Table of Contents	ix
List of Tables	xiii
List of Figures.....	xv
List of Abbreviations	xix
The Arabic Transliteration System.....	xxi
Chapter 1 Introduction.....	- 1 -
1.1 Aims of the study.....	- 2 -
1.2 Motivation.....	- 2 -
1.3 Research questions.....	- 3 -
1.4 Research contributions and limitations.....	- 4 -
1.4.1 Expected contributions	- 4 -
1.4.2 Limitations.....	- 5 -
1.5 Introducing the theoretical framework	- 6 -
1.6 Natural language processing, programmes and data analysis tools.....	- 7 -
1.6.1 Natural language processing with Python	- 8 -
1.6.2 Linguistic resources and computational tools	- 10 -
1.6.3 Data analysis and visualisation tools in the qualitative analysis ..	- 12 -
1.7 Introducing the qualitative analyses of textual data.....	- 14 -
1.8 Introducing the criteria for evaluation of the translations.....	- 17 -
1.9 Introducing the research corpora	- 20 -
1.9.1 The Holy Qur'an	- 20 -
1.9.2 M. M. Pickthall (London, 1930).....	- 21 -
1.9.3 Abdullah Yusuf Ali (Lahore, 1939)	- 22 -
1.9.4 A. J. Arberry (London, 1957).....	- 22 -
1.9.5 Saheeh International (Jeddah, 1997)	- 23 -
1.9.6 M. A. S. Abdel Haleem (Oxford, 2004).....	- 23 -
1.10 Overview of this thesis	- 24 -

Chapter 2 Natural Phenomena in the Qur'an: A Descriptive Overview with NLP Applications.....- 25 -

Introduction	- 25 -
2.1 Compiling the list of nature terms and their contextual meanings	- 28 -
2.1.1 An overview of the available lists of natural phenomena	- 28 -
2.1.2 A reading of nature in the Qur'an	- 31 -
2.1.3 The contextual meanings of nature as a theme in the Qur'an	- 33 -
2.1.4 The compiled working list of natural phenomena in the Qur'an	- 40 -
2.2 The ontology of nature terms	- 41 -
2.2.1 The available ontologies of concepts in the Qur'an	- 42 -
2.2.2 The proposed conceptual ontology of nature terms	- 45 -
2.3 The disambiguation of nature terms	- 52 -
2.3.1 Stem-based disambiguation.....	- 54 -
2.3.2 Word sense disambiguation.....	- 57 -
2.4 A preliminary statistical exploration of nature terms in the Qur'an	- 60 -
2.4.1 The frequency distribution for Ali's translation as a bag of words and the central theme of the Qur'an	- 60 -
2.4.2 The frequency distribution for the list of natural phenomena in the divisions of Ali's translation	- 63 -
2.4.3 A collocation extraction of nature terms in Ali's translation	- 66 -
2.4.4 The 30 most frequent natural phenomena terms in the Arabic Qur'an	- 68 -
2.5 Conclusion	- 68 -

Chapter 3 Literature Review:.....- 71 -

Determining the Theoretical Framework and Methodology

Introduction	- 71 -
3.1 Defining 'corpus linguistics', 'collocation', 'lexico-grammar', and 'SP.'	- 72 -
3.1.1 Collocation: "knowing a word by the company it keeps"	- 78 -
3.1.2 Lexico-grammar and collocation.....	- 95 -
3.1.3 Semantic prosody: a collocational phenomenon and cohesive device	- 100 -
3.2 Corpus linguistic studies on the Qur'an.....	- 113 -
3.2.1 Corpus linguistic studies on nature in the Qur'an	- 113 -
3.2.2 Corpus linguistic studies on SP in the Qur'an.....	- 115 -
3.2.3 NLP projects and the Qur'an.....	- 119 -
3.3 The evaluation of translation	- 122 -

3.4	Conclusion	128 -
Chapter 4 Methodology		129 -
	Introduction	129 -
4.1	Part 1: The corpus-based analysis of SP of natural phenomena	131 -
4.1.1	Finding SP: a corpus-based method and a computational approach ..	132 -
4.1.2	Text pre-processing	133 -
4.1.3	Statistical analysis	141 -
4.1.4	Finding the SP of nature in the Qur'an.....	148 -
4.1.5	Data analysis in this research	153 -
4.2	Part 2: The evaluation of the Qur'an translations via SP of nature terms-	163 -
4.2.1	Why semantic prosody?	164 -
4.2.2	The evaluation of the five translations	165 -
4.3	Conclusion	170 -
Chapter 5 Results and Discussion		171 -
	Introduction	171 -
5.1	Results of exploring the SP of natural phenomena in the Qur'an	172 -
5.1.1	Frequencies of natural phenomena in the Qur'an	172 -
5.1.2	Collocations of natural phenomena in the Qur'an.....	174 -
5.1.3	Lexico-Grammatical (LG) patterns of natural phenomena in the Qur'an	- 179 -
5.1.4	The SP of natural phenomena in the Qur'an	182 -
5.2	Results of the evaluation of the translations	190 -
5.3	Discussion of the results	196 -
5.3.1	Is there SP in the representation of nature in the Qur'an?.....	197 -
5.3.2	What are the lexico-grammatical patterns, evaluative prosodies, and	discourse prosodies of nature in the Qur'an?
5.3.3	Which of the five translations is most congruent, and which is the most	divergent from the representation of nature in the Qur'an?
5.3.4	How can variances in the representation of SP of nature in the English	renderings of the Qur'an be explained in terms of consistency and
	accuracy?	207 -
5.4	Summary	214 -
Chapter 6 Conclusion		216 -
6.1	Findings and implications	216 -
6.2	Challenges and reflections	219 -

6.3 Suggestions for future studies.....	219 -
Bibliography	221 -
Appendix A: The Arabic Qur'an List of Stop-Words.....	249 -
Appendix B: The English List of Stop-Words in the Translations of the Qur'an-	279 -
Appendix C: The Lexico-grammatical Patterns of Nature in the Qur'an	285
Appendix D: The Lexico-grammatical Patterns of Nature and Their Discourse	
SPs in English Translations	293
Appendix E: The Evaluative SP of Nature in the Qur'an Translations	329
Appendix F: A Diachronic Analysis of English Stop-words in the English	
translations of the Qur'an (1930-2014).....	335
Appendix G: Examples of the Five Most Frequent Discourse Prosodies in the	
Qur'an	343

List of Tables

Table 1: Examples of language processing tasks and corresponding NLTK modules with examples of functionality (table taken from Bird et al., 2009, xiv).....	9 -
Table 2: Types of textual data and data analyses in this research (classification based on Stubbs, 2001, pp. 66-7, <i>my hyphenations</i>).....	14 -
Table 3: Three approaches to content analysis (table taken from Hsieh and Shannon 2005, p.1286).....	15 -
Table 4: Yahya's natural phenomena in other chapters of his book and examples from the Qur'an (table taken from Yaya, 2003, pp.224-28).....	30 -
Table 5: Raw frequencies of الطبيعة <i>tabi'i</i> or words related to the natural phenomenon (table taken from Badawi, 2012, p.561)	31 -
Table 6: Nature in the most frequently repeated verses of the Qur'an	32 -
Table 7: The list of natural phenomena in the Qur'an and their semantic categories-	48 -
Table 8: The difference in the 20 most frequent nature terms between Makkan and Medinan surahs	65 -
Table 9: A sample of trigrams of natural phenomena terms in Ali's translation..	67 -
Table 10: The 30 most frequent natural phenomena terms in the Qur'an	68 -
Table 11: A contingency table for the observed frequencies (table taken from Brezina, 2018, p.70 and based on the notation in Evert 2008)	86 -
Table 12: A contingency table for the expected frequencies (table taken from Brezina, 2018, p.71 and based on the notation in Evert 2008)	86 -
Table 13: An example of a collocation parameters notation (CPN) (table adopted from Brezina et al., 2015, p.157; Brezina, 2018, p.75)	88 -
Table 14: English lexico-grammatical patterns across collocations (table taken from Benson et al., 1986, p.160)	97 -
Table 15: Theories of lexis and grammar to interpret collocations	109 -
Table 16: Summary of preferred translations for verb-preposition constructions under study (Younis, 2018, p. 140).....	127 -
Table 17: Compiling the Qur'an stop-list	136 -
Table 18: The word-level root-disambiguation of the word <i>day</i> in Arabic	139 -
Table 19: Examples of the synonym level root-disambiguation of the words ...	140 -
Table 20: The root-based disambiguation of natural phenomena terms.....	140 -
Table 21: A sample of the syntactic categorisation of collocates of natural phenomena	149 -
Table 22: The Python and LncsBox LLR scores of bigrams of the word 'earth' in Ali's translation.....	156 -

Table 23: The collocation parameters notation (CPN) of natural phenomena in the Quran	157 -
Table 24: Statistical overview of datasets in this research	165 -
Table 25: The translations of nature terms in the five selected translations	165 -
Table 26: The frequency counts of nature terms in the Qur'an and its translations-	167 -
Table 27: The 30 most frequent natural phenomena in the Qur'an	172 -
Table 28: Example 1- the collocations of the node word شمس <i>šams</i> 'sun' in the Qur'an	174 -
Table 29: Example 1- Verses with collocational co-occurrences of the node word شمس <i>šams</i> 'sun' in the Qur'an and their English interpretation	175 -
Table 30: Example 2- the collocations of the node word ماء <i>mā</i> 'water' in the Qur'an	175 -
Table 31: Example 2- the collocational co-occurrences of the node word ماء <i>mā</i> 'water' in the Qur'an and their English Interpretation	177 -
Table 32: LG pragmatic functions of natural phenomena and their relevant collocates	179 -
Table 33: Samples of the LG patterns of nature in the Qur'an and their functions-	181 -
Table 34: The evaluative prosodies (EP) and discourse prosodies (DP) of nature in the Qur'an	182 -
Table 35: The evaluative prosody of the words ريح <i>ryḥ</i> 'wind' and رياح <i>ryaḥ</i> 'winds' in the Qur'an	185 -
Table 36: The comparison of the ratios of the most frequent natural phenomena in the Qur'an	190 -
Table 37: The frequency distribution of the word 'fire' in the Qur'an and its translations	192 -
Table 38: The statistical representation of the lexico-grammar of nature in the Qur'an in comparison with its translation	195 -
Table 39: The comparison of congruency of evaluative SP (as in Ebeling, 2014)-	196 -
Table 40: The comparison of congruency of discourse SP	196 -
Table 41: Examples of translations with lexical gaps	208 -
Table 42: Examples of the inconsistency of terminology in the English translation of nature terms in the Qur'an	212 -

List of Figures

Figure 1: A model for the theoretical framework	6 -
Figure 2: The download page for NLTK	9 -
Figure 3: Concordance lines for <i>catch fire</i> (figure taken from Kilgarriff et al., 2014)	13 -
Figure 4: The data analyses of SP of nature in the Qur'an and its translations	14 -
Figure 5: The flow of tasks of producing a list of nature terms in the Qur'an	27 -
Figure 6: The theme of nature in the Qur'an	33 -
Figure 7: The Quranic Ontology. Each node in the graph represents a unique concept in the Quran (Figure taken from: http://corpus.quran.com/ontology)	44 -
Figure 8: The ontology of concepts in the Qur'an.....	46 -
Figure 9: The concepts of nature in the Qur'an	46 -
Figure 10: An elaborated ontology of nature in the Qur'an.....	47 -
Figure 11 : A Comparison between raw frequencies of natural phenomena in the Arabic text and Ali's translation.....	52 -
Figure 12: A comparison between frequency results after stemming (lemmatization)- 53 -	
Figure 13: The rule-based stemming model	55 -
Figure 14: The pre-processing model	56 -
Figure 15: The 30 most frequent words in Ali's translation of the Qur'an	62 -
Figure 16: The types of the phraseme (figure based on the classification of Mel'Čuk, 2012, p.32, pp. 31-56).....	79 -
Figure 17: The Extended Lexical Unit (ELU)	94 -
Figure 18: Halliday and Matthiessen's lexico-grammar cline (2004, p.43)	96 -
Figure 19: The defining features of SP in this research.....	104 -
Figure 20: Class hierarchy in the Qur'anic nature ontology (figure taken from Sadi et al., 2016)	114 -
Figure 21: Triple in verse (2:50) (figure taken from Sadi et al., 2016)	115 -
Figure 22: The SP of the word مُبِين <i>mubeen</i> 'clear' in the Qur'an (figure taken from Al- Nasser and Khashan 2008, p.11)	116 -
Figure 23: An ontology graph based on the chapter 'Joseph' from the Qur'an .-	120 -
Figure 24: The three corpus linguistic features of the methodology	130 -
Figure 25: Finding the SP of natural phenomena in the Qur'an	131 -
Figure 26: A model for text pre-processing the research corpora	133 -
Figure 27: The modification of the available Python Arabic stop-words.....	136 -
Figure 28: Statistical analysis of collocations of nature in the Qur'an	142 -

Figure 29: Illustration of surface co-occurrence for the word pair (<i>hat, roll</i>) ...-	144 -
Figure 30: A concordance of the nature term <i>light</i> in the Qur'an.....-	150 -
Figure 31: The discourse prosodies of the word <i>light</i>-	151 -
Figure 32: The discourse prosodies of the word <i>water</i>-	152 -
Figure 33: A dataset of the pragmatic functions to find discourse SP in the Arabic text-	153 -
Figure 34: The data analysis of SP of natural phenomena in the Qur'an	154 -
Figure 35: The frequencies of natural phenomena nodes	155 -
Figure 36: A sample of the statistical representation of the collocations of nature in the Qur'an in accordance with Evert's model (2008).....-	156 -
Figure 37: Collocates of العالمين <i>al'ālamīn</i> , translated as 'beings' by Arberry in the Quran : 6a-LL (6.63), 5L-5R, C3-NC3; function words removed; ranked collocates considered	158 -
Figure 38: The LG patterns of natural phenomena in the Qur'an and its translations. -	158 -
Figure 39: The annotation of the evaluative prosodies of the natural phenomena-	160 -
Figure 40: An example of dual annotation of discourse prosodies.....-	160 -
Figure 41: The annotation of the discourse prosodies of nature in the Qur'an in Arabic-	161 -
Figure 42: Alignment of concordances of the word <i>moon</i> via LancsBox v 4.0 .-	169 -
Figure 43: The comparison of the SP of nature terms between the Qur'an and its translations	170 -
Figure 44: Flow of reporting results (repeated Figure 24).....-	171 -
Figure 45: The semantic categories of natural phenomena in the Qur'an	173 -
Figure 46: Collocates of شمس <i>šams</i> 'sun' in the Qur'an: 6a-LL (6.63), 5L-5R, C3- NC3; function words removed; ranked collocates considered	174 -
Figure 47: Collocates of ماء <i>mā'</i> 'water' in the Qur'an: 6a-LL (6.63), 5L-5R, C3- NC3; function words removed; ranked collocates considered	176 -
Figure 48: Collocates of مطر <i>maṭar</i> 'rain' in the Qur'an: 6a-LL (6.63), 5L-5R, C3- NC3; function words removed; ranked collocates considered	178 -
Figure 49: The comparison based on the relative frequencies of natural phenomena in the Arabic text.....-	191 -
Figure 50: The semantic categories of natural phenomena	193 -
Figure 51: The evaluative SP of nature in the Qur'an	198 -
Figure 52: The discourse prosodies of nature in the Qur'an.....-	198 -
Figure 53: The collocational network of discourse SP of nature nodes in the Qur'an .-	201 -

- Figure 54: The collocational network of discourse SP of nature nodes in Haleem's translation.....- 202 -
- Figure 55: A LancsBox portrait of the alignment of GraphColls of the word قمر *qmr* 'moon' in the Arabic and English corpora (Haleem's translation).....- 203 -
- Figure 56: The analysis of evaluative SP of natural phenomena in the Qur'an .- 204 -
- Figure 57: The SP congruency scores of the English translations.....- 207 -
- Figure 58: A comparison based on the LLR scores of the collocations of the word 'sun'- 210 -
- Figure 59: A comparison based on the LLR scores of the collocations of the word 'people'- 211 -

List of Abbreviations

BNC	British National Corpus
CA	Content Analysis
DP	Discourse Prosody
ELU	Extended Lexical Unit
EP	Evaluative Prosody
IR	Information Retrieval
LG	Lexico-Grammar
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
QA	Qur'an Analysis
QAC	Qur'anic Arabic Corpus
SP	Semantic Prosody
SPs	Semantic Prosodies
ST	Source Text
TT	Target Text
WSD	Word Sense Disambiguation

The Arabic Transliteration System¹

Arabic letter	ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ظ	ع	غ	ف	ق	ك	ل	م	ن	هـ	و	ي
Transliteration	or ā	b	t	ṭ	j	ḥ	x	d	ḍ	r	z	s	š	ṣ	ḍ	ṭ	ḏ	ġ	f	q	k	l	m	n	h	w or ū	y or ī

The *ḥarakāt*, *fathā*, *kasra* and *ḍamma* are transliterated as a, i, u. A *šadda* results in a geminate (consonant is written twice), except in the case of the article, which is written with ‘sun letters’ assimilated (aš-šams). An alif marking (a:) is transliterated as ā. *tā’ marbūṭa* (ة) as word-final -h or -t. *’alif maqṣūra* (ى) appears as ā, rendering it indistinguishable from *alif*. Long vowels (i:) and (u:) are transliterated as ī and ū. The Nisba suffix appears as -iyy-; the nunation is ignored in the article and prepositions.

¹ <http://quranic-research.net/transliteration/index.html> is a website used in transliterating the Arabic words whenever needed into English via this system.

Chapter 1 Introduction

This chapter provides an overview of this thesis, which explores the Semantic Prosody (SP) of natural phenomena in the Qur'an and employs the resulting datasets of this exploration to evaluate the English renderings of SP of nature in the following five translations of the Qur'an: Pickthall (1930), Yusuf Ali (1939/ revised edition 1987), Arberry (1957), Saheeh International (1997), and Abdel Haleem (2004) (Haleem henceforth). SP is defined as the spreading of connotative colouring (Partington 1998, p.65; 2004b, pp.131-32); it is an *evaluative prosody*, in that it refers to lexical items that collocate with semantic classes of words that are *positive*; *negative*, or *neutral*. In addition, it is a *discourse prosody*, in that it has the communicative purpose of revealing the speaker's or writer's attitudes. Moreover, although it is commonly explored as a collocational phenomenon (as in Louw 2000, p.50) and as an aspect of meaning that can be investigated using corpora (as in Partington 1998), there has been very little and limited work done on collocation and SP in the Qur'an (e.g., Al-Nasser and Khashan, 2008; Al-Ubaidi, 2013; Al-Sofi et al., 2014; and Younis 2018).

Building on previous research (Chapter Three), the present study aims to shed light on the collocational behaviour of words in the Qur'an, with specific attention to exploring nature as a Qur'anic theme (Fazlur Rahman, 2009). Using Corpus linguistic techniques, this research examines collocations² describing words referring to natural phenomena, such as *earth*, *day*, *sky*, *garden*, and *mountain* as they appear in the Qur'an (Chapter Four). It relies on the quantitative and qualitative data analyses of these collocations to reveal the SP of natural phenomena in two ways: as a collocational phenomenon, and as a cohesive device that endows the text with consistency and harmony (Morley and Partington, 2009, p.139). In this sense, this thesis presents a novel approach to analyse a specific theme by statistically exploring collocations of the theme's relevant concepts, such as the ones in the Qur'anic Ontology of Concepts;³ and focuses on yielding up patterns of SP in the context of references to nature in the Qur'an. In addition, it explores collocations to unveil the communicative purpose of the SPs of words referring to natural phenomena as they appear

2 Among several definitions, *collocation* is defined as a phenomenon which describes words that tend to occur in proximity (co + location) to one another because they have "affinity" to, or "affiliation with", one another (See also Weisser, 2016, p. 198).

3 The *Qur'anic Ontology of Concepts* uses knowledge representation to define the key concepts in the Qur'an and shows the relationships between these concepts using predicate logic; that is, the type of relationships between concepts is 'is-a'. Available from: [<http://corpus.quran.com/ontology.jsp>], [Accessed 15 December 2016].

in the Qur'an through a detailed analysis of the co-occurrences of natural phenomena as integral constituents of the lexical unit (See Sinclair's theory of the Extended Lexical Unit 2004a, p.141).⁴ In the context of the aforementioned approach, the analysis of SP is further extended to include the five English translations of the Qur'an (Chapter Four). Furthermore, the resulting datasets of this exploration were analysed to indicate the most congruent of the five translations in its representation of the SP of nature in the Qur'an based on congruency scores of the two features of SP (the results of the evaluative and discourse prosodies are in Chapter Five).

Following this general description of the flow of tasks in this thesis, this introductory chapter will present the aims, motivation, research questions, and limitations as well as the expected contributions in Sections 1.1-1.4. It will then give a brief account of some of the components of this research to familiarise the reader with the theoretical framework and some parts of the methodology (i.e., the parallel corpora⁵ and the implemented methods of qualitative data analysis) (Sections 1.5-1.9). Finally, the structure of the thesis will be outlined in detail in the last section of this chapter (Section 1.10).

1.1 Aims of the study

This corpus-based study has both a primary aim and a secondary aim, which can be summarised as follows:

- The primary aim is to explore the semantic prosody of nature terms in the Qur'an and the communicative and pragmatic functions of such a prosody.⁶
- The secondary aim is to evaluate five English translations of the Qur'an using Arabic-English congruency in the semantic prosody of nature terms as a discriminating factor.

1.2 Motivation

To the researcher's knowledge, there has been a limited number of studies on SP in the Qur'an, despite its importance in understanding the meanings conveyed in the Qur'an and

⁴ Sinclair defines *collocation* in the light of his lexical model, which states that a given lexical item in the Extended Lexical Unit (ELU) is characterised in terms of "five categories of co-selection", two of which are obligatory: the core and the semantic prosody (2004a, p.141). See also Section 3.1.1.

⁵ *Parallel corpora* can be bilingual or multilingual, i.e., source text in language A and translations in languages B, C, D, etc. (Olohan, 2004, p.25). They are also defined as: the original text and its translation aligned sentence by sentence (Teubert, 1996, p.249).

⁶ *Semantic prosody* is said to express the "attitudinal or pragmatic function of an item [i.e., its communicative purpose]; without it, the string of words is not put to use in a viable communication" (Sinclair, 1996, pp.87-8).

in maintaining the accuracy of lexical choice in its translation. Therefore, the researcher's motivation was to provide evidence of the importance of SP as a linguistic phenomenon in the Qur'an. Moreover, although the theme of nature has been commented on in several previous studies (e.g., Abdul Wadud, 1971;1996; 1998; Bell, 1987; Bell et al., 1991; Robinson, 1999; Mohamed, 2014), none has provided insights into the collocational behaviour of words in the Qur'an that describe natural phenomena. In addition, it was found that one of the latest trends of SP study is to explore it as "a tool for accuracy" (e.g., Ebeling, 2014, p.161; Younis 2018, p.120) and for comparison between texts in cross-linguistic and translation studies.⁷ Hence, it was deemed feasible to expand this corpus-based research to include an aspect of descriptive translation studies.⁸ Accordingly, five translations of the Qur'an were chosen to be examined for congruency in the representation of the theme of nature via the exploration of the SP of this theme.

Furthermore, another popular trend in corpus-based studies is to employ a computational method in a mixed approach, which produces quantitative and qualitative analyses (as described in Adolphs and Carter, 2002, p.7; Hunston, 2002, p.249; and Baker, 2016, p.139). Thus, this trend in interdisciplinary research of corpus linguistics that incorporates computational and statistical data analysis tools was another motivation for the researcher to explore textual data using Corpus Linguistic techniques coded in Python and other data analysis tools. This study will be pioneering research on SP in the Qur'an, given the limited number of previous studies on the subject. Finally, the researcher, whose background is in Applied Linguistics, has always been fascinated with the elegance and eloquence of the language of the Qur'an. Therefore, it was an excellent opportunity for her to experience first-hand the grandeur of expression in the Qur'an.

1.3 Research questions

This research focuses on four questions that derive from the aims of the study. The first two research questions cover the primary aim of this thesis, which is to explore the SP of nature in the Qur'an; the third and fourth questions focus on the secondary aim which is to

⁷ *Accuracy* of translation can generally refer to being precise in rendering from one language into another, i.e., making the correct choices of equivalent terms in the target language text to convey the meaning of the source language text. See also (Baker, 1993, pp.233-52).

⁸ *Descriptive Translation Studies*, as the name suggests, is interested in describing translations and translation practice as it occurs or has occurred, the role and nature of translation and the impact of translation activity in wider cultural, social, historical contexts, etc. (Olohan, 2004, p.199).

evaluate five translations of the Qur'an based on the results of the SP analysis of each of the datasets in this research. They are the following:

- 1- Is there SP in the representation of nature in the Qur'an?
- 2- What are the lexico-grammatical patterns, evaluative prosodies, and discourse prosodies of nature in the Qur'an?
- 3- Which of the five translations is the most congruent and which is the most divergent from the representation of nature in the Qur'an?
- 4- How can variances of the representation of SP of nature in the English translations of the Qur'an be justified in terms of consistency and accuracy?

1.4 Research contributions and limitations

This section presents the expected contributions of this research and then highlights the limitations of this corpus-based study.

1.4.1 Expected contributions

The impact of this research can be said to benefit three disciplinary fields of study: corpus linguistics, Qur'anic studies, and translation studies. In addition, it is sub-divided into four domains of contribution: theory, methodology, insights, and language resources. On the theoretical level, it contributes to corpus linguistics, in that it establishes a theoretical framework encompassing three constituents (collocation, lexico-grammar, and SP). This combination of three related theories can be used for future corpus-based studies on SP.

Similarly, the methodology of this research, which demonstrated the usefulness of Corpus linguistic techniques for exploring concepts such as root-based disambiguation and statistical analysis to find collocations and SP, can be applied to other corpus-based studies on SP. In the same domain of contribution, this mixed-method approach used to examine the theme of nature in the Qur'an is a novel approach that can be utilised in Qur'anic studies to explore other prominent themes in the Qur'an.

Moreover, the findings of this research can be useful in providing insights into the importance of the analysis of SP for translation studies. It contributes by emphasising the significance of raising the translator's awareness of this linguistic phenomenon to achieve accuracy in translation. By alluding to the differences (i.e. congruency and divergence) between translators in the representation of the SP of nature as a Qur'anic theme, the translator can perceive the importance of presenting this theme in the same light of the

Qur'anic context to maintain accuracy and consistency. To put it differently, this research highlights the importance of SP by suggesting that it should be investigated, and its analysis be utilised as a tool for achieving congruency in translation.

Finally, on the level of language resources, this research produces a bilingual list of stop-words⁹ that are inclusive of all the function words in the Qur'an. It also presents six SP tagged datasets of the Qur'an and its five translations; each illustrating the annotation of the evaluative prosodies (i.e., *positive, negative, and neutral*) and discourse prosodies (i.e., *glorifying of God, reward in the afterlife, punishment of the present life, miracle story, etc.*) as well as the lexico-grammatical patterns of nature in the Qur'an (e.g., *N+N* (noun+ noun), *N+V* (noun+ verb), etc.). Another linguistic resource that this thesis presents includes a machine-readable¹⁰ (i.e., a plain text file in this research in [.txt] format) English translation of Tafsīr (the Interpretation of the Holy Qur'an) by Al-Jalalayn,¹¹ which is, as (Hamza, 2008, ii) writes in the introduction to its translation, "one of the most popular Tafsīrs in the Islamic world, perhaps even the most popular Tafsīr". It used the *al-Ijmaliy* (translated as the general method) method of interpretation,¹² whereby the meanings of the Qur'an are explained by using straightforward language (Embong and Hanapi, 2017, p.527). By converting a pdf file of the English translation *Tafsīr Al-Jalalayn*¹³ into a plain text file to be paralleled with the original Arabic text,¹⁴ this thesis produces a verse-by-verse English interpretation of the Qur'an that can be aligned with the original text in future corpus-based studies to explore the meanings of the Qur'an.

1.4.2 Limitations

One limitation of this study is that it places emphasis on nature, which is one of eight themes to formulate the fundamental message of the Qur'an (Malik, 1997; Fazlur Rahman, 2009).

9 In computing, *stop-words* are words which are filtered out before or after processing of natural language data (text). For some search engines, these are some of the most common, short function words, such as *the, is, at, which, and on*. See also (Rajaraman, 2011, pp.1-17).

10 *Machine-readable* data, or computer-readable data, is data (or metadata) in a data format that can be automatically read and processed by a computer, such as CSV, JSON, XML, TXT, etc. Available from: [<https://opendatahandbook.org/glossary/en/terms/machine-readable/>],[Accessed 25 June 2019].

11 It is a fifteenth-century Qur'anic commentary or Tafsīr of 'the two Jalāls' (al-Jalālayn) — the Egyptian scholar Jalāl al-Dīn Muḥammad b. Aḥmad al-Maḥallī (d. 864 AH / 1459 CE), and his (also Egyptian) student, the scholar, Jalāl al-Dīn 'Abd al-Raḥmān Abī Bakr al-Suyūfī (Al-Khalidi, 1996, p.31).

12 An *interpretation method* is a discipline which places specific procedures to interpret the words in the Qur'an. Every process that involves revealing the meaning of words in the Qur'an is categorised as an interpretation method. The *al-Ijmaliy* method is one of the Qur'an interpretation methods in the modern era, which interprets the words of the Qur'an by demonstrating meanings in a simple way and by using straight-forward language. Some interpretations of the Qur'an composed by using this method are *Tafsīr al-Jalalayn* by alSuyuti and al-Mahalliy, *al-Wajiz fi Tafsīr al-Kitāb al-'Aziz* by al-Wahidiy al-Naysabury and *Sofwat al-Bayan lima'ani al-Qur'an* by Husin Makhluḥ. See also Embong and Hanapi (2017, pp.521-35).

13 Hamza, F. 2008. *Tafsīr Al-Jalalayn (an English translation)*. [Online Version]. Pdf file is available from: [<http://altafsir.com>].

14 Available from: [<http://tanzil.net/docs/>],[Accessed 15 December 2016].

Hence, the analysis of SP as a collocational phenomenon in this research is limited to this theme without any direct association with the other themes. Another limitation is that in its evaluation of the selected translations of the Qur'an, this study considered the role of SP as one factor in maintaining lexical cohesion; it might consequently overlook other factors.

1.5 Introducing the theoretical framework

This section presents the theoretical framework of this research, which was drawn from previous literature on collocation and semantic prosody; namely the model on which the methodology was based to address the research questions. This model focuses on an area that only a few researchers of the language of the Qur'an have investigated, which is the importance of the collocational structure in the construction of meaning (e.g., concepts and themes in the Qur'an). Hence, the exploration of the SP of nature in the Qur'an depends on the analysis of collocational patterns of words, which are relevant to natural phenomena. It also operates in an interwoven pattern of theoretical constituents which start with collocations of a lexical item (i.e., word relevant to natural phenomena); moves to Sinclair and Halliday's functional lexico-grammatical structures of the collocations found and their meanings, and ends with semantic prosody (See Figure 1).

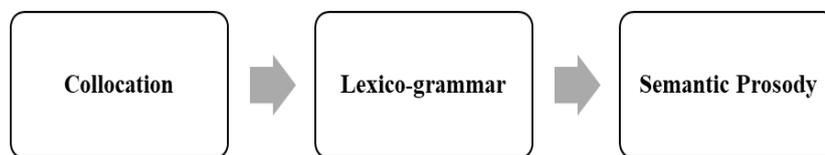


Figure 1: A model for the theoretical framework

The model designed for this research is also based on Stubbs's approach to the study of words and phrases in *lexical semantics*.¹⁵ This approach is summarised in two points: *meaning is use* and *corpus semantics*. *Meaning is use* entails that the meaning of words and phrases differs according to their use in different linguistic and social contexts. In addition, *corpus semantics* refers to an approach to analysing language in which "observational data from large text collections are used as the main evidence for the uses and meanings of

15 The study of how the words of a language denote either things in the real world or concepts. Available from: [<https://www.thefreedictionary.com>]. See also (Stubbs, 2001, p.22).

words and phrases” (Stubbs, 2001, p.22).¹⁶ This applies to the Qur’an in the sense that it is a text with a network of recurrent concepts:

Distribution of a particular concept or subject over many scattered verses within different chapters is very evident in the Qur’an. Often a concept summarized in one verse is elaborated in another verse. Historical events, stories of prophets, emphasis on a command, attributes and qualities of God, description of paradise and hellfire, are some of the common subjects that are often repeated in the Qur’an. However, each repetition adds new meanings absent in other instances, and the overall subject could be fully understood when all instances are taken into consideration (Sharaf and Atwell, 2012, p.2295).¹⁷

This property of the Qur’an makes it an attractive text for exploring SP through the analysis of the frequent patterns in which these “concepts” occur. It also makes possible the formation of a model that can shed light on the different concepts relevant to the theme of nature in the Qur’an. It is a model that helps to provide evidence of whether or not natural phenomena in the Qur’an are semantically prosodic as they appear in different contextual environments. Also, it establishes a scale for weighing the meaning of nature as a theme in the Qur’an via SP and comparing it to its parallel representations in the five translations of the Qur’an (See Section 3.1 for a detailed discussion of the components of the model of this research).

1.6 Natural language processing, programmes and data analysis tools

The first observation common to most recent research on collocation, particularly studies that explore the inseparability of the lexical item from its grammar like the present thesis,¹⁸ is that Natural Language Processing (NLP) is empirical and based on large amounts of naturally occurring text. In addition, frequency of occurrence and co-occurrence of language items are crucial, and corpora and corpus tools (software packages, computer scripts or online search interfaces) are used to identify which items are common in which contexts and in which types of discourse (Römer 2009, p. 148). Therefore, my corpus-based research employs a state-of-the-art technology as in the NLP programming language with Python in a quantitative approach to initially explore SP of natural phenomena in the Qur’an. Moreover, it relies on linguistic resources and tools (e.g., concordance lines) to

¹⁶ See also: [<https://www.uni-trier.de/fileadmin/fb2/ANG/Linguistik/Stubbs/stubbs-2001-words-phrases-ch-1.pdf>], [Accessed 10 July 2018].

¹⁷ Available from: [<http://textminingthequran.com/papers/qursim.pdf>], [Accessed 10 July 2018].

¹⁸ For example, Sinclair, 1991; Goldberg, 1995; 2006; Hunston and Francis, 2000, Biber et al., 2004; Hoey, 2004; 2005; and Hoey et al., 2007.

then qualitatively interpret the results of the exploration aligned with the theoretical approach to language in this research (following the definition of corpus linguistics by McEnery and Hardie (2012), Sinclair (2004c), Thompson and Hunston (2006), and Gries (2010b).

This section will give a brief overview firstly, of NLP with Python as the chosen NLP programming language; it will then introduce the linguistic resources and computational tools used to define the list of natural phenomena and conduct the analyses of texts in this research. They are introduced in this chapter to give the reader an understanding of the computational and corpus-based methodology employed in this thesis. In addition, an introductory overview of the features of Python will be provided to highlight the nature of the tasks and the data analysis (i.e., statistical analysis) that are discussed in-depth in the upcoming chapters.

1.6.1 Natural language processing with Python

Natural Language Processing (NLP) is an interdisciplinary subfield of computer science that is concerned with the processing of natural languages from the computational perspective (Mitkov, 2005).¹⁹ The nature of projects in this field is often expected to be managed by multidisciplinary teams such as computer scientists and linguists. To many people in academia, NLP is known by the name of ‘Computational Linguistics’ (Bird et al., 2009, x).²⁰ In this research, Python is the programming language used to statistically analyse the parallel corpora (quantitative analysis), chiefly because it is “a simple yet powerful programming language with excellent functionality for processing linguistic data” (as in Bird et al., 2009, x). In this regard, it is implemented by using the available linguistic libraries to manipulate textual data, the most popular of which is the Natural Language Toolkit (NLTK) with its corpora for many languages other than English (e.g., Arabic). Figure 2 shows the download page for NLTK.

¹⁹ Available from: [<https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199276349.001.0001/oxfordhb-9780199276349>], [Accessed 19 October 2016].

²⁰ Also available from:

[https://doc.lagout.org/programmation/python/Natural%20Language%20Processing%20with%20Python_%20Analyzing%20Text%20with%20the%20Natural%20Language%20Toolkit%20%5bBird%2c%20Klein%2c%26%20Loper%202009-07-10%5d.pdf], [Accessed 19 December 2016].

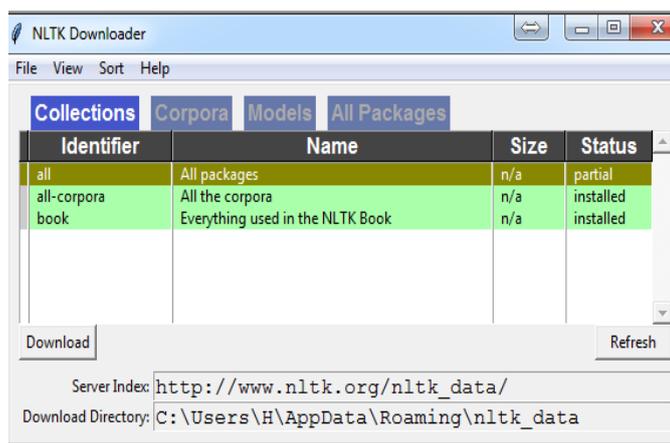


Figure 2: The download page for NLTK ²¹

NLTK has various modules with different functions, which help to explore linguistic aspects manifested in the analysis of syntactic and semantic dimensions of texts. Examples of these modules with their functions are shown in the following table from Bird *et al.* (2009).

Table 1: Examples of language processing tasks and corresponding NLTK modules with examples of functionality (table taken from Bird et al., 2009, xiv)

<i>Language Processing Task</i>	<i>NLTK Modules</i>	<i>Functionality</i>
Accessing corpora	nlk.corpus	Standardised interfaces to corpora and lexicons
String processing	nlk.tokenize, nlk.stem	Tokenizers, sentence tokenizers, stemmers
Collocation discovery	nlk.collocations	t-test, chi-squared, point-wise mutual information
Part-of-speech tagging	nlk.tag	n-gram, Brill, HMM, TnT
Classification	nlk.classify, nlk.cluster	Decision tree, maximum entropy, naive Bayes, EM, k-means
Chunking	nlk.chunk	Regular expression, n-gram, named entity
Evaluation metrics	nlk.metrics	Precision, recall, agreement coefficients
Parsing	nlk.parse	Chart, feature-based, unification, probabilistic, dependency

Finally, NLTK is primarily employed for first defining the list of natural phenomena compiled in this research (Chapter Two), and then for natural language processing over six texts: the Qur'an and its five selected translations. In this textual analysis, it is used for pre-processing texts; processing texts; finding frequencies of words referring to natural

²¹ Available from [<http://www.nltk.org/>], [Accessed 19 October 2016].

phenomena in the Qur'an; finding their collocations, namely bigrams;²² and applying an association measure²³ on these bigrams (Chapter Four).

1.6.2 Linguistic resources and computational tools

The linguistic resources and computational tools used in this corpus-based study are employed to produce six datasets to be evaluated for their representation of SP of nature in the Qur'an. These resources and tools with brief descriptions are listed below.

1- <http://tanzil.net/>

Tanzil.net is a linguistic resource drawn on in this study, namely a website from which machine-readable versions of the Qur'an as a raw data source can be obtained. It is a part of the Tanzil Project, which consists of the original verses in Arabic as well as 42 manual translations of the entire book. In this research, four of the translations were downloaded as Text documents from this website, except for Haleem's translation which is not available on the website; hence, was converted from a pdf file into a Text document to apply the same algorithm to it. They were run against Python via NLTK to obtain the statistical analysis of natural phenomena terms. In addition, they were copied into Excel for an alignment task²⁴ to find SP in the Qur'an and evaluate its English renderings. Furthermore, their versions with the appended Arabic root-disambiguated nature terms and stem-disambiguated translated English nature terms²⁵ were run in *Sketch Engine*²⁶ to produce concordance lines for each of the natural phenomena to visually represent the collocational behaviour of these natural terms in context (their SP).²⁷ Finally, these versions were run in

22 *Bigrams* are pairs of words, which can be found using association measurement functions found in the *nltk.metrics* package (Perkins, 2010, pp.21-3). This research will focus on bigrams and trigrams that are statistically proven collocations; not simply contiguous sequence of words.

23 *Association measures* (sometimes called collocation measures) are statistical measures that calculate the strength of association between words based on different aspects of the co-occurrence relationship. There are many different association measures, each producing (slightly) different lists of collocates (Evert, 2008; Gabalasova et al., 2017). See also Brezina (2018, pp. 66-71).

24 *Alignment* is a mechanism in corpus linguistics whereby a segment of text in the target text can be identified as the translation of the segment, or vice versa. This usually requires the two sets of texts to be aligned. Users of parallel corpora are often interested in retrieving instances of lexis or grammatical constructions in the source language together with their translations. (Olohan, 2004, p.198).

25 The corpora were pre-processed via Python and the root- disambiguated forms of the nature terms were appended to the text to explore with accuracy their frequencies, collocations, etc. [See Chapter Four].

26 The word *sketch engine* evolved from the program Bonito. It is a web-based Concordancing program. The sampler version which can be found at <http://www.sketchengine.co.uk/> uses the British National Corpus. It is a leading corpus tool widely used in lexicography. The Sketch Engine website offers many ready-to-use corpora, and tools for users to build, upload and install their own corpora. Available from: [<https://www.sketchengine.eu/>],[Accessed in October 2016 onwards].

27 More on root-based and stem-based disambiguation and the rationale behind applying these processes will be seen in Chapters Two and Four.

*LancsBox*²⁸ to draw the collocational networks and visualise their parallel concordance lines of nature in the Qur'an and its translations.

2- <https://www.altafsir.com>

The Qur'an tafsirs²⁹ in <https://www.altafsir.com> is another linguistic resource, which is a website that is employed as an aid to understanding the meanings of the Qur'an. The website *altafsir.com* was commissioned by the Royal Aal al-Bayt Institute for Islamic Thought³⁰ in Jordan and developed and maintained by the Integrated Technology Group. This is composed of a library of Islamic resources (i.e., old and contemporary works of Islamic theology and Tafsirs) related to the meanings of the Qur'an, such as *Ibn Abbās*;³¹ *Ibn Kathīr*;³² *Tafsīr al-Jalalayn*; *Asbab Al-Nuzul*.³³ This library was consulted while conducting the qualitative analysis of this research to determine the two forms of SP (i.e., evaluative and discourse prosodies) of the theme of nature in the Qur'an.

To put it differently, the researcher consulted them to arrive at a sort of consensus on the agreed contextual meanings of occurrences of nature in the Qur'an (prosodies). However, it should be mentioned that out of the resources mentioned here, *Tafsīr al-Jalalayn*, one of the commentaries of the Qur'an, was chosen as a primary source in the alignment of the parallel corpora for its popularity (as in Hamza, 2008, ii) and simplicity. The choice of *Tafsīr al-Jalalayn* is purely premised on a linguistic perspective and not as a reflection of the ideological background of its authors. In practical terms, the preference to use this tafsīr is its linguistic interpretation, where the general method focuses on the meanings of the verses tailored with brief syntactic and morphological explanations. The researcher found this feature especially useful in the alignment of verses of the Qur'an via Microsoft Excel, where each verse is adjacent to its tafsīr both in its Arabic and English versions.

28 *LancsBox* is a new-generation software package for the analysis of language data and corpora developed at Lancaster University. Available from: [<http://corpora.lancs.ac.uk/lancsbox/>], [Accessed 04 April 2018 onwards].

29 The word تفسیر *tafsīr* 'commentary or interpretation' is derived from the root فسر *fassara* 'to comment or to interpret'. Someone who writes Tafsīr is a مفسر *mufassir* 'commentator of Qur'an'. An example of a commentary is Ibn Kathīr's interpretation of the Qur'an.

30 *The Royal Aal al-Bayt Institute for Islamic Thought*, established in 1980 by the late H.M. King Hussein of Jordan, is a non-political and international charitable trust based in Jordan comprising of 70 to 100 members of the world's top Islamic scholars, who meet or correspond on a regular basis. It recognises all seven traditional madhhabs (or legal schools) of Islam. See also [<https://www.altafsir.com/aboutfoundation.asp>], [Accessed in 15 December 2016 onwards].

31 Attributed to the Companion Abdullah Ibn Abbas (d. 68/687) and Muhammad ibn Ya'qub al-Firuzabadi (d. 817/1414) and translated into English by Guezou in 2007, *Tanwīr al-Miqbās min Tafsīr Ibn 'Abbā* is one of the most pivotal works for understanding the environment which influenced the development of Qur'anic exegesis.

32 Ibn Kathīr, I.1983. *Tafsīr Al-Qurān*. Al-cAZīm: vols.1-4. Beirut: DarAlqalam.

33 Alī ibn Ahmad al-Wahidi is the earliest scholar of the branch of the Qur'anic sciences known as *Asbāb al-Nuzūl* (i.e. the contexts and occasions of the Revelation of the Qur'an). Available from: [<https://www.altafsir.com/Books/Asbab%20Al-Nuzul%20by%20Al-Wahidi.pdf>], [Accessed 15 December 2016 onwards].

Thus, this research not only relies on a web source to find meanings of the Qur'an but also provides a machine-readable commentary of the Qur'an (Verse-by-Verse) to align it in a spreadsheet containing the original verses and their English renderings. This means that the exploration of the meaning of a natural phenomenon term in the qualitative analysis of the Qur'an is immediate, and the congruency of the translations with the Arabic representation is systematic. In any event, the use of this commentary and others is an integral part of determining the following results of this research: the pragmatic functions of natural phenomena in the Qur'an (their discourse prosodies); their connotative colourings (their evaluative prosodies); and in turn the evaluation of the English renderings of SP of natural in the phenomena in the Qur'an (congruency vs divergence).

1.6.3 Data analysis and visualisation tools in the qualitative analysis

Since the nature of the data analysis in this thesis is both quantitative and qualitative, it requires two types of data analysis and visualisation. On the quantitative level, statistical visual tools, including frequency distribution graphs and collocation tables, are implemented in the course of analysing the produced datasets. On the other hand, the qualitative level, as mentioned above, employs the use of concordance lines³⁴ in examining the collocations and LG patterns of nature in the Qur'an and LanksBox to envisage the representation of collocations of nature as a network of meanings via collocational graphs and networks, which are defined as:

Summaries of complex meanings of words in texts and corpora. These networks can provide useful information about key topics [the theme of nature in the Qur'an] in texts and discourses as well as their connection (Brezina, 2018, p.79).

1- *Concordancing: A Core Function in Sketch Engine*

Also referred to as *Key Word In Context* (KWIC),³⁵ a concordance is a list of all of the occurrences of a particular search term in a corpus, presented within the context in which they occur- usually a few words to the left and right of the search term (Baker et al., 2006, p.42). The rationale behind the use of this core function in *Sketch Engine* is that, as with

³⁴ *Concordancing* is a [qualitative] analysis technique that allows linguists to investigate the occurrences and behaviour of different word forms in real-life contexts. This is quite different from more "traditional approaches" in linguistics that simply depend on the intuition of native speakers in order to determine the "correct" usage. See also Weisser (2016, pp. 67-79).

³⁵ *Key Word In Context concordance* is the preferred format for displaying concordance data because it is easy to observe the context of to the right and left. Available from: [https://www.sketchengine.eu/my_keywords/kwic/], [Accessed 19 October 2016 onwards].

collocations, concordances provide information about the “company that a word keeps” or its *discourse prosody*³⁶ (See Figure 3 for a sample of concordance lines in *Sketch Engine*). This research, therefore, employs this function in conjunction with the use of alignment to conduct the qualitative analysis which attempts to uncover the SP of nature in the Qur’an and its translations; that is, by examining the surroundings of the words describing natural phenomena.

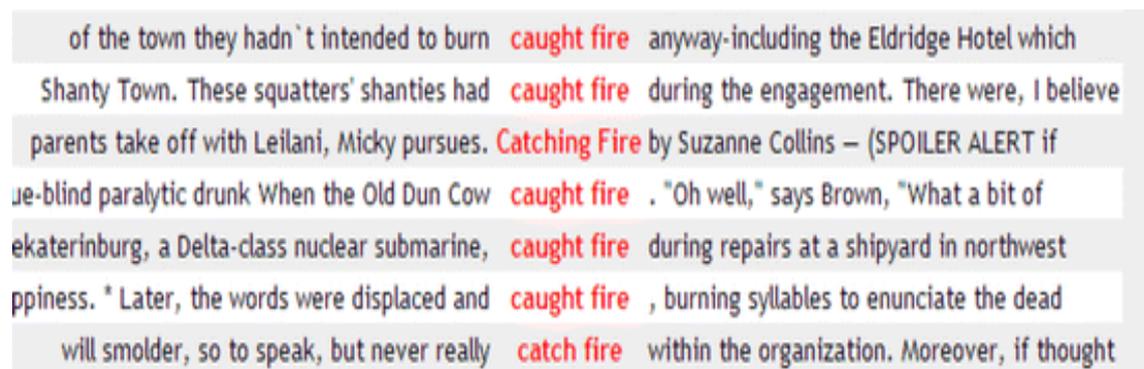


Figure 3: Concordance lines for *catch fire* (figure taken from Kilgarriff et al., 2014)

Following pioneering corpus linguists (e.g., Renouf and Sinclair, 1991; Stubbs, 1995; 2001; Louw, 2000; Sinclair, 2003), concordance lines are used to navigate the context of the collocations (i.e., bigrams) found in the statistical analysis of the collocation extractions. A seven-step procedure is followed for the analysis of a selection of concordance lines for a specific node (Sinclair, 2003, xvi-xvii) and will be described later in this thesis (Chapter Four).

2- *LancsBox v. 4.x*

As previously indicated, on the qualitative level, this research employs *LancsBox v. 4.x*,³⁷ which is the fourth version of a new-generation software package for the analysis of language data and corpora developed by Lancaster University. It is a recently developed and appraised data visualisation tool, which the researcher found useful for the data visualisation of collocational networks of natural phenomena³⁸ and the aligned concordance lines of nature in the Qur’an and its translations.

³⁶ *Discourse prosody* is a term reported by Stubbs (2001) relating to the way that words in a corpus can collocate with a related set of words or phrases, often revealing (hidden) attitudes (as cited in Hunston 2007, p.251).

³⁷ See Section 4.1.5 for the rationale of opting for Lancsbox instead WordSketch in Sketch Engine.

³⁸ This is done via employing the Loglikelihood statistic, which is built in the software. [Chapter Four]

1.7 Introducing the qualitative analyses of textual data

This section gives a preview of the qualitative analysis that is inherent in the mixed approach used in this research incorporating quantitative and qualitative analyses. The flow of tasks in the textual analyses to explore collocations and identify the SPs of nature in the Qur'an and its translations following this approach is shown in the following figure.

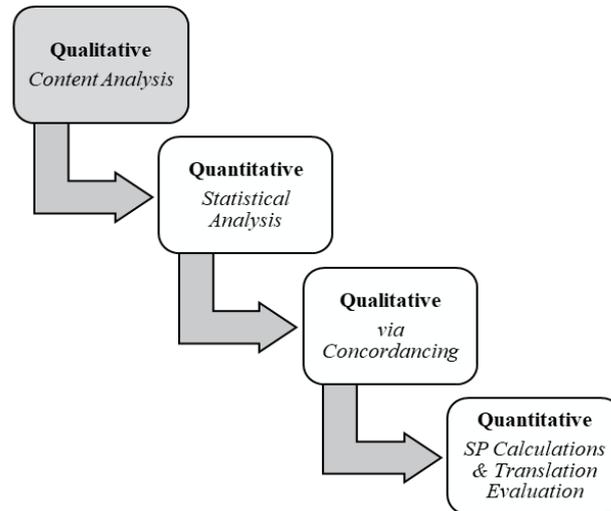


Figure 4: The data analyses of SP of nature in the Qur'an and its translations

Furthermore, this approach employs Stubbs's (2001) classification of textual data, where each task of the data analysis is applied to a specific type of textual data, as seen in the following table:

Table 2: Types of textual data and data analyses in this research (classification based on Stubbs, 2001, pp. 66-7, *my hyphenations*)³⁹

<i>Type of Data</i>	<i>Description</i>	<i>Textual Data Analysis</i>
first-order data	raw corpus data	content analysis [qualitative]
second-order data	corpus data as manipulated by a basic concordance program	collocation-via-concordance [qualitative]
third-order data	corpus data that has been manipulated using statistical analyses to present patterns within the data	collocation-via-significance [quantitative]

As shown in Table 2, the qualitative analyses in this research are applied on two types of data: the first-order data shown in the raw corpus (the Arabic Qur'an)⁴⁰ and the second-order data in the corpus as processed by a concordance program (Sketch Engine).

³⁹ See also McEnery and Hardie (2012, p.127). Available from: [https://epdf.pub/corpus-linguistics-method-theory-and-practice.html], [Accessed 04 May 2019].

⁴⁰ See also Chapter Two.

This section of the introductory chapter will focus on the qualitative part solely because, while the adopted tasks of the quantitative analysis of this research have been drawn from a previously tested model of statistical analysis to investigate collocations (e.g., Evert, 2005; 2008; Bartsch and Evert, 2014), the qualitative analysis of words related to nature in the Qur'an is used differently in two instances in this thesis. Firstly, it is used in the content analysis (CA) in the preliminary stage (pre-methodology) of the research to categorise the contextual meanings of nature in the Qur'an. Secondly, it is used in the methodology of this research to tag the evaluative and discourse prosodies of nature in the Qur'an and its translations, an approach that has not been explicitly used before in this manner by researchers on SP in general and SP in the Qur'an in particular (e.g., Louw, 1993; 2000; Partington, 1998; 2004b; Sinclair, 2004a; Younis, 2018). The following lines will present a brief description of the two instances of the qualitative analyses in this research; details of the first will be found in Chapter Two and the second in Chapter Four.

In its first instance, the qualitative analysis or CA is used to elicit the contextual meanings of words describing natural phenomena in the Qur'an based on the previous literature; it is not within the framework of the methodology chapter only because it occurs at the preliminary stage. CA is a technique for making replicable and valid inferences from data and their context "to provide knowledge, new insights, a representation of facts, and a practical guide to action" (Krippendorff, 2018, p. 403). The purpose, as Krippendorff (2018) argues, is to obtain a condensed and broad description of the phenomenon (in this case, *nature* in the Qur'an), and the outcomes of the analysis are concepts or categories describing it. Usually, these concepts or classes are used to build up a model, a conceptual system, or a conceptual map of categories from data which are "texts to which meanings are conventionally attributed: verbal discourse, written documents, and visual representation" (Krippendorff, 2018, p. 403). Furthermore, Hsieh and Shannon (2005) state that the current applications of CA rely on one of three distinct approaches to interpreting meaning from the content of text data: conventional, directed, or summative. They are described in the following table:

Table 3: Three approaches to content analysis (table taken from Hsieh and Shannon 2005, p.1286)

<i>Type of CA</i>	<i>Study Starts With</i>	<i>Timing of Defining Codes or Keywords</i>	<i>Source of Codes or Keywords</i>
Conventional	Observation	Codes are defined during data analysis	Codes are derived from data

<i>Type of CA</i>	<i>Study Starts With</i>	<i>Timing of Defining Codes or Keywords</i>	<i>Source of Codes or Keywords</i>
Directed	Theory	Codes are defined before and during data analysis	Codes are derived from theory or relevant research findings
Summative	Keywords	Keywords are identified before and during data analysis	Keywords are derived from the interest of researchers or review of literature

In its preliminary stage (as seen in the grey box in Figure 4), this research implements the *summative approach* and builds a relatively exhaustive list of terms for natural phenomena. However, instead of calling them *keywords*,⁴¹ the list of natural phenomena is referred to as a list of terms, since they do not represent all the themes of the Qur'an. They only represent nature as a Qur'anic theme and the underlying meanings related to it. These terms and their contextual meanings, derived from the review of literature on Qur'anic studies about nature and the researcher's close reading of the Qur'an, were identified before and during the data analysis. Then, CA is conducted to provide an understanding of these terms inductively;⁴² data is organised by coding to create categories and abstractions (Hsieh and Shannon, 2005, p. 1281; Bernard and Ryan, 1998). *Coding* means that notes and headings are written while examining the literature on nature in the Qur'an; *abstraction* means generating categories which depict the different contexts in which nature occurs in the Qur'an (Bernard and Ryan, 1998, p.608 and p. 619).

On the other hand, in its second instance in this research and, as seen in Figure 4, the task of qualitative data analysis falls within the approach of analysing *collocation-via-concordance*,⁴³ which is a non-statistical technique where a linguist uses his/her intuitive scanning in the inspection of concordances. This *hand-and-eye* technique⁴⁴ is usually implemented in other *neo-Firthian*⁴⁵ research and has been scrutinised and praised by linguists, who used it such as Stubbs (1995, pp. 27-8) and Younis (2018, p.126). Louw (2007b), who is also in favour of this approach, believes that to uncover the hidden meanings of any text it is a requirement that the "corpus stylisticians" use a concordance.

41 *Keywords* in corpus linguistics are defined as words in a corpus whose frequency is unusually high (positive keywords) or low (negative keywords) in comparison with a reference corpus. Available from: [https://www.kent.edu/appling/corpus-linguistics-glossary], [Accessed 04 May 2019].

42 *Inductive content analysis* is a qualitative method of content analysis that researchers use to develop a theory and identify themes from raw textual data. It relies on inductive reasoning, in which themes emerge from the raw data through repeated examination and comparison, and reduces the material to a set of themes or categories. See also Thomas (2006, pp. 237-46).

43 In contrast to the *collocation-via-significance*; both are commonly used techniques of analysing collocation. See also McEnery and Hardie (2012, pp. 122-66).

44 See also McEnery and Hardie (2012, p.125, *my italics in the above*).

45 The phrase *neo-Firthian* collectively refers to a group of scholars who preferred Firth's approach to collocation and discourse.

He states that “the literary world of any text is assembled afresh every time that text is read”, and to demonstrate this process:

Corpus stylisticians require only three things: (i) the literary text in a machine-readable form which allows us to read the text by random access as well as linearly in its traditional paper or hard copy form; (ii) a reference corpus of natural language of both spoken and written and containing fiction and non-fiction; (iii) concordance software containing collocator and a facility for the co-selection expressions and which produces raw data as its output (p.104).

However, since this research attempts to employ a mixed approach in exploring the collocations of natural phenomena in the Qur’an, and in lieu of extracting collocations manually from concordance lines as in Stubbs (1995) and Younis (2018), the statistically verified patterns (i.e., collocations) previously found in the quantitative analysis as seen in Figure 4, are utilised to provide explicit criteria (e.g., the statistically verified collocate sets for each node) for coding the data with the appropriate SP for each of the nature terms. This approach also agrees with Stubbs’s (2001) later viewpoint on the importance of the primacy of the human analyst over the statistical results of collocation, and that he/she should constantly be checking the outcomes of the mathematical calculation of collocation against concordance and raw text (Stubbs, 2001, p.71). Hence, the purpose of this task is to employ the statistical results of collocation in identifying the SP of nature in the Qur’an and its translations. The revealed SP meanings [i.e., evaluative and discourse prosodies] of nature terms in this task are based on the [statistically verified] collocations of the words referring to natural phenomena (i.e., nodes) with other words that unveil these meanings (i.e., collocates). Finally, a detailed discussion of this second instance of qualitative analysis will be provided in Section 4.1.5 of this thesis.

1.8 Introducing the criteria for evaluation of the translations

This section discusses the challenging issue of the evaluation of translation, which is conducted in the second phase of this research. It introduces this part of the methodology because its criteria should be defined before elaborating on the method of this process. The evaluation of translation is defined as “placing a value on a translation in terms of grade or pass mark” (McAlester, 1999, p.169). However, this definition is said to be confined to translation teaching and training; “an unnecessary limitation of the wider concept of evaluation in translation studies” (McAlester, 1999, p.169). In a broader sense and in harmony with the method and theory of the present research, House (1997) puts forward

the functional pragmatic model of translation evaluation, which focuses on the preservation of “meaning” across two different languages and cultures. In this sense, she views translation as “the re-contextualisation of a text in L1 [the first language of the source text such as Arabic] by a semantically and pragmatically equivalent text in L2 [the second language of the target text such as English]”, and identifies three aspects of meaning particularly important for translation evaluation which are: the semantic, pragmatic, and textual aspects (House, 1997, p.30-6). Following these aspects, the analysis of SP features, as a textual phenomenon and a means for translation evaluation in this research, relies on both the semantic and pragmatic aspects of meaning (cf. Partington, 1998; Stubbs, 2001; Hunston, 2002; Sinclair, 2004b).

Furthermore, Hewson (2011) argues that the evaluation of translation is a conscious act that examines “degrees of similarity to or divergence from the source text’s perceived interpretative potential” and that it attempts to set out this interpretative potential of a translation in “the light of an established interpretative framework whose origin lies in the source text” (pp.6-7, as cited in Al Ghamdi, 2015, p.14). This is the very definition of the type of evaluation this study implements to decide on the congruency of the translations in the representation of the theme of nature in the Qur’an. In essence, the second aim of this research is to describe translated texts based on the results of the source text by using parallel corpora; that is, to identify differences but not judge them. Furthermore, Al Ghamdi (2015) notes that this type of method requires standards (i.e., criteria) on which the evaluation is based to draw objective conclusions and avoid subjectivity in evaluating the quality of the translations.⁴⁶ Thus, the criteria chosen for this research upon which the evaluation is based are *accuracy* and *consistency* of terminology in translation (as in Al Ghamdi, 2015, p.10).

Accuracy of translation can generally refer to being precise in rendering from one language into another, i.e., making the correct choices of equivalent terms in the target language text to convey the meaning of the source language text (Baker, 1993, p.233). It does not mean word-for-word translations; instead, it can be achieved through functional or communicative means to convey the source/original meaning to the target reader (Hatim and Munday, 2004, p.11). Al Ghamdi (2015), who employed the same criteria in evaluating

⁴⁶ As part of his PhD thesis in the University of Leeds, Al Ghamdi (2015) conducted a study on evaluating translations of the Qur’an in their translation of the Divine Names of Allah. Available from: [<http://etheses.whiterose.ac.uk/9592/1/Saleh-PhD-Thesis.pdf>], [Accessed 18 May 2019].

translations of the Qur'an in their translation of the Divine Names of Allah, states that when it comes to assessing the translations of terms in the Qur'an (i.e., natural phenomena in the current study), the linguistic features (e.g., semantic and syntactic features) of these terms are essential in the evaluation process (p.15). In this research, the linguistic features found to be crucial are frequencies, collocations in the form of Lexico-Grammatical (LG) patterns, and SP. They are measured using quantitative statistics via frequency finders, collocation extraction, and an association measure to calculate the statistically most significant collocations (For details on accuracy and SP see Section 4.2).

Moreover, *consistency* of terminology in translation, as Al Ghamdi (2015, p.27) sees it, refers to lexical cohesion which is a significant feature of literary texts. He asserts that consistency or uniformity of terminology is one of the characteristics of any good translation.⁴⁷ In the light of this notion, a data evaluation method which focuses on consistency is used to compare the Qur'an translations. Similarly, he states that translators of the Qur'an are expected to maintain a high level of consistency and uniformity or, as he calls it, "lexical cohesion" in their choices of the suitable rendering of the language of the Qur'an (Al Ghamdi, 2015, p.28). Hence, this study devises the analysis of SP as a cohesive lexical tool and adopts the viewpoint of Morley and Partington (2009) who describe SP as the mechanism showing evaluative meaning, or what they call "evaluative harmony", which endows discourse with cohesion. They write:

Semantic prosody is the mechanism which shows how one elemental type of meaning — evaluative meaning— is frequently shared across units in discourse and, by ensuring consistency of evaluation or evaluative harmony, plays a vital role in keeping the discourse together, in its cohesion; which, of course, in turn, helps maintain comprehensibility for the listener (p.139).

In sum, to draw objective conclusions and avoid subjectivity when evaluating the five selected translations, criteria of assessment are chosen to suit the purpose of this research. Thus, the functional or communicative means to convey source/original meaning in the definition of the accuracy criterion and the lexical cohesion and harmony in the definition of consistency can be seen through the SP meanings. This research uses results of the SP of natural phenomena in the Qur'an and its five translations as a means of evaluating congruency vs divergence (as in Ebeling, 2014). These criteria were chosen due to their

⁴⁷ Al Ghamdi states that he follows the Newmark (1988) tradition of translation criticism.

relevance to the adopted definition of SP in this research, which is that it is both a collocational phenomenon with *evaluative prosody* and a cohesive device with *discourse prosody*.

1.9 Introducing the research corpora

The compiled data in this research consists of parallel corpora or, as Hofland and Johansson (1998) collectively call it, a *translation corpus*. These are corpora that make up a translation corpus “consisting of a set of texts in one language and their translations in another language” (Olohan 2004, p.24). They consist of six machine-readable Text Documents downloaded from Tanzil.net, except for Haleem (2004) as mentioned previously. In this regard, they contain the Arabic Qur’an and five of its English translations: Pickthall (1930); Yusuf Ali (1939/ revised edition 1987); Arberry (1957); Saheeh International (1997); and Haleem (2004). The first text is the Arabic Holy Qur’an (in two versions aligned next to each other: with and without *taškīl*, which is translated as diacritics); its analysis is employed to address the question whether semantic prosody as a linguistic phenomenon is present in the Qur’an. Furthermore, it is the source text based on which the comparison (e.g., the comparison of frequencies and SPs) among the translations is established. Finally, the translations have been chosen based on the quality of reviews they attained (e.g., Robinson, 1999, p.291 describes Ali's and Pickthall's as the most accurate of their time) and the fact that each of them has enjoyed a good reputation amongst Muslims at different historical periods (as in Al Ghamdi, 2015, p.32), is the diversity of the translators’ backgrounds.

1.9.1 The Holy Qur’an

The Holy Qur’an, consisting of 114 surahs (chapters) and 6,236 ayahs (verses), is considered by many scholars a rich source for linguistic and stylistic research. Although most of its language is said to be familiar to Non-Muslim and Muslim Arabs whose literature and speech have similar features; it still has its unique form inherently obtained from the Divine choice of words. It was promised in the Qur’an that it should not be altered or replaced: “We have, without doubt, sent down the Message; and We will assuredly guard it (from corruption)” (15:9).⁴⁸ This provides a solid ground for a researcher to observe the behaviour of lexical items in different environments and how their semantic features work

⁴⁸ From Yusuf Ali’s English Translation of the Qur’an (2013). Downloaded from: <http://tanzil.net/>. Reference to the verses of the Qur’an in this thesis will take the following system: (number of surah: number of ayah or ayahs).

in the representation of meaning. In this regard, the central theme that prevails throughout the Qur'an is the exposition of reality and the invitation to worship God (Malik, 1997, p.104); all other themes revolve around this theme. From the very beginning of the Qur'an to its end, the different topics it addresses (i.e., nature) are said to be intimately connected with its central theme. Also, it is a book with several types of traditional divisions as follows: chronological based on historical revelation; length-based rule; by subject; and *Juz*'s.⁴⁹ The most popular division is the one based on chronology, that is, the division of the Qur'an into Makkan and Medinan based on the place of revelation.⁵⁰ The following sub-sections of this chapter present the five English translations that were chosen for this study.

1.9.2 M. M. Pickthall (London, 1930)

Muhammad Marmaduke Pickthall (1875–1936), a British Muslim novelist, had an interest in the languages of the Middle East and mastered languages such as Arabic, Turkish, and Urdu. Pickthall's translation deserves credit for keeping close to the original Arabic (Haleem, 2004, xxviii). Besides, his work has the merit of being a first-rate English interpretation of the Qur'an at the time of its publication (Kidwai, 1987, p.66). In the preface of his book, Pickthall writes:

The Qur'an cannot be translated. The book is here rendered almost literally, and every effort has been made to choose befitting language. However, the outcome is not the Holy Qur'an, that inimitable symphony, the very sounds of which move men to tears and ecstasy. It is only an attempt to present the meaning of the Qur'an -and peradventure something of the charm in English. It can never take the place of the Qur'an in Arabic, nor is it meant to do so (Pickthall, 1930, vii).⁵¹

These lines reflect his view on the translation of the Qur'an when he states that "the Qur'an cannot be translated" and that a translation of the Qur'an can never be produced, "nor is it meant to do so". They show that he endorses the position of Muslim scholars that the Qur'an is untranslatable;⁵² however, he maintains that the general meaning of the text could still be conveyed to English speakers. In his translation, he writes an introduction for each

49 The main division of the Qur'an is into chapter (surah) and verse (ayat). The Qur'an is additionally divided into 30 equal sections, called *juz'* (plural: *ajiza*). The divisions of *juz'* do not fall evenly along chapter lines. These divisions make it easier to pace the reading over a month's period, reading a fairly equal amount each day. See also: [http://corpus.quran.com/documentation/versemarks.jsp].

50 See also the Index for the division of the Quran in *King Fahd Complex for Printing the Qur'an*. [Available from: https://epub.qurancomplex.gov.sa/issues/hafs/standard39/].

51 Available from: [http://www.khayma.com/librarians/call2islaam/quran/pickthall/], [Accessed 18 May 2019].

52 See also Al-Jabari (2008, pp.17-21).

chapter and includes footnotes. Although its language might be “artificially archaic”, Pickthall’s translation was highly regarded by various readers (Haleem, 2004, xxviii; Khan et al., 2013, p.571).

1.9.3 Abdullah Yusuf Ali (Lahore, 1939)⁵³

Perhaps the most popular English translation to this day is the one written by Abdullah Yusuf Ali (1872-1952), *The Holy Qur’an: Translation and Commentary* (1st edition 1934 and revised in 1939-40), which was until the 1990s perhaps the most popular English version among Muslims. There are two main revised versions of Ali’s translation: the *King Fahd Qur’an Printing Complex* (1987)⁵⁴ and the Amanah Corporation’s (1989).⁵⁵ Both versions found merits in Ali’s original translation (Jassem and Jassem, 2001, p.36); yet the former is the version used in this research, only because it is the available machine-readable text on Tanzil.net.⁵⁶ As for the publication of this particular version, the committee in the *King Fahd Qur’an Printing Complex* chose Ali’s translation,⁵⁷ for its distinguishing characteristics, such as the “highly elegant style” and “a choice of words close to the meaning of the original text”(Deedat, 2012, p.49). In his article, “Assessing English Translations of the Qur’an”, Wild (2015) states that Ali “sought to convey the music and richness of the Arabic with English poetic versification”.⁵⁸ Also, it has been said that Ali was “gifted with an eloquent, vivid writing style”, and that the footnotes, despite their abundance, are helpful for those who want to fully understand the background of the text (Haleem, 2004, xxviii).

1.9.4 A. J. Arberry (London, 1957)

Arthur John Arberry was born into and lived in a conservative and well-educated English Christian family from which he established a strong desire for higher education (Skillitee, 1970, p.363). As a Cambridge University graduate, he spent several years in the Middle East perfecting his Arabic and Persian language skills. For a short period of his life, he

53 The revised version of this translation was used in this research; it is the document retrieved from Tanzil.net as: Ali, A. Y. (1987). *The Holy Qur’an: English Translation of the Meanings and Commentary*. King Fahd Holy Qur’an Printing Complex.

54 *King Fahd Complex for the Printing of the Holy Qur’an* (Arabic: مجمع الملك فهد لطباعة المصحف الشريف) is a printing plant located in Medina, Saudi Arabia that publishes the Qur’an in Arabic and other languages. The company produces about 10 million copies a year, and has published 55 different translations of the Qur’an in 39 languages. [Online] Available from: [https://qurancomplex.gov.sa/].

55 In 1989, Saudi Arabia’s Ar-Rajhi Banking Company funded the U.S.-based Amanah Corporation’s project to revise the translation of the Qur’an (Mohammed, 2005, p.58).

56 See also [http://tanzil.net/docs/translations_resources].

57 Alongside Al-Hilali and Khan (1983).

58 Available from: [https://www.meforum.org/717/assessing-english-translations-of-the-quran], [Accessed 18 May 2019].

worked as a Professor of Classics at Cairo University; in 1946, he was a professor of Persian at the University of London and transferred to Cambridge the following year to become a Professor of Arabic there until his death in 1969. His translation of the Qur'an, *The Koran Interpreted* (1957) was the first English translation by a scholar specialised in Arabic and Islamic studies. "He rendered the Qur'an into understandable English and separated text from tradition; it has earned the admiration of intellectuals worldwide" (Al-Jabari, 2008, pp.35-6). In addition, having been reprinted several times, his translation remains the reference of choice for most academics (Halimah, 2014, p.124). Finally, it has been said that unlike his predecessors, Arberry "paid close attention to the rhetorical features of the Qur'an and aimed to reproduce these splendid features" (Sadiq, 2010, p.8).

1.9.5 Saheeh International (Jeddah, 1997)

Saheeh International is a team of three American women whose names are not mentioned in this translation of the Qur'an. It is considered one of the most prominent and widely read English-language Qur'anic translations. The translators converted to Islam in the 1980s and lived and worked in Saudi Arabia. Their translation, entitled *The Qur'ān*, published by Abul Qasim Publishing House,⁵⁹ like Ali's, contains many notes but in parentheses in the running text. In the foreword, the translators say that their adopted word order mimics that of the Arabic text. They also avoided using transliterated Arabic terms. Also, they shift pronouns in many verses to match the Qur'anic style. They also note that particular words and phrases in the source text have various shades of meaning. Therefore, they argue that:

Any translation, which can reflect but one emphasis, must necessarily appear as a severe limitation. Although additions in brackets and explanatory notes are a partial remedy, it remains to be said that nothing can take the place of an in-depth study through the Arabic medium itself (Saheeh International, 1997, viii).

This translation was widely acclaimed and respected for accurately reflecting the meaning of the original Arabic meanings and was recommended by various well-known scholars.⁶⁰

1.9.6 M. A. S. Abdel Haleem (Oxford, 2004)

Abdel-Haleem is an Egyptian-British scholar and the editor of the *Journal of Qur'anic Studies*. He has an interest in Arabic-English lexicography and Qur'anic studies and has

⁵⁹ See also [<https://darulqasim.org/>].

⁶⁰ For example, Muhammad al-Munajjid, Zakir Naik, Yusuf Estes, and Jamal Badawi. See also: [<http://www.islamicbookstore.com>].

published several works in this field. He is recognised for his services to Arabic culture and literature and inter-faith dialogue (Shah 2010, p.1). In 2004, Oxford University Press⁶¹ published Haleem's translation of the Qur'an under the title: *The Qur'an: A New Translation*. Following his criticism of the previous translations of the Qur'an, Haleem (2004) claims that he opted for an explanatory translation of the Qur'anic meanings since the peculiarities of the English versions of his predecessors mostly resulted from "excessive literalism and adherence to the syntactical and stylistic peculiarities for the (Qur'anic) Arabic, (...) which is very concise, idiomatic, figurative and elliptic"(p.67). Finally, Haleem's translation has been described as the best-appraised translation of the Qur'an to be introduced to English readers by an Arab translator (Shah, 2010, p.2 ; Rippin, 2016, p.12).

1.10 Overview of this thesis

This thesis consists of six chapters. **Chapter One** is a general introduction that outlines the aims of this research, its motivation, the research questions, and related concepts of data analysis. **Chapter Two** presents an overview of natural phenomena in the Qur'an and reports the process of compiling and defining the list of natural phenomena via the NLP applications implemented in this research. **Chapter Three** presents a literature review of studies relevant to the theoretical background in this research which encompasses three linguistic features: collocations, Lexico-grammatical (LG) patterns, and SP. In addition, it reports on previous corpus-based studies on the Qur'an, on computational linguistic studies on natural phenomena in the Qur'an, and several methods of evaluation of translation. **Chapter Four** presents the corpus-based methodology implemented in this study by describing the two phases of this research and processes related to each. **Chapter Five** reports the findings of this research and discusses their relevance to the previous literature and theoretical background of this research. Finally, the conclusions and implications of this study will be summarised in **Chapter Six**, and some suggestions for future research will be made.

61 *Oxford University Press* is a department at the University of Oxford, which has worldwide publications. It is said to be the largest university press in the world, publishing in 70 languages and 190 countries. Available from: [<http://global.oup.com/?cc=gb>]

Chapter 2 Natural Phenomena in the Qur'an: A Descriptive Overview with NLP Applications

Introduction

Nature in the Qur'an, just as *God* and *Eschatology*,⁶² is found to be one the of the “predominant themes”; several nature terms emerge as statistically significant keywords in corpus comparison of the Qur'an with a large reference set (Brierley et al., 2018, p.427). In addition, the mentions of natural phenomena in the Qur'an with their “order; proportion; and beauty” portray them as *signs* of God's creation with its finitude in contrast to His “infinity and unity” (Fazlur Rahman, 2009, pp.66-7). The depiction of such order and proportion in the creation of natural phenomena can be seen in the following lines of the Qur'an:

وَالشَّمْسُ يَجْرِي لِمُسْتَقَرٍّ لَهَا ذَلِكَ تَقْدِيرُ الْعَزِيزِ الْعَلِيمِ ﴿٣٨﴾ وَالْقَمَرَ قَدَرْنَاهُ مَنَازِلَ حَتَّىٰ عَادَ كَالْعُرْجُونِ الْقَدِيمِ ﴿٣٩﴾

لَا الشَّمْسُ يَنْبَغِي لَهَا أَنْ تُدْرِكَ الْقَمَرَ وَلَا اللَّيْلُ سَابِقُ النَّهَارِ وَكُلٌّ فِي فَلَكٍ يَسْبَحُونَ ﴿٤٠﴾

﴿٤٠﴾

Yusuf Ali: And the sun moves (along its course) to its resting place that is the measuring (or determination) of the All-Mighty, the All-Knowing, and for the moon We have appointed certain stations until it returns like an old curved stick. It is not for the sun to overtake the moon, nor for the night to overstrip the day, each coursing in its own orbit.

Verses (36:38-40)

In this Qur'anic excerpt, the *sun* and the *moon* are mentioned to emphasise God's power in the perfection and order of His creation. Both natural phenomena are described as “floating” uniformly, each in its orbit, causing the occurrence of *day* and *night*. Man can then benefit from these regular movements to tell day from night; the allusion to both the *sun* and *moon* in these verses is shown in the natural causation sense to show God's power in the present order and beauty in the creation of the universe (Fazlur Rahman, 2009, p.67).

In the context of this research, nature as a theme in the Qur'an is described in this chapter by reporting Alshahrani and Brierley (2017),⁶³ which presents an overview of this theme

62 In the history of religion, the term *eschatology* refers to conceptions of the last things: immortality of the soul, rebirth, migration of the soul, and the end of time. Definition available from: [<https://www.britannica.com/topic/eschatology>].

63 Alshahrani, H. and Brierley, C. 2017. An overview of natural phenomena in the Qur'an. In: *Postgraduate Research Conference May 5th, 2017. University of Leeds, Leeds*.

in the Qur'an. It is a pilot study that can be situated between the literature review in the following chapter because it reviews some of the previous studies on natural phenomena in the Qur'an and the pre-methodology stage where the list of natural phenomena was produced before the implementation of the actual corpus-based methodology. Moreover, the rationale behind devoting a chapter to the preliminary work (i.e., the tasks of the pre-methodology stage) of this research is its importance as for most researchers in corpus linguistics (e.g., Sinclair, 1991; Goldberg, 2006; Hunston and Francis, 2000; Biber et al., 2004; Hoey, 2004), data and observation come first, and theory comes second. Hoey's Lexical Priming theory and Hunston and Francis's Pattern Grammar,⁶⁴ for instance, provide explanations for what is observed in the language. Biber and his colleagues look at "descriptive facts that require explanation" (as cited in Römer, 2009, p. 148). Hence, this chapter presents an essential exploratory study which observes and prepares the data that is later used in the methodology and explained in terms of the aligned theoretical framework of this research.⁶⁵ It describes the experimental work done to compile, classify, disambiguate, and statistically examine the terms which represent nature in the Qur'an. The researcher, who had no training in NLP with Python before commencing this research, conducted several experiments to explore its functions and utilise its application in her study. These experiments were the primary aspects on which the rest of the NLP workflow of this thesis stands. Hence, a chapter presents this study because it mimics the tasks in the preliminary stage of data observation in this research (i.e., the significant discovery stage before the actual first tasks of the methodology to find SP of nature in the Qur'an and the theoretical background employed to interpret the results of this research).

The tasks in the researcher's pilot study will be discussed throughout this chapter, two of which will include references to previous commentaries⁶⁶ on nature as a theme in the Qur'an; the other two on some NLP applications to collect and verify a list of natural phenomena. The compilation of terms describing natural phenomena was done to conduct a summative content analysis (CA) which categorises the contextual environments of nature in the Qur'an. It was based on the exploratory overview of the previous literature on nature as a Qur'anic theme as well as a close reading of the Qur'an. ~~In this regard,~~ After examining the previous research, the researcher applied a close reading of the Qur'an and

64 In the 1920s, Palmer started what would become a cornerstone of British Applied Linguistics: he devised lists of the most frequently used words and phrases, constructed what he later termed *Pattern Grammar* (which was then refined by A. S. Hornby in 1954 and taken up by Hunston and Francis in 2000) (as in Pace-Sigge, 2013, p.150).

65 Based on the definition of corpus linguistics by Thompson and Hunston (2006, p.8).

66 These studies on nature in the Qur'an in this chapter are not corpus linguistic studies unlike the other mentioned studies on the same Qur'anic theme in the following chapter which resemble this thesis in that they are purely corpus-based.

Ali's translation (i.e., beginning with reading the verses with the attention to nature terms, then viewing the text as a computed list of words via NLTK in Python) to elicit a list of natural phenomena terms in the Qur'an. Hence, by inspecting the available previous lists of nature terms and breaking the entire text into individual words using NLTK in Python (Section 2.1.3), it was possible to manually extract a comprehensive list of nature terms from the Qur'an as a whole. However, it should be highlighted that to cover all the nature terms in the Qur'an was beyond the scope of this study; therefore, a plausible exhaustive list of nature terms was compiled. The researcher then employed this list in a CA to produce the categories of contexts in which nature occurs. The *ontology*⁶⁷ of nature terms, partially adopted from the Qur'anic Arabic Corpus (QAC),⁶⁸ was then produced. After that, the linguistic features (i.e., stems, definitions with examples, synonyms, and antonyms) of these terms in the ontology were identified to reduce their variant forms (i.e., word disambiguation), and raw frequencies were found followed by a preview of their most frequent collocations via NLTK in Python. The importance of this ontology lies in the fact that it is the provider of the 30 most frequent nature terms which were considered for analysis in this research; they were analysed for their frequencies, collocations, and SP. Thus, the role of this chapter is to report the flow of tasks taken to produce and explore a semantic ontology of a compiled list of 154 terms denoting words relevant to natural phenomena in the Qur'an (See Figure 5 for illustration).

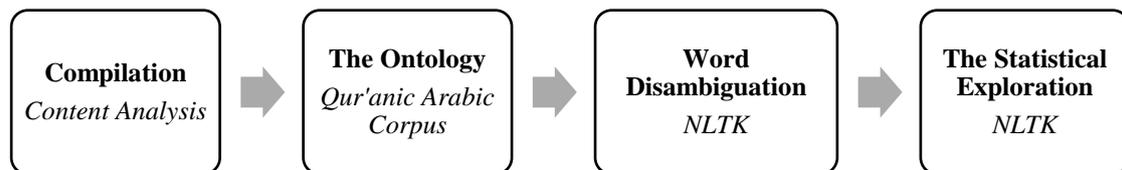


Figure 5: The flow of tasks of producing a list of nature terms in the Qur'an

To summarise, this chapter will discuss a brief account of the following four points: compiling the list of nature terms; producing an ontology of nature terms; disambiguating

⁶⁷ The word *ontology* is used with different senses in different communities. The most radical difference is perhaps between the philosophical sense and the computational sense. The adopted definition in this research is: the study of attributes that belong to things because of their very nature. See also (Giaretta and Guarino, 1995, pp.307-17). Ontologies consist of four components; concepts, relations, axioms, and instances (Gruber, 1993, p.2).

⁶⁸ The *Qur'anic Arabic Corpus* (<http://corpus.quran.com>) is an annotated linguistic resource with multiple layers of annotation including morphological segmentation, part-of-speech tagging, and syntactic analysis using dependency grammar. See also Dukes and Habash (2010, p.2530).

the list of natural phenomena via the WordNet language package in NLTK; and providing statistical profiling of nature in the Qur'an via NLTK.

2.1 Compiling the list of nature terms and their contextual meanings

This section focuses on gathering a working list of natural phenomena⁶⁹ and their contextual meanings which involved, firstly, the elicitation of a list of nature terms based on the previously compiled lists of natural phenomena in the Qur'an and a close reading of the Qur'an (i.e., verses and words) with consulting the Tafsirs, whenever needed.⁷⁰ Secondly, this was followed by the identification of the contextual meanings of nature via CA and the previous literature on nature in the Qur'an.⁷¹ A brief account of each will be given in the subsections below with the last subsection featuring the first compiled working list of nature terms in this research.

2.1.1 An overview of the available lists of natural phenomena

As the first step of compiling a list of nature terms for the CA at this stage, the available lists from previous studies were explored to collect a comprehensive list of words representing nature terms that occur in the Qur'an. In the first list, one of the few lists on nature as a Qur'anic theme (Abdul-Raof, 2001, pp. 8-9), examples of natural phenomena that appear in the Qur'an are compiled in one group of terms (group number 9), and they include the following:

1. *Animals in the Qur'an*: cow, livestock, camel, she-camel, pig, dog, elephant, calf, lion, monkey, donkey, mule, wolf, goat, sheep, ewe, horse, wild beast, frog.
2. *Insects in the Qur'an*: spider, fly, mosquito, ant, lice, locust, moth, serpent, snake, bee, woodworm
3. *Birds in the Qur'an*: hoopoe, quail, crow, birds of prey.
4. *Trees and plants in the Qur'an*: olive tree, lote tree⁷², grapevine tree, tamarisk tree, date palm, gourd tree, Zaqqum tree⁷³, thorny plant, tree of eternity, grass, vegetation, field.

69 It has been modified several times throughout the research, but the original list resides here in this chapter.

70 A marked corpus of the lines of the Qur'an with their tafsirs and the nature terms that appear in the nature related verses is available with the electronic appendices of this research.

71 No corpus-based method, such as collocation extraction to identify meanings of nature in the Quran, is used yet until Section 2.4.3.

72 ^{السدرۃ} Lote tree is a location in the afterlife that is mentioned in the Qur'an, and is usually considered to be part of Paradise Available from: [<http://corpus.quran.com/concept.jsp?id=lote-tree>], [Accessed 04 July 2018].

73 ^{الزقوم} Zaqqum tree is a tree that springs out of the bottom of Hell-fire, and is nourished by the fire, for it was created from fire. Quran Tafsir Ibn Kathir. Retrieved from: [http://www.qatafsir.com/index.php?option=com_content&task=view&id=1934&Itemid=93], [Accessed 10 July 2018]

5. *Fruits and vegetable in the Qur'an*: olives, figs, green herbs, onion, garlic, bitter fruit, sweet basil
6. *Sea animals in the Qur'an*: narcissus, musk.
7. *Planets in the Qur'an*: sun, moon, earth, shooting stars, stars, swimming orbits.
8. *Categories of people in the Qur'an*: man, woman, male, female, elderly, those who evoked God's anger, those who have gone astray, deaf, dumb, blind, lame, just, unjust, pious, believer, unbeliever, immigrant, supporter, idolater, wrongdoer, wretched, criminal, poor, rich, orphan, needy, wayfarer, slave, prisoner of war, winner, successful, loser, humble, doer of good, arrogant, disobedient, liar, corrupter, etc.
9. *Natural phenomena in the Qur'an*: day, morning, light, shadow, night, darkness, the passing of the night, fire, smoke, wind, whirlwind, clouds, cloud mass, hail, rain, water, spring, sea, river, bank, flood, waves, foam, torrent, land, sand, plain, valley, earthquake, heat, coolness, lightning, thunderbolt, mountains, mirage, heap of sand, shake, blast, elevation, dust, clay, stone, rock, fragments, horizons, sleep, lethargy, dreams, creation, the heavens and earth, falling fragments from the sky, setting of stars, the glow of sunset, etc.
10. *Time in the Qur'an*: dawn, morning, day time, night time, darkness, sunrise, sunset, summer, winter, the beginning of the day, end of the day, day, the crescent moons.
11. *Metals in the Qur'an*: gold, silver, iron, copper, shackles.
12. *Body parts in the Qur'an*: body, head, face, cheek, eyes, tears, mouth, lips, tongue, nose, tooth, neck, aorta, chest, throat, elbow, hand, arm, finger, nail, ear, leg, foot, ankle, heart, back, belly, intestines, hearing, sight, womb, embryo, flesh, bone, rib, skin, wound, disease.
13. *Directions in the Qur'an*: east, west, right, left.
14. *Colours in the Qur'an*: black, white, yellow, green, red.

The list shows that the category named *natural phenomena* is confined mostly to *weather phenomena*; it does not include *animals, plants, man*, which are categorised as distinct groups of words, and can be, by definition, said to be components of nature.⁷⁴ This fact is taken into consideration when compiling the list of natural phenomena in this research. In “Natural Phenomena Mentioned in the Qur’an”, another study on nature in the Qur’an, a linguistic resource for all topics and concepts related to the Qur’an is provided (Yahya, 2003, pp. 224-28). In this linguistic resource, natural phenomena are listed by quoting the excerpts of verses in the Qur’an in which they appear. This list of natural phenomena

⁷⁴ *Nature* is defined as: the animals, plants, rocks, etc. in the world and all the features, forces, and processes that happen or exist independently of people, such as the weather, the sea, mountains, the production of young animals or plants, and growth. (Definition of ‘nature’ from the Cambridge Academic Content Dictionary © Cambridge University Press). Available from: [<https://dictionary.cambridge.org/dictionary/english/nature>], [Accessed 10 July 2018].

mentioned in the Qur'an, which was useful along with other concepts from the book in compiling the list for this research, includes the following five components:

1. Wind
2. Lightning
3. Formation of day and night
4. The creation of the heavens
5. Periods of months and weeks

In addition, other chapters of this book concern different aspects of natural phenomena with their contexts which can be shown as follows:

Table 4: Yahya's natural phenomena in other chapters of his book and examples from the Qur'an (table taken from Yaya, 2003, pp.224-28)⁷⁵

<i>Natural Phenomena</i>	<i>Examples from the Qur'an</i>
Landscapes	mountains, tracts, valleys, gardens and springs, seas, earth and sky, and roads
Water	water in Paradise, water in Hell, Allah sends down water from the sky, Allah sends drinking water, Allah produces rain, Allah gives life to the earth with water, the barrier between two seas, ships sailing on the sea, and various sources of water, Allah made the sea subservient to people, Allah brings forth crops by water, the metaphor of the of the world, Allah's throne, man's creation, Allah drowned disbelieving nations, the rain of those who are warned, flood of 'Iram, Allah created everything from water ..., etc.
Jinn	jinn are created to worship Allah, jinn is created out of the fire, jinn has limited power, jinn transmits the Qur'an to their people, jinn whispering into the heart, ..., etc.
Angels	angels are honoured servants, angels that came into the prophet Ibrahim, angels of death, angels of Hell, angels of Paradise, ..., etc.
Time	<i>The relativity of time in the Qur'an</i> the man whom Allah caused to die a hundred years and then brought back to life, companions of the cave, comparison of time in the Hereafter with that in the world, a day in the sight of Allah is as a thousand years, the length of time in which angels ascend to the presence of Allah <i>Time mentioned in the Qur'an</i> verses in the Qur'an mentioning night and day angels' glorification, the time that the prophet Musa (as) fixed for Pharaoh, the prophet Musa (as) travels with his people at night, the time when Pharaoh's troops, started to pursue the Prophet Musa (as) and his people, the Prophet Zakariya (as) did not speak to people for three nights, the wind made subservient to the Prophet Sulayman, the morning of those who were warned Satan's, asking for a reprieve until the day of Rising, the Coming Hour of Rising, ..., etc.
Punishment	drowning, earthquakes, drought - lack of crops, troublesome animals, flood calamities, an awful blast, rain of stones of clay
Fruits, vegetables, trees, and plants	date, pomegranate, banana, fig, grape, trees, olives and others, plants used in the metaphors in the Qur'an, those have healing, plants in Paradise, plants in Hell, thorny bush, Zaqqum
Animals	wild animals, calf, livestock, cow, ewe, pork, donkey, horse, mule, camel, bird, crow, quail, fish, lion, elephant, gnat, locust, bee, spider, worm, snake, fly, metaphors, animals made subservient to people

Again, the list of natural phenomena in this study is not exhaustive; it lacks, for example, concepts that denote *astronomical bodies* (e.g., *moon, sun, stars*, etc.). Moreover, a preview

⁷⁵ Available from: [<http://www.harunyahya.com/en/Books/664/general-knowledge-from-the-quran/chapter/8262>], [Accessed 20 December 2016].

of raw frequencies of yet another incomplete list of natural phenomena in the Qur'an from Badawi (2012) is shown in the table below.

Table 5: Raw frequencies of الطبيعة *tabi'i* or words related to the natural phenomenon (table taken from Badawi, 2012, p.561)

Words	Transliteration	Translation	Frequency
السماء	<i>Sama''a, sama''i, sama''u.</i>	sky / heaven	120
نهر	<i>Anhara, anharu, (47), Anharan (3), Naharan (1)</i>	river	51
شمس	<i>Syamsa, syamsi, syamsu (32), syamsan (1)</i>	sun	33
قمر	<i>Qamara, qamari, qamaru (26), qamaran (1)</i>	moon	27
بر	<i>Barri, barru (1)</i>	land	13
طور	<i>Tura, turi</i>	mount (Sinai), a large mountain with trees	10
رواسي	<i>Rowasiya</i>	mountains	9
جبل	<i>Jabala, jabalin</i>	mount, hill without a tree	6

In brief, these lists of natural phenomena which were found in the previous literature, although very resourceful for this study in verifying the composition of the list of natural phenomena, they were not fully adopted for they lacked several vital concepts relevant to words referencing natural phenomena in the Qur'an.

2.1.2 A reading of nature in the Qur'an

Following the overview of the previously compiled lists of nature terms, the researcher performed a close reading of the Qur'an to produce an exhaustive list of natural phenomena that occur in the Qur'an. In this close reading, two corpora, obtained from *Tanzil.net*, were considered: the Arabic Qur'an and Yusuf Ali's translation. The purpose of choosing two corpora was to have a general idea of how these terms occur in both the source text and its translation. Also, two types of reading were performed: a reading of the texts as verses, and another reading of the texts as individual words. Therefore, 6,236 verses and 77,430 words⁷⁶ of the Qur'an,⁷⁷ as well as the translated verses and English equivalents in Ali's translation, were read to mark the verses and words that have occurrences of natural phenomena in both corpora. On examining the most frequently repeated verses of the

⁷⁶ As per the Tanzil.net downloaded document.

⁷⁷ Based on the splitting of the Arabic text into words via NLTK in Python.

Qur'an,⁷⁸ the researcher observed that words referring to natural phenomena occur in 23 out of the 95 most frequent verses of the Qur'an (See Table 6).

Table 6: Nature in the most frequently repeated verses of the Qur'an

<i>Verse</i>	<i>Frequency</i>	<i>Translation</i> ⁷⁹	<i>Word/s</i>
ويل يومئذ للمكذبين	10	Woe, that Day, to the deniers.	day
وما أسألكم عليه من أجر إن أجري إلا على رب العالمين	5	And I do not ask you for it any payment. My payment is only from the Lord of the worlds.	worlds
فسجد الملائكة كلهم أجمعون	2	So the angels prostrated themselves, all of them together	spirit
إن المتقين في جنات وعيون	2	Indeed, the righteous will be among gardens and springs,	gardens-springs
إلى يوم الوقت المعلوم	2	Until the Day of the time well-known.'	day
في جنات النعيم	2	In the Gardens of Pleasure	The Garden (Heaven)
فإذا سويته ونفخت فيه من روحي فقعوا له ساجدين	2	So when I have proportioned him and breathed into him of My (created) soul, then fall down to him in prostration.'	spirit
كلوا واشربوا هنيئًا بما كنتم تعملون	2	(Being told), 'Eat and drink in satisfaction for what you used to do.'	food-drink in Heaven
في جنة عالية	2	In an elevated garden,	The Garden (Heaven)
فهو في عيشة راضية	2	So he will be in a pleasant life -	life
سيح لله ما في السموات وما في الأرض وهو العزيز الحكيم	2	Whatever is in the heavens and whatever is on the earth exalts Allah, and He is the Exalted in Might, the Wise.	earth-sky
فأخذتهم الرجفة فأصبحوا في دارهم جاثمين	2	So the earthquake seized them, and they became within their home (corpses) fallen prone.	earthquake
فألقي عصاه فإذا هي ثعبان مبين	2	So Moses threw his staff, and suddenly it was a serpent, manifest.	staff (stick)
ونزع يده فإذا هي بيضاء للناظرين	2	And he drew out his hand; thereupon it was white for the observers.	hand-white
متاعا لكم ولأنعامكم	2	As provision for you and your grazing livestock.	cattle
ذلك بما قدمت أيديكم وأن الله ليس بظلام للعبيد	2	That is for what your hands have put forth and because Allah is not ever unjust to (His) servants.'	hand
ولا يحض على طعام المسكين	2	And does not encourage the feeding of the poor.	food
قل إنني أخاف إن عصيت ربي عذاب يوم عظيم	2	Say, 'Indeed I fear if I should disobey my Lord, the punishment of a tremendous Day.'	day
يا بني إسرائيل اذكروا نعمتي التي أنعمت عليكم وأني فضلتكم على العالمين	2	O Children of Israel, remember My favour that I have bestowed upon you and that I preferred you over the worlds.	worlds
في جنات وعيون	2	Within gardens and springs	gardens-springs
وأمطرنا عليهم مطرا فساء مطر المنذرين	2	And We rained upon them a rain (of stones), and evil was the rain of those who were warned.	rain
ثم أغرقنا الآخرين	2	Then We drowned the others.	drowning
الذين يقيمون الصلاة ويؤتون الزكاة وهم بالأخرة هم يقوتون	2	Who establish prayer and give zakat, and they, of the Hereafter, are certain (in faith).	resurrection

78 See also : <http://www.quraananalysis.com>

79 (Ali, 1987)

Based on this observation, it can be said that nature presents approximately a quarter of the most frequent and recurrent concepts (themes) in the Qur'an. In addition, the close reading of the words of the Qur'an and its translation required that NLTK was run against the two corpora to compute a set of individual tokens⁸⁰ of the texts to manually mark the occurrences of the terms in the list of natural phenomena as seen in the code below.

```
from nltk.tokenize import sent_tokenize, word_tokenize
print (word_tokenize (data))
```

2.1.3 The contextual meanings of nature as a theme in the Qur'an

Several previous studies show that the theme of natural phenomena is recurrent in the Qur'an (Golshani, 1986; Yahya, 2003; Fazlur Rahman, 2009; Mohamed, 2014; Brierley et al., 2018). Based on the qualitative analysis of the previously described CA, the researcher found that the portrayal of this theme is related to the central theme of the Qur'an and is in three primary contexts as follows: *natural signs* to glorify God's creation of the world; *portent signs* or *historical signs* representing an interruption of the natural order in present life punishment (e.g., floods, hurricanes, violent earthquakes, torrential rains falling where there usually is little or no rain) and the calamities of the Day of Judgment; and *super-natural signs* which appear in the narration of miracle stories (See Figure 6).

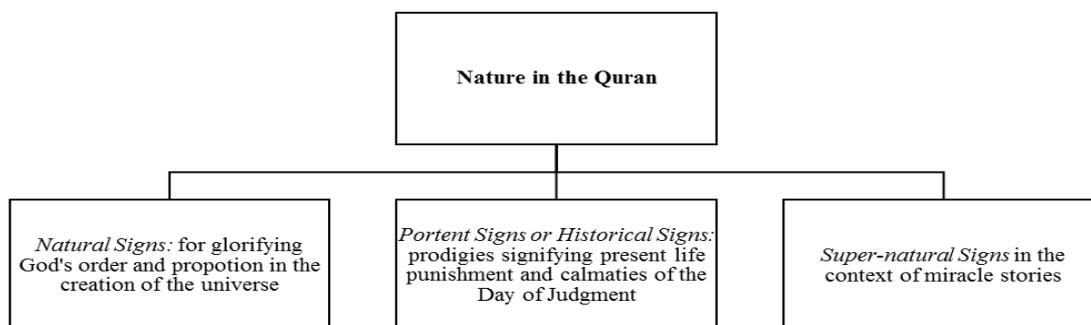


Figure 6: The theme of nature in the Qur'an⁸¹

To elaborate on the above, this subsection will give a brief account of some of the previous comments on the presence of nature as a predominant theme in the Qur'an. In *Major*

⁸⁰ A *token* is the smallest unit that each corpus divides to. Typically, each word form and punctuation (comma, dot, etc.) is a separate token. A text is divided into tokens by a tool called a tokenizer which is often specific for each language. *Corpus Linguistics- A Short Introduction*. Available from: [https://wmtang.org/corpus-linguistics/corpus-linguistics/], [Accessed 04 May 2019].

⁸¹ Summarised from the previous literature such as: Golshani, 1986; Bell et al., 1991; Robinson, 1999; Yahya, 2003; Fazlur Rahman, 2009; Al-Qinai, 2012; Nöldeke et al. 2013; Mohamed, 2014; and Brierley et al. , 2018.

Themes of the Qur'an, a chapter is devoted to 'Nature' as a significant theme (Fazlur Rahman, 2009). Fazlur Rahman (2009) states that "Qur'anic cosmogony" is minimal. However, even with it being so little, there are frequent allusions to nature. He writes:

But if the Qur'an has little to say about cosmogony, it makes frequent and repeated statements about nature and natural phenomena, though these statements invariably relate nature to God or to man, or to both (p.65).

To put it differently, there are not numerous references to the origin of the universe and solar and earth-moon systems in the Qur'an. However, natural phenomena are frequently seen throughout the Qur'an. The Qur'anic verses which have allusions to natural phenomena relate them to God, man, or to both. Furthermore, the linguistic representation of nature in the Qur'an is a combination of "natural causation language" and "divine causation" or "religious language" in different contexts, and clearly with varying purposes in view. For example, the most fundamental disparity between God and His creation (nature) is that whereas God is infinite and absolute, every creature (natural phenomenon) is finite. All things have potentialities, but no amount of potentiality may allow what the Qur'an means when it says that everything except God is "measured out" (*qadar* or *qadr*, *taqdir*, etc.). In other words, everything is dependent upon God.

When God creates anything, He places within its powers or laws of behaviour, called in the Qur'an 'guidance', 'command', or 'measure' whereby it fits into the rest of the universe: 'He gave everything its creation then guided (it)' (Fazlur Rahman, 2009, p.67).

Furthermore, the Qur'an mentions natural phenomena because they are *signs* of God's ultimate power and command over the universe. They serve to prove that God is the Master of all creations; hence, all humans must believe in Him: "The gigantic machine, the universe, with all its causal processes, is the prime *natural sign* or *proof of its Maker*" (Fazlur Rahman 2009, p.68, *italics in original*).

Similarly, natural phenomena are referred to as *signs* in two studies that outline the historical background and themes of the Qur'an (Bell et al., 1991; Robinson, 1999, p. 28, p.78, and p.83). In this sense, Bell et al. (1991) suggest that the composition of the Qur'an falls into three main periods, in one of which nature is a component as follows:

1. An early period: this consists mainly of lists of natural phenomena, which are 'signs' of Allah's power and generosity and exhortations to worship Him.

2. The 'Qur'an' period: this began with the institution of the prayer ritual and covered the latter part of Muhammad's activity in Makkah and the first year or two of his residence in Medina.
3. The 'Book' period: this began around the end of the second year of the 'Hijrah' (after the break with the Jews and the victory at Badr)⁸² (Bell et al., 1991, p.6).

Moreover, in relation to the notion of natural phenomena as *signs*, Robinson (1999) claims that fifteen of the early Makkan surahs include nature terms with contextual environments that are associated with the Creator's power and beneficence. These surahs mention the *sun* and *moon*, as seen in the previously mentioned example, sometimes stressing that God made them pursue a regular course, enabling humans to calculate the passage of time. Similarly, the *stars* that help humans navigate are mentioned, and references to *day* and *night* are frequent. In the repeated occurrences of *day* and *night*, the stress is on God, causing this regular alteration, which affords humans opportunities to sleep and to seek their livelihood. There is also one reference to God causing the dawn to break, and several to the *sun* and *moon* and *stars* to bow to their Creator. In the contexts mentioned above of nature in the Qur'an, it "labours these points to transform the consciousness of the pagan Arabs (disbelievers) who worshipped the heavenly bodies" (Robinson, 1999, p.152).

Furthermore, Nöldeke et al. (2013) claim that the style of the Qur'an is in large parts prosaic but shows "poetic power" when it refers to descriptions of Heaven and Hell and allusions to God's working in *Nature* (pp. 106-07). Another observation about the stylistic function of natural phenomena is seen in their appearance in metaphors in the Qur'an to perform the following functions:

Metaphors of natural phenomena fulfil two opposing functions. On the one hand, metaphors act as heralds of goodness for those who believe in God and have strong faith in Him, His messages and His messengers. On the other hand, they act as a form of punishment for those who disbelieve in Him and deny His messages and His messengers (Mohamed, 2014, p.23).⁸³

Furthermore, Al-Qinai (2012) comments that the *earth*, *heavens*, animate and inanimate objects (birds, trees, mountains, thunder) are treated in the Qur'an as living beings that praise God and understand when spoken to by God and his messengers (p.84). The

⁸² *Battle of Badr*, (624 CE), in Islamic history, first military victory of the Prophet Muhammad. It damaged Meccan prestige while strengthening the political position of Muslims in Medina and establishing Islam as a viable force in the Arabian Peninsula. Available from: [<https://www.britannica.com/event/Battle-of-Badr>], [Accessed 15 December 2016].

⁸³ Available from: [<http://eujournal.org/index.php/esj/article/viewFile/7441/7167>], [Accessed 15 December 2016].

examples below illustrate the Qur'anic description of inanimate phenomena such as *mountains* and *thunder* as animate ones glorifying and praising God.

فَفَهَّمْنَاهَا سُلَيْمَانَ وَكَلَّمْنَا هَارُونَ وَخَلَقْنَا لَهُ مِنْ هَاجِلِهِ ذُرِّيَّةً وَأَوْفَىٰ بَعْدَ ذَلِكَ أَيْمَانَ
وَالطَّيْرُ وَكَانَ فَلَجًا
﴿٧٩﴾

Yusuf Ali: To Solomon, We inspired the (right) understanding of the matter: to each (of them) We gave Judgment and Knowledge; it was Our power that made the hills and the birds celebrate Our praises, with David: it was We Who did (all these things).
Verse (21:79)

إِنَّا سَخَّرْنَا الْجِبَالَ مَعَهُ يُسَبِّحْنَ بِالْعُشِيِّ وَالْإِشْرَاقِ
﴿١٨﴾

Saheeh International: Indeed, We subjected the mountains (to praise) with him, exalting (Allah) in the (late) afternoon and (after) sunrise.
Verse (38:18)

وَيُسَبِّحُ الرَّعْدُ بِحَمْدِهِ وَالْمَلَائِكَةُ مِنْ خِيفَتِهِ وَيُرْسِلُ الصَّوَاعِقَ فَيُصِيبُ بِهَا مَن يَشَاءُ وَهُمْ يُجَادِلُونَ فِي اللَّهِ وَهُوَ شَدِيدُ الْمِحَالِ
﴿١٣﴾

Saheeh International: And the thunder exalts (Allah) with praise of Him - and the angels (as well) from fear of Him - and He sends thunderbolts and strikes therewith whom He wills while they dispute about Allah, and He is severe in assault.
Verse (13:13)

From another viewpoint, namely a scientific one, and in relation to the above context of glorifying God, Golshani (2003) claims that more than 750 Qur'anic verses are dealing with natural phenomena, and these involve a vital message for man (pp.8-9). The following are three essential points of that message with several examples from the Qur'an and their translations (Golshani, 2003, pp.10-8):

- 1) In the verses containing references to natural phenomena, the study of all aspects of nature and the discovery of the mysteries of the creation is recommended.

وَفِي خَلْقِكُمْ وَمَا يَبُذُّ مِنْ دَابَّيْهِ آيَاتٌ لِّقَوْمٍ يُوقِنُونَ
﴿٤﴾

Saheeh International: And in the creation of yourselves and what He disperses of moving creatures are signs for people who are certain (in faith).
Verse (45:4)

قُلْ انظُرُوا مَاذَا فِي السَّمَوَاتِ وَالْأَرْضِ وَمَا تُعْزِي الْأَبْنُوتُ وَالنُّذُرُ عَنْ قَوْمٍ لَا يُؤْمِنُونَ
﴿١١﴾

Saheeh International: Say, 'Observe what is in the heavens and earth.' But of no avail will be signs or warner to a people who do not believe. Verse (10:101)

قُلْ سِيرُوا فِي الْأَرْضِ فَانظُرُوا كَيْفَ بَدَأَ الْخَلْقَ ثُمَّ اللَّهُ يُنشِئُ النَّشْأَةَ الْآخِرَةَ إِنَّ اللَّهَ عَلَىٰ كُلِّ شَيْءٍ قَدِيرٌ ﴿١٠١﴾

Saheeh International: Say, (O Muhammad), 'Travel through the land and observe how He began creation. Then Allah will produce the final creation. Indeed Allah, over all things, is competent.' Verse (29:20)

According to the Holy Qur'an, man can get access to cognition of nature if he uses his senses and intellect. In this regard, a scholar like Al-Biruni,⁸⁴ who believed the universe could not be eternal giving as evidence the terrestrial changes in the rock and land surface due to natural forces like wind, water and fire, explicitly stated that the incentive behind his research in scientific fields is Allah's words in the Qur'an which persuade human beings to think about the creation of the heaven and the earth, a creation which is purposeful and not vain (as cited in Douglas, 1973, p. 209):

الَّذِينَ يَذْكُرُونَ اللَّهَ قِيَمًا وَقُعُودًا وَعَلَىٰ جُنُوبِهِمْ وَيَتَفَكَّرُونَ فِي خَلْقِ السَّمَوَاتِ وَالْأَرْضِ رَبَّنَا مَا خَلَقْتَ هَذَا بَطْلًا سُبْحَانَكَ فَقِنَا عَذَابَ النَّارِ ﴿١٣١﴾

Saheeh International: (Those) who remember Allah while standing or sitting or (lying) on their sides and give thought to the creation of the heavens and the earth, (saying), 'Our Lord, You did not create this aimlessly; exalted are You (above such a thing); then protect us from the punishment of the Fire. Verse (3:191)

- 2) Verses with references to natural phenomena state that everything in this world is orderly and purposeful and that there is no fault in the works of Allah:

الَّذِي لَهُ مُلْكُ السَّمَوَاتِ وَالْأَرْضِ وَلَمْ يَتَّخِذْ وَلَدًا وَلَمْ يَكُن لَّهُ شَرِيكٌ فِي الْمُلْكِ وَخَلَقَ كُلَّ شَيْءٍ فَقَدَرَهُ بَدَلًا ﴿٢٥٢﴾

Saheeh International: He to whom belongs the dominion of the heavens and the earth and who has not taken a son and has not had a partner in dominion and has created each thing and determined it with (precise) determination. Verse (25:2)

84 Abu Riahan Muhammad bin Ahmed Al-Biruni (973-1048), a Muslim scholar, was perhaps the greatest scientist of his day and a productive scholar in many fields of knowledge. He was interested in astronomy, astrology, language, philosophy, religion, and literature, as well as, history and geography. See also (Douglas, 1973, p.209-10).

الَّذِي خَلَقَ سَبْعَ سَمَاوَاتٍ طِبَاقًا مَا تَرَى فِي خَلْقِ الرَّحْمَنِ مِنْ تَفَوُّتٍ فَارْجِعِ الْبَصَرَ هَلْ تَرَى
مِنْ فُطُورٍ ﴿٦٣﴾

Saheeh International: (And) who created seven heavens in layers. You do not see in the creation of the Most Merciful any inconsistency. So return (your) vision (to the sky); do you see any breaks? Verse (67:3)

وَرَى الْجِبَالِ تَحْسَبُهَا جَامِدَةً وَهِيَ كَمَثَرِ السَّحَابِ صُنِعَ اللَّهُ الَّذِي لَنْفَنَ كُلَّ شَيْءٍ إِنَّهُ خَبِيرٌ
بِمَا تَعْمَلُونَ ﴿٨٨﴾

Yusuf Ali: Thou seest the mountains and thinkest them firmly fixed: but they shall pass away as the clouds pass away: (such is) the artistry of Allah, who disposes of all things in perfect order: for he is well acquainted with all that ye do. (Verse 27:88)

- 3) The Qur'an invites man to recognise the laws of nature (i.e., Allah's patterns in the universe) and to exploit them for the welfare of human beings and without transgressing the limits of the Shari'ah.⁸⁵

الشَّمْسُ وَالْقَمَرُ بِحُسْبَانٍ ﴿٥﴾ وَالنَّجْمُ وَالشَّجَرُ يَسْجُدَانِ ﴿٦﴾

وَالسَّمَاءَ رَفَعَهَا وَوَضَعَ الْمِيزَانَ ﴿٧﴾ أَلَّا تَطْغَوْا فِي الْمِيزَانِ ﴿٨﴾

Saheeh International: The sun and the moon (move) by precise calculation, and the stars and trees prostrate. And the heaven He raised and imposed the balance. That you not transgress within the balance.

Verses (55:5-8)

It can be said accordingly that the scientific view of verses describing natural phenomena in the Qur'an is summarised in the following: the importance of the discovery of nature through using man's senses and intellect,⁸⁶ and that the Qur'an gives the correct world order with its perfection and balance.

Also, just as natural phenomena with the regularity in which they operate to serve as God's natural signs for humans, their course can be diverted, suppressed, or temporarily suspended in phenomena such as floods, hurricanes, violent earthquakes, or torrential

⁸⁵ The term *Shari'ah*, often translated as 'Islamic law', refers to commands, prohibitions, guidance and principles that God has addressed to mankind pertaining to their conduct in this world and salvation in the next. See also (Kamali, 2008, p.4).

⁸⁶ This view of nature influenced the scholars of the so-called 'Golden Age of Islam' to undertake scientific activities that resulted in the vast corpus of scientific works of that era (Faruqi, 2007, p.463).

rainfall where there is normally little or no rain. These distinctive signs are *portent* or *historical signs* often coming at some point when a people is irretrievably on a wicked course of action (Fazlur Rahman 2009, p.47). Furthermore, in the Qur'an, the course of nature is said to be suspended or diverted in these calamities to indicate the advent of the Day of Judgment. Examples of both cases respectively are shown in verses with the word *earth* as seen below:

أَفَلَمْ يَرَوْا إِلَى مَا بَيْنَ أَيْدِيهِمْ وَمَا خَلْفَهُمْ مِنْ السَّمَاءِ وَالْأَرْضِ إِنَّ نَسْأًا نَحْصِفُ بِهِمْ
الْأَرْضَ أَوْ نَسْقِطُ عَلَيْهِمْ كِسْفًا مِنَ السَّمَاءِ إِنَّ فِي ذَلِكَ لَآيَةً لِكُلِّ عَبْدٍ مُنِيبٍ ﴿٤١﴾

Pickthall: Have they not observed what is before them and what is behind them of the sky and the earth? If We will, We can make the earth swallow them, or cause obliteration from the sky to fall on them. Lo! herein surely is a portent for every slave who turneth (to Allah) repentant. Verse (39:4)

إِذَا رُجَّتِ الْأَرْضُ رَجًا ﴿٤١﴾

Saheeh International: When the earth is shaken with convulsion.

Verse (56:4)

The final context of nature in the Qur'an is seen in the presence of natural phenomena as *supernatural signs* in miracle stories. The Qur'an narrates several miracle stories about messengers. When Abraham is thrown into the *fire*, it miraculously becomes "cool and safe" for him (21:69); and Mary conceives Jesus as a virgin (3:45–8, 19:18–21). Similarly, Jesus heals the blind and sick with his touch, and revives the dead (5:110), while Solomon understands the language of the *birds* (27:16) and has miraculous means of transportation (34:12), and so on. Examples of these verses are shown below:

قُلْنَا يَا نَارُ كُونِي بَرْدًا وَسَلَامًا عَلَىٰ إِبْرَاهِيمَ ﴿٦٩﴾

Yusuf Ali: We said, 'O Fire! be thou cool, and (a means of) safety for Abraham!'

Verse (21:69)

وَوَرِثَ سُلَيْمَنُ دَاوُدَ وَقَالَ يَا أَيُّهَا النَّاسُ عُلِّمْنَا مَنْطِقَ الطَّيْرِ وَأَوْتِينَا مِنْ كُلِّ شَيْءٍ إِنَّ هَذَا
هُوَ الْفَضْلُ الْمُبِينُ ﴿١٦﴾

Saheeh International: And Solomon inherited David. He said, 'O people, we have been taught the language of birds, and we have been given from all things. Indeed, this is evident bounty.'
Verse (27:16)

To conclude, the exploratory account of the previous studies on nature in the Qur'an has provided the needed ground of understanding to commence the scrutinisation of the initially compiled list [nature terms for the CA of nature contextual meanings] to, later on, conduct the corpus-based methodology of this research. This was done with bearing in mind their recurrences and the meanings with which they associate and/or their role in the formulation of the central theme of the Qur'an.

2.1.4 The compiled working list of natural phenomena in the Qur'an

After examining the presence of nature in the Qur'an, a bilingual working list of 154 words denoting natural phenomena was formed, then verified against the words and lists found in the previous literature. It should be mentioned that this list includes both the singular and plural forms as well as recurrent words such as *sign*, *signs*, *creation*, and *creations* which are used to refer to natural phenomena abundantly in the Qur'an. The following is the working list of terms describing natural phenomena in this research.

1. Angel (الملائكة)	39. Ear (أذن)	78. Hoopoe (هدهد)	117. Food (طعام)
2. Ant (نمل)	40. Earth (أرض)	79. Herb (بقل)	118. Clay (صلصال)
3. Ape (قرد)	41. Dog (كلب)	80. Honey (عسل)	119. Heaven (سماء/ الجنة)
4. Beast of the Earth (دابة الارض)	42. Dust (تراب)	81. Heart (قلب)	120. Wind (ريح)
5. Zaqquq Tree (زقوم)	43. Resurrection (القيامة)	82. Heel (كعب)	121. Mud (طين)
6. Bee (نحلة)	44. Desert (صحراء)	83. Hell (جهنم)	122. Sea (بحر)
7. Jinn (الجن)	45. Date Palm (نخلة)	84. Gold (ذهب)	123. Water (ماء)
8. Wolf (ذئب)	46. Constellation (البروج)	85. Grain (حبة)	124. Mountain (جبل)
9. Valley (واد)	47. Coral (مرجان)	86. Grape (عنب)	125. Breasts (صدور)
10. Tree (شجرة)	48. Crow (غراب)	87. Green (أخضر)	126. Earthquake (زلزال)
11. Tongue (لسان)	49. Colour (لون)	88. Eden (جنات عدن)	127. Man (رجل-إنسان)
12. Thunder (رعد)	50. Clay (صلصال)	89. Garlic (ثوم)	128. Woman (امرأة)
13. Sun (الشمس)	51. Cloud (سحاب)	90. Glass (زجاج)	129. Waves (امواج)
14. Star (نجم)	52. Camel (جمل)	91. Goat (ماعز)	130. Storm (عاصفة)
15. Spider (عنكبوت)	53. Bird (طائر)	92. Ginger (زنجبيل)	131. Shower (مطر)
16. Soil (ثرى)	54. Boat (سفينة)	93. Cow (بقرة)	132. Light (نور)
17. Snake (حية)	55. Bone (عظم)	94. Mankind (بشر)	133. Darkness (ظلام/ظلمات)
18. Sirius (الشعرى)	56. Soul (نفس)	95. Finger (أصابع)	134. Day (يوم)
19. Silver (فضة)	57. Body (بدن)	96. Heart (قلب)	135. Dawn (فجر)
20. Silk (حرير)	58. Onion (بصل)	97. Blood (دم)	136. Stone (حجارة)
21. Ship/ark (سفينة)	59. Olive (زيتون)	98. Hand (يد)	137. Rock (صخر)
22. Sheep (غنم)	60. Oil (زيت)	99. Face (وجه)	138. Stick/rod (عصا)
23. Saqar (سقر)	61. Night (ليلة/ليل)	100. Head (رأس)	139. Land (بر)
24. Salsabil (سلسبيل)	62. Eagle (نسر)	101. Leg (ساق)	140. River (نهر)
25. Ruby (ياقوت)	63. Mule (بغل)	102. Steed (خيل)	141. Cold (برد)
26. Rain (مطر)	64. Mount (طور)	103. Fruit (ثمر)	142. Hot (حر)
27. Pomegranate (رمان)	65. Mosquito (بعوضة)	104. Whale (حوت)	143. Shade (ظل)
28. Pig (خنزير)	66. Moon (القمر)	105. Garden (جنة)	144. Flood (طوفان)
29. Pearl (لؤلؤ)	67. Milk (لبن)	106. Fish (سمك)	145. Hour (ساعة)
30. Paradise (الجنة)	68. Louse (القمل)	107. Sky (سماء)	146. People (ناس)
31. Firdous (الفردوس)	69. Loti Tree (السدره)	108. Morning (صباح)	147. Fountains (عيون)
32. Fly (ذباب)	70. Locust (جرادة)	109. Evening (مساء)	148. Fire (نار)
33. Forehead (جبهة)	71. Lion (أسد)	110. Life (حياة)	149. Foot (قدم)
34. worlds (عالمين)	72. Lightning (برق)	111. Death (موت)	150. Skin (جلد)
35. Year (عام)	73. Lentil (عدس)	112. reast (صدر)	151. Sign (آية)
36. Frog (ضفدع)	74. Leaf (ورقة)	113. Men (رجال)	152. Signs (آيات)
37. Embryo (أجنة)	75. Flame (لظى)	114. Women (نساء)	153. Create (يخلق)
38. Eye (عين)	76. Iron (حديد)	115. Tilth (حصاد)	154. Creation (خلق)
	77. Ink (حبر)	116. Cattle (أنعام)	

Moreover, it should be acknowledged that the *Qur'anic Arabic Corpus Topics*⁸⁷ is an excellent resource to check the accuracy of the linguistic representation of some of the terms (the shaded words in the list above were also found in the Arabic Qur'anic list of topics). Hence, this was the first step of performing a qualitative analysis, where a list of nature terms was compiled. The following step was to build an ontology of natural phenomena terms in the Qur'an.

2.2 The ontology of nature terms

The term *ontology*, a word coined in the early seventeenth century, is used with different senses in different communities. Two uses of this word are distinguishable as follows: a countable noun, as in *an ontology*, and an uncountable noun, as in *ontology*. In the first case, it refers to a special kind of information object or computational artefact and is a way to formally model the structure of a system, i.e. relevant entities and relations that emerge from its observation, and in which the existence of the object with its concepts is pragmatic (e.g. for Artificial Intelligence systems, what “exists” is that which can be represented) (Giaretta and Guarino, 1995, p.1). On the other hand, in the second case, where the existence of the object is not always necessary, ontology refers to a philosophical discipline, namely the branch in philosophy dealing with the nature and structure of “reality”. Aristotle famously discusses this subject in *metaphysics*⁸⁸ and defines it as:

... the science of being *qua* being, i.e. the study of attributes that belong to things because of their nature. Unlike experimental sciences, which aim at discovering and modelling reality under a certain perspective, ontology focuses on the nature and structure of things per se, independently of any further considerations, and even independently of their actual existence. For example, it makes perfect sense to study the ontology of unicorns and other fictitious entities: although they do not have an actual existence, their nature and structure can be described in terms of general categories and relations (as cited in Giaretta and Guarino, 1995, pp.1-3).⁸⁹

In the context of this research, the definition of the ontology of nature as a Qur'anic theme is a combination of both uses of this term. It focuses on the nature and structure of things (concepts of natural phenomena represented by words), and these things pragmatically exist,

87 Available from: [<http://corpus.quran.com/topics.jsp>].

88 *Metaphysics*, or what Aristotle himself called his ‘first philosophy’, involves a study of the universal principles of being, the abstract qualities of existence itself. Available from: *The Stanford Encyclopaedia of Philosophy*, [<https://seop.illc.uva.nl/entries/aristotle-metaphysics/>],[Accessed 10 July 2017].

89 Available from: [<https://userpages.uni-koblenz.de/~staab/Research/Publications/2009/whatIsAnOntology.pdf>], [Accessed 10 July 2019]

whether in the present life (e.g. weather and geographical location) or, as claimed by the Qur'an, in the afterlife (e.g. flowing rivers in the reward of Paradise and the blazing fire as the afterlife punishment). This section introduces the proposed ontology of nature terms in this research. It was assembled by examining the available ontologies of concepts in the Qur'an to select the appropriate relevant network of meanings relevant to nature as a theme, then using the compiled working list in the previous section; the ontology was built to represent nature as a Qur'anic theme. It is an ontology that borrows concepts from the previous work on Qur'anic ontologies and is inclusive of the 154 natural phenomena terms that were compiled in this preliminary stage of the research.

2.2.1 The available ontologies of concepts in the Qur'an

It was found that numerous studies have been conducted to create semantic ontologies for the Qur'an (e.g. Saad et al., 2010; Dukes, 2011; Hammo et al., 2012; Muhammad et al., 2012; Mukhtar et al., 2012; Yauri et al., 2012; Abbas et al., 2013; Ahmad et al., 2013; Alrehaili and Atwell, 2014; Abed, 2015; Alromima et al., 2015; Ismail et al., 2016; Ta'a et al., 2013; Alqahtani and Atwell, 2018; Nawi, 2018). A survey of the recent Qur'an ontology research projects concluded that some of the ontologies built for the Qur'an are incomplete and/or focused on a specific, limited domain (Alrehaili and Atwell, 2014).⁹⁰ Besides, Alrehaili and Atwell (2014, p.26) found that there is no clear agreement on the semantic annotation format,⁹¹ technology to be used, or how to verify or validate the results. Some examples of these studies are discussed below with an indication of the portions of the ontology adopted for this study.

In an attempt to classify the Qur'an into retrievable information, a study was conducted to present a mining technique named the *Qur'an Mining Technique* (QMT) (Hammo et al., 2012). It relies on the following predefined set of 10 themes for the Qur'an (p.191):⁹²

- T1. Human and Ethical Relations
- T2. Human and Social Relations
- T3. Working
- T4. Faith

90 Available from: [<http://eprints.whiterose.ac.uk/78272/1/alrehaili14lrerel2v4%20Camera-Ready.pdf>], [Accessed 05 in July 2017].

91 In another study, Alqahtani and Atwell (2015), claim that there is no agreement in the techniques used to retrieve information from the Qur'an based on annotation; they can be semantic-based or keyword-based. Semantic-based search techniques are concept-based and retrieve results by matching the contextual meaning of terms as they appear in a user's query, whereas Keyword-based search techniques return results according to the letters in the word(s) of a query (as cited in Ahmad et al., 2016).

92 The themes were identified by Al-Kabi et al. (2005).

- T5. Judicial and Law Relations
- T6. Organising Financial Relations
- T7. Al-Jihad⁹³
- T8. Science and Art
- T9. Holy Qur'an
- T10. Unseen World

Moreover, Abbas (2009a) is another study conducted to build a Qur'an corpus augmented with a conceptual ontology, which is "imported" from the recognised expert source: *Mushaf Al Tajweed* by Habash (2001).⁹⁴ To build a computational conceptual ontology, Abbas (2009a, pp.7-8) borrows Habash's classification of the Qur'an into the following 15 concepts:⁹⁵

1. Pillars of Islam
2. Faith
3. The Call for Allah
4. The Holy Qur'an
5. Jihad
6. Action (Work)
7. Man, and Moral Relations
8. Man, and Social Relations
9. Organising Financial Relationships
10. Trade, Agriculture, Industry, and Hunting
11. Judicial Relationships
12. General and Political Relationships
13. Science and Art
14. Religions
15. Stories and History

On a less comprehensive level, Mukhtar et al. (2012) produce a dataset that contains concepts from only the second chapter of the Qur'an, known as the *Vocabulary of Qur'anic Concepts*. The researchers here used six different English translations of the Qur'an and applied a domain-independent tool called *Termine* to extract the concepts. However,

93 From the Arabic word *jihād*, literally 'effort', expressing, in Muslim thought, struggle on behalf of God and Islam. It is also defined as a struggle or fight against the enemies of Islam. Also, a greater jihad is defined as the spiritual struggle within oneself against sin. Available from: [<https://www.lexico.com/en/definition/jihad>]

94 Dr. Mohamed Habash (who is the director of the Islamic Studies Center in Damascus) compiled *Mushaf Al Tajweed*, which is a version of the Quran that shows readers how to read the Quran correctly. 'Mushaf' is another *name* for the word 'Quran' book in Arabic and 'Al-Tajweed' refers to the rules for correctly reciting Quran in Arabic. He classifies the Quran into fifteen main themes/topics, covering nearly 1100 subthemes/subtopics, and arranges the verses in accordance with this classification. See also Habash (2001).

95 Habash (2001) as cited in Abbas (2009a, p.7).

Termine was designed for the extraction of multi-word terms, while the Qur'an has numerous single-word concepts (e.g., 'Allah' and 'Muhammad'). Another interesting yet not inclusive ontology was developed according to the themes of Al-Qur'an knowledge as described in *Syammil Al-Qur'an Miracle the Reference* (Ta'a et al.,2013). الإيمان *'al- 'īmān* 'faith' and الأخلاق *'al- 'aḥlāq* 'morals or ethics' are the main themes that were chosen as the research scope for constructing the ontology, and a hierarchy for these themes was developed as main concepts with other themes as sub-themes derived from the main themes. The concepts and properties produced were defined using Malay; hence, this ontology is not suitable for this study which focuses solely on the Arabic version and English translations of the Qur'an. Moreover, a suitable choice, had it been available, would have been the bilingual Arabic-English Qur'an ontologies from different datasets related to the Qur'an that is being developed by Alqahtani and Atwell (2018) but not yet made accessible.

Out of the various ontologies found, one ontology that has been scrutinised and appraised⁹⁶ is that of Dukes (2011) known as the *Qur'anic Arabic Corpus* (QAC) project.⁹⁷ This online Qur'an has been annotated at several levels and includes an ontology that defines 300 concepts in the Qur'an, and captures interrelationships using predicate logic. The number of relationships is 350, and the type of relationships between concepts is 'Is-a'. Dukes (2011) built this Qur'anic ontology of concepts (See Figure 7) based on the knowledge contained in traditional sources of Qur'anic analysis, including the sayings of the Prophet Muhammad ﷺ (translated as 'peace be upon him'), and Tafsīr books.

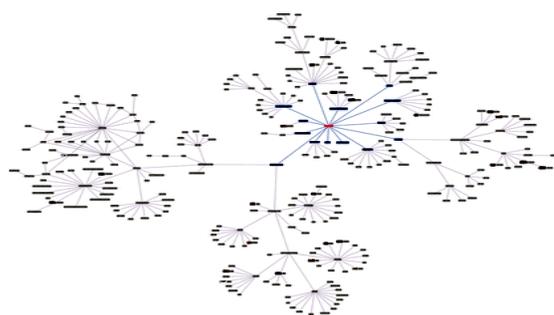


Figure 7: The Quranic Ontology. Each node in the graph represents a unique concept in the Quran (Figure taken from: <http://corpus.quran.com/ontology>)

⁹⁶ See also (Alrehaili and Atwell, 2014).

⁹⁷ Dukes, K. 2011. The Qur'anic Arabic Corpus. Online. Available from; [<http://corpus.quran.com>], [Accessed: 19 October 2016 onwards]

QAC also contains other analyses of the Qur'an, such as a morphological analysis and a dependency parse-structure⁹⁸ analysis. Finally, his work provides a semantic ontology which uses knowledge representation to define the key concepts in the Qur'an. These cover the following twelve categories:

1. Artefact⁹⁹
2. Astronomical Body
3. Event
4. False Deity
5. Holy Book
6. Language
7. Living Creations
8. Location
9. Physical Attribute
10. Physical Substance
11. Religion
12. Weather Phenomena

To sum up, the present research used the previous work on Qur'anic ontology to compile an ontology named 'Nature in the Qur'an' with concepts or classes adopted and composed partially from the existing QAC ontology developed by Dukes (2011) to classify the compiled list of nature terms specified in the previous section. Ontologies of natural phenomena in the Qur'an will be discussed further in Section 3.2.1, which provides an overview of the corpus linguistic studies conducted on the theme of nature in the Qur'an. The following subsection describes the method of partially adopting the QAC ontology of concepts and categorising the list of nature terms per the concepts in QAC with examples from both the Arabic and English lists.

2.2.2 The proposed conceptual ontology of nature terms

The proposed ontology of nature terms is based on the ontology of concepts in QAC (See Figure 8).

⁹⁸ *Dependency parsing* is an approach to automatic syntactic analysis of natural language inspired by the theoretical linguistic tradition of dependency grammar. The basic assumption underlying dependency grammar is the idea that syntactic structure essentially consists of words linked by binary, asymmetrical relations called dependency relations (or dependencies for short). A dependency holds between a syntactically subordinate word, called the *dependent*, and another word on which it depends, called the *head*. See also Kübler et al. (2009, pp.1-2).

⁹⁹ *Artefact*, sometimes spelled 'artifact' in American English, is an object that is made by a person, such as a tool or a decoration, especially one that is of historical interest. [<https://dictionary.cambridge.org/dictionary/english/artefact>]. [Accessed 04 July 2019]

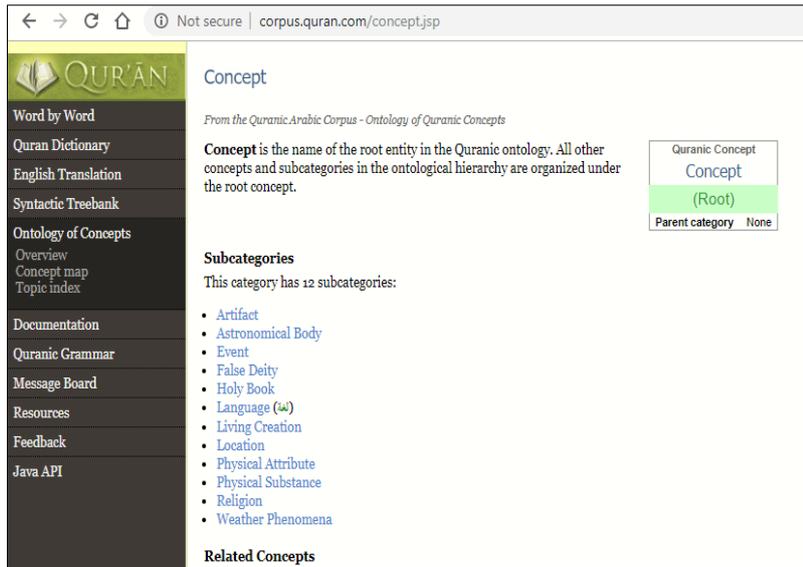


Figure 8: The ontology of concepts in the Qur'an

From the above classification in the Qur'anic Arabic Corpus, concepts of nature in the Qur'an were adopted and grouped as seen in the following figure:

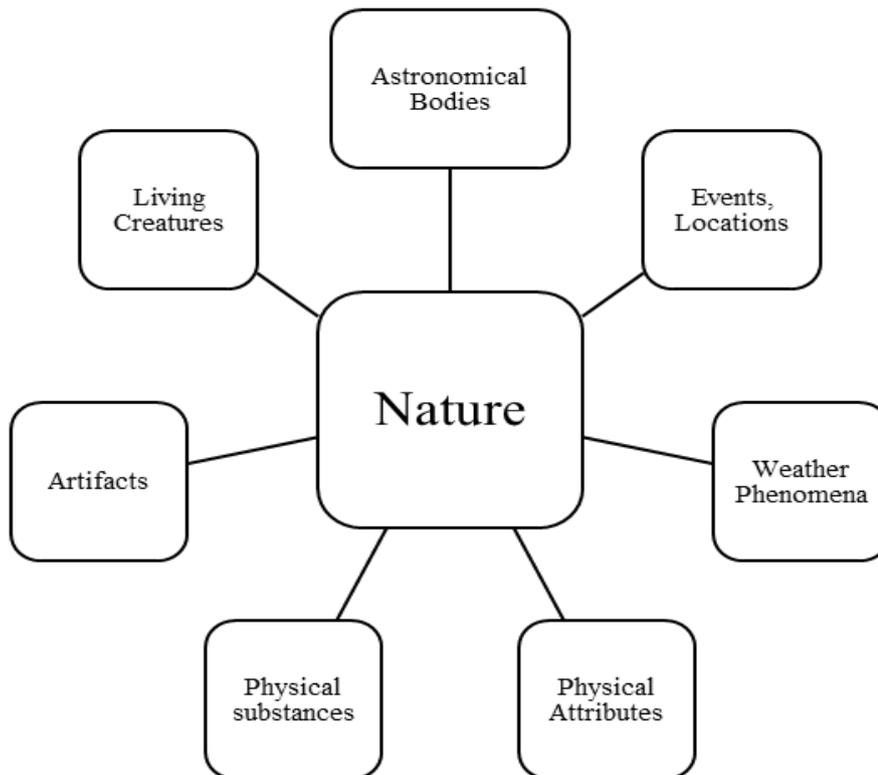


Figure 9: The concepts of nature in the Qur'an¹⁰⁰

100 Compiled based on [<http://corpus.quran.com/concept.jsp>].

A more elaborate visual representation of the concepts of nature in the Qur'an with the distribution of nature terms from the working list in the previous section is presented in the following figure:

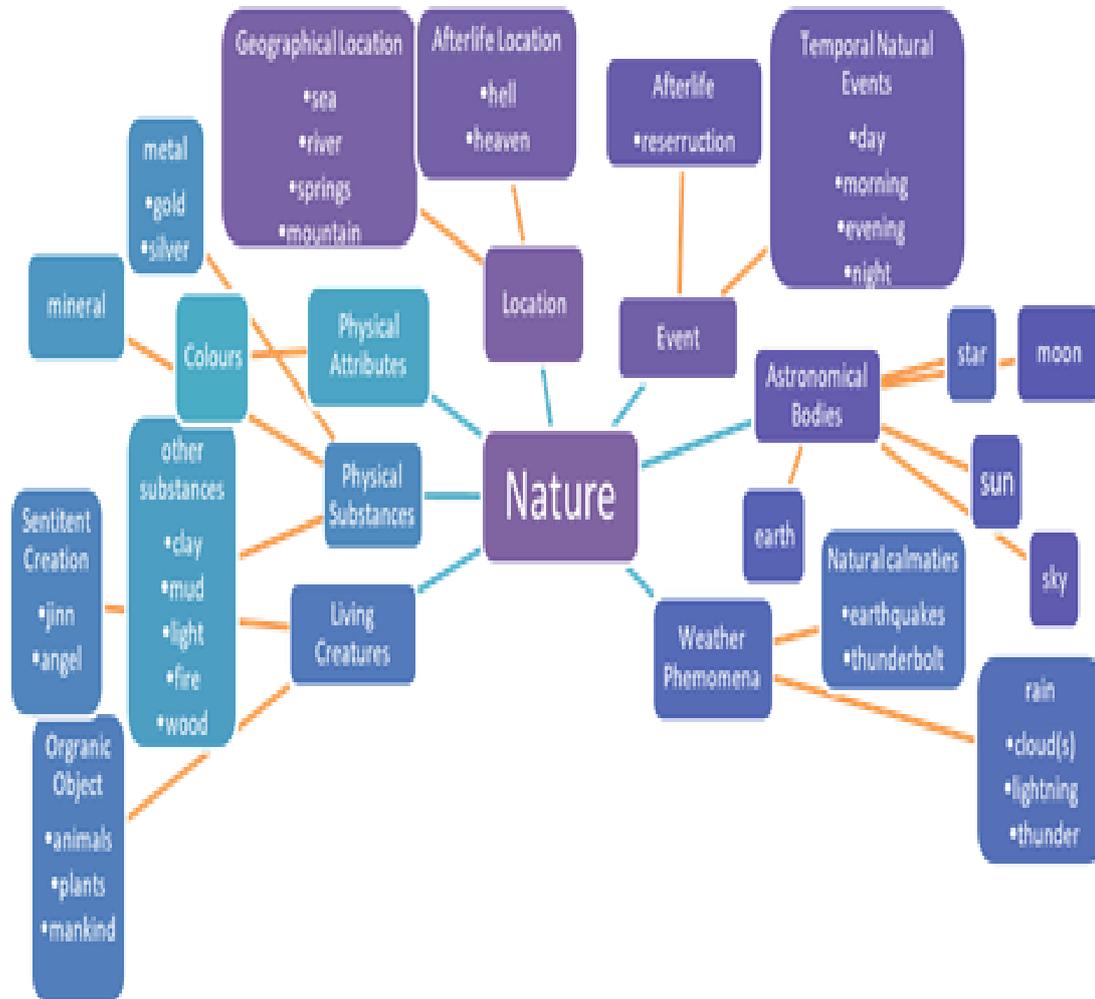


Figure 10: An elaborated ontology of nature in the Qur'an

Figure 10 shows that the established ontology of nature terms in this research, which is the outcome of the qualitative analysis discussed previously in Section 2.1, blends with Dukes's categorisation of Qur'anic concepts (2011) presented previously in the list of concepts in Figure 8. Moreover, it is clear that while the first list of natural phenomena [Section 2.1.4] is useful in providing examples for these phenomena, a more holistic approach is taken to produce the second list which contains more precise and easy-to-detect terms. These terms in Figure 10 were classified both in Arabic and English to be used in the exploration of SP of nature in the Arabic Qur'an and its five translations (See Table 7).

Table 7: The list of natural phenomena in the Qur'an and their semantic categories

<i>Semantic Category</i>	<i>Subcategory/s</i>	<i>Nature Term</i>	<i>English Tokens</i>	<i>Arabic Tokens</i>
Astronomical Bodies	astronomical objects	moon	moon	القمر/قمر
		star	star	نجم/النجم / النجوم
		earth	earth	الأرض/ أرض
		planet	planet/s	كوكب/ كواكب
		sky/s ¹⁰¹	sky/s, constellation, pieces of the sky	سماوات/سماء/ السماء/ السماوات/ المعارج/سقف/ كسفا
		the universe	worlds, universe, orbit	العالمين/ قلك (مسار)
Artefact	human-made construction	sun	sun	شمس/ الشمس/ سراجا
		ship	ship, boat, ark	الفلك/ السفينة/ذات ألواح/الجوار / مساكن/بيت/جدار
Events	writing	pen	pen, ink	القلم/ مداد
	afterlife event	day of resurrection	day of judgment, last day, day of resurrection	اليوم الاخر / الاخرة/يوم القيامة/الفرع/ الواقعة/ الساعة/ القارة/ الراجفة/ الرادفة/ الطامة/ الحطمة/ الصور
Physical Attributes	attributes of physical objects	temporal	day, dawn, night,	يوم /يوم فجر / ضحى/الغداة/الإصباح/ الصبح/بكرة/شفق/فلق/ظلام/الليل/الشروق / عشيا/ نور
		natural events	noon, morning daylight, light, darkness	بيضاء/أصفر /أحمر/أسود/أخضر
Physical Substances	naturally occurring substances	metal	brass, gold, iron, silver	نحاس/ذهب/فضة/ حديد
		fire	fire	نار/نور/شهاب
		mineral	ruby	ياقوت/كنز/لؤلؤ
		other	stone/s, rocks, dust, soil, oil, coral, pearl, silk, clay, glass, atom	حجارة/صخر/تراب/صلصال/طين/زيت/صعيدا/ز جاج/حرير / ذرة
Weather Phenomena	weather	weather	rain, lightning, thunder, storm, water, cloud/s	مطر /برق /رعد/صاعقة/ماء/سحاب/غمام/ودق / برد
		natural calamities	wind/s, tornado, earthquake, drowning, blast	صيحة/رياح/زلزال/عاصفة/ غرق/صاعقة/ قاصف/ خسف/ صيب
Living Creations	organic object	biological organisms	Animals, bird/s. Insect/s, fish, cow, calf, swine, apes, horses, wolf, donkey, snake, plant, garden, palm tree/s	حيوانات/ طيور/حشرات/عجل/خنزير/كلب/ناقة/ نباتات/جنات/جنة/نخل/ شجر/غنم
		body parts	ear, finger, forehead, forelock, heart, bone, tongue, right-hand, left- hand, throat, heal, wing, gut, stomach, wombs, blood	أذن(سمع) / أصبع/جباه/ناصية/قلب/عظم/لسان/بنان/صدر اليمين/الشمال/حلقوم/كعب/جناح/أرحام/دم/أعين) بصر(وجه/أرجل/جلد/قدم
		disease	leprosy, blind, mute, deaf	برص/اصم/بكم/عمي
		Food	meat, cattle, fruits	لحم/أنعام/فاكهة/ثمرات/طعام

101 's' next to a natural phenomenon term in this table and elsewhere in this thesis indicates that it represents both the singular and plural form of the word.

<i>Semantic Category</i>	<i>Subcategory/s</i>	<i>Nature Term</i>	<i>English Tokens</i>	<i>Arabic Tokens</i>	
Locations		humanity	soul/s, mankind, children of Adam, stages of human fetus, woman, man, people	نفس/بشر/نطفة/علقة/مضغة/اجنة/اطوار/حملا/ أمشاج/طفل/ذرية/ امرأة/نساء/رجل/رجال/الناس/امرء	
		sentient creations	other living entities	angel/s, devil/s, Satan, jinn, beast of the earth,	الملائكة/الشیطان/الجن/داية
		geographical location	township	town, city	مدينة/مدائن/قرية
		land	land, valley, desert	العراء/وادي/سبل/بر	
		waters	sea/s, river/s, wave/s	بحر/نهر/موج	
		east-west road	east, west road/s	مشرق/مغرب/مشارق/مغارب	
		mountain	mount, aljudi, almarwa, Arafat, Sinai, cave, spring	جبل/جبال/طور/كهف/غار/لحطمة/عين/ عيون/عرفات/الصفاء/المروة	
		afterlife location	Hell	hell, ladha, Sijeen, Zaqqum tree, saqar	جهنم/لظى/سقر/سجين/جحيم
			Heaven	firdous, garden of Eden, salsabil Paradise, springs	الجنة/فردوس/جنات/عدن/سلسبيل/الغرفة
	(Others)	life /death	life	life, death	حياة/موت/دنیا

As seen in the table above, eight¹⁰² of the concepts in QAC were borrowed to establish an ontology of nature-related concepts in this research. The compiled list of natural phenomena was arranged as per the ontology of concepts which is further categorised into the following semantic sub-categories or attributes:

1. astronomical objects
2. human-made construction
3. writing
4. afterlife event
5. physical event
6. attributes of physical objects
7. naturally occurring substances
8. weather phenomena
9. organic object
10. sentient creations¹⁰³
11. geographical location
12. afterlife location
13. life /death

102 The ninth is not in QAC and given the name 'Others' to include the miscellaneous terms in the compiled list of natural phenomena (e.g., life and death).

103 *Sentient Creation* is a category which refers to the angels, jinn, humanity and other sentient creations mentioned in the Qur'an. Available from: [<http://corpus.quran.com/concept.jsp?id=sentient-creation>].

computationally explored to find the SP of nature in the Qur'an as the primary aim of this thesis.

2.3 The disambiguation of nature terms

Bearing in mind Sinclair's views on lemmatization (1991, p.8; pp.41-2; pp.173-4), especially that each word form is "potentially a unique lexical unit, and that forms should only be conflated into lemmas when their environments show a certain amount and type of similarity", the researcher first computed the raw frequencies of word forms referring to nature in the Qur'an.¹⁰⁶ However, once these raw frequencies were put against their counterparts in Ali's translation, there was a variation especially in the number of word forms referring to living creatures (See Figure 11).

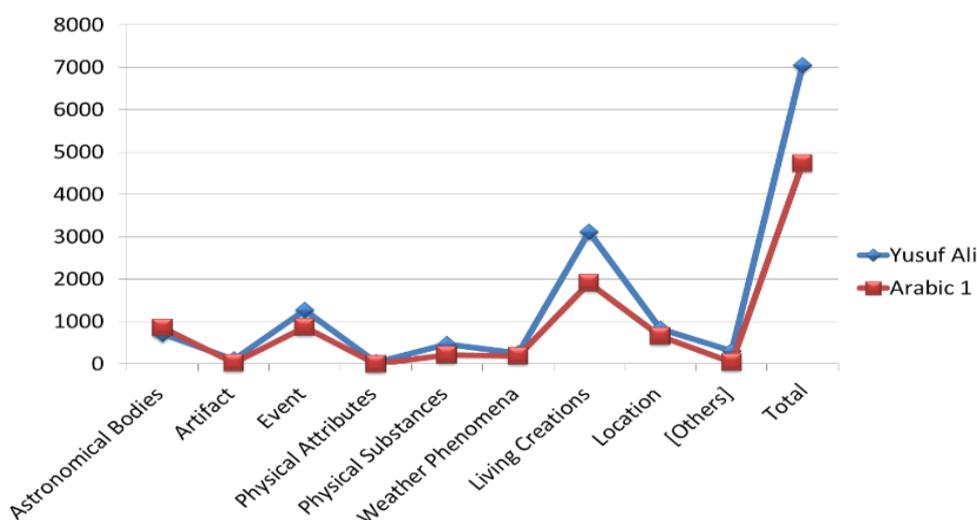


Figure 11 : A Comparison between raw frequencies of natural phenomena in the Arabic text and Ali's translation¹⁰⁷

A close examination of the raw counts revealed that the program did not capture some of the word forms, simply because they were not added to the aggregated initial list in Section 2.1.4. Hence, this was an issue of ambiguity that did in fact fall into the category of Sinclair's exception on lemmatization. When looking at the two lists in Figure 11, it was predicted that the missing word forms relating to nature were occurring and counted in the translation, and that they would have similar environments to the ones already counted in the Arabic Qur'an (as in Sinclair, 1991, p.8). This was proven to be true in a consecutive

¹⁰⁶ See also Alshahrani, H. and Brierley, C. 2018. The root-based disambiguation of words in the annotation of semantic prosody of nature as a Qur'anic theme. In: *The Workshop on Computational Approaches to Morphologically Rich Languages*, 3 July 2018, University of Leeds, Leeds.

¹⁰⁷ Counts of the semantic categories of natural phenomena. See also Section 2.2.2.

experiment of disambiguation and frequency counts as seen in the results shown Figure 12 below, where the counts were more similar. The following lines will elaborate on the ambiguity in the word forms in the Qur'an and how this issue was addressed in this research.

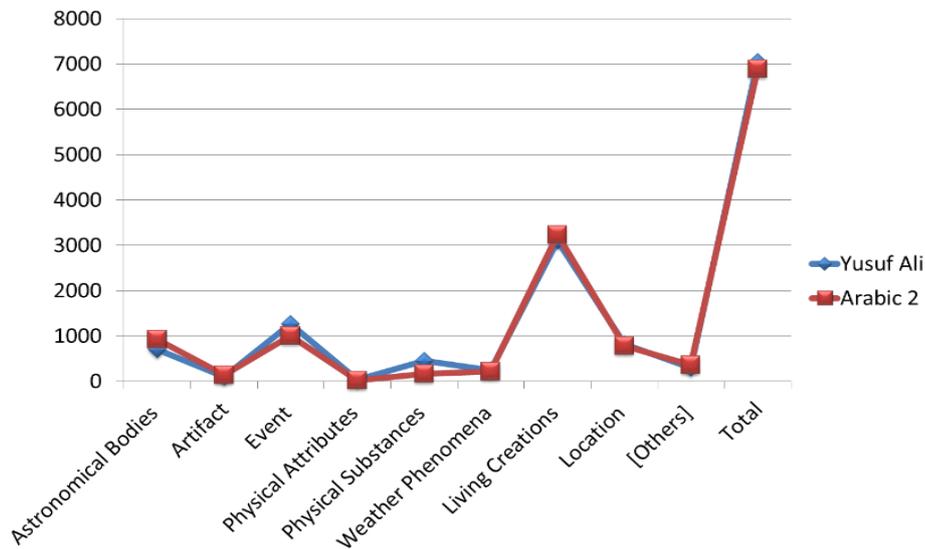


Figure 12: A comparison between frequency results after stemming (lemmatization)

The ambiguity of some words in the Qur'an can generally be caused by factors such as figurative language (outlined as an issue in translation with proposed solutions by Newmark, 1988, p.104) and polysemy;¹⁰⁸ internal word structure; syntactic, semantic, constituent boundary; and anaphoric ambiguity (See also Farghaly and Shaalan, 2009, p.9). For instance, it can be the origin of the variation in translation as explained in the following illustration of the difference in translations of the Qur'an:

The referential versatility of Qur'anic lexemes¹⁰⁹ undergoes semantic reduction when Qur'anic translators render a polysemy as monosomy. For example, the word *الفلق* *alflq* (113:1) is translated as 'daybreak' by Pickthall and Arberry, 'dawn' by Asad, Shakir and Ali whereas it can be generally used to refer to the process of splitting as in cell multiplication including seeds, molecules and biological entities (AbdulRaof, 2004 as cited in Al-Qinai 2012, p.82).

108 A *polysemic word* is a word that is capable of having several possible meanings, usually related by contiguity of meaning within a semantic field. Available from: [<https://www.collinsdictionary.com/dictionary/english/polysemic>], [Accessed July 2018].

109 *Lexeme*, a term sometimes used interchangeably with the term 'lemma' (as in Biber et al., 1999, p. 54), refers to a group of inflectional forms related to one stem that belong to the same word class (Kucera and Francis, 1967, p.1). It entails grouping together forms that have the same base and differ only with respect to aspects of grammar which do not affect word class or basic word sense – such as the singular and the plural forms of the same noun, the present and past tense of the same verb, the comparative and superlative forms of the same adjective (Brezina, 2018, p.40-1).

To address an issue as those mentioned above, which not only affects the translation of the Qur'an but also can affect the accuracy of NLP applications to find frequencies of words and extracting their collocations, word disambiguation of nature terms was used in two ways: stemming these terms to their smallest unit to reduce their various forms, and assigning a pre-defined sense to each of the versatile realisations (i.e., variety of forms) of each nature term. In this regard, following the step of compiling the bilingual list of nature terms in the Qur'an and establishing its ontology, it was necessary to disambiguate each of the terms to achieve accurate results when computing these words to find their frequencies, collocations, and SP. The various morphological forms of the same term in the text were reduced by stemming,¹¹⁰ and the terms are each recognised in terms of the main concept (i.e., the reduced form) which also includes the various semantic forms (i.e., synonyms). However, it should be noted that both parts of the exploration were primarily based on Ali's translation in the early stages of this research and not on the Arabic Qur'an, because it has been said that "the Arabic language still lacks many NLP resources available for a language like English (e.g., taggers, parsers, WordNet, FrameNet, etc.)" (Sharaf and Atwell 2012, p.131).¹¹¹The following subsection will describe the word disambiguation of the English terms in two NLP tasks via NLTK in Python: stem-based disambiguation by using the Porter Stemmer of English words, and word-sense disambiguation by using WordNet and the Lesk Method. The resulting reduced and unified forms are appended back into the text to produce more accurate results of the frequencies and collocations of nature terms.

2.3.1 Stem-based disambiguation

A clear definition of stem-based disambiguation is that it is a form of stemming that entails the following:

(It is) the process of stripping affixes (i.e., including both prefixes and suffixes) from words to form a stem, and it is often applied to words in Information Retrieval (IR)¹¹² systems so that

¹¹⁰ *Stemming* is a pre-processing step in Text Mining applications as well as a very common requirement of Natural Language Processing functions. The main purpose of stemming is to reduce inflectional forms and sometimes derivationally related forms of a word to a common base form (i.e., reducing different grammatical forms/ word forms of a word like its noun, adjective, verb, adverb etc. to its root form. See also (Jivani, 2011, p.1930).

¹¹¹ To go around this shortage, and suggest a solution for this issue for the Arabic Qur'an, this research proposes a linguistic approach of root-based disambiguation to normalise the different forms of the same terms in Arabic (e.g., synonyms from a reliable linguistic source, morphological and syntactic variations from QAC)(See Section 4.1.2).

¹¹² *Information retrieval* (IR) is finding material of an unstructured nature that satisfies an information need from within large collections. (Manning et al., 2010, pp.100-03).

words with almost the same meaning but superficial spelling differences are grouped together as the same concept (Wiese et al., 2011, p.496).

In NLTK, stemmers remove morphological affixes from words, leaving only the word stem. A sample of lines of stemming code via Python is seen below:

```
from nltk.stem import PorterStemmer  
ps= PorterStemmer ()
```

The 1980 Porter Stemmer, also known as Snowball in its developed version,¹¹³ is the rule-based stemmer¹¹⁴ chosen for this task because it is widely used for its simplicity and efficiency (Willett 2006; Wiese et al., 2011). The function of this stemmer is to separate affixes in five steps where inflectional affixes¹¹⁵ are handled in the first step; the derivational affixes¹¹⁶ are handled in the next three steps, and the recoding is conducted in the last step. To compute this task, two adopted stemming models by Pooja et al. (2016, pp. 7-9) were followed. They are the rule-based stemming model and the pre-processing model, as shown in the figures below. Figure 13 illustrates the steps of conducting the stemming of Ali's translation (as a list of words from which the stemmed natural phenomena were extracted), and Figure 14 presents an elaboration of the pre-processing task.

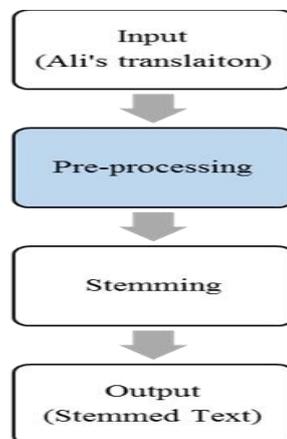


Figure 13: The rule-based stemming model

113 Porter (2001), available from: [<http://snowball.tartarus.org/texts/introduction.html>], [Accessed 15 March 2018].

114 *Rule-based stemmers* are algorithmic approaches, using minimal resources, that elaborate suffix replacement (or suffix removal) and conflate all the inflected forms of words (usually) in a single stem. (Karanikolas, 2015, p.315).

115 An *inflectional affix* is an affix that does not change the word class of its stem; typically located farther from its root than a derivational affix; and produces a predictable change of meaning. Available from: *Glossary of Linguistic Terms*. [<https://glossary.sil.org/>],[Accessed 04 July 2019].

116 A *derivational affix* is an affix by means of which one word is formed (derived) from another. The derived word is often of a different word class from the original. Available from: *Glossary of Linguistic Terms* [<https://glossary.sil.org/>],[Accessed 04 July 2019].

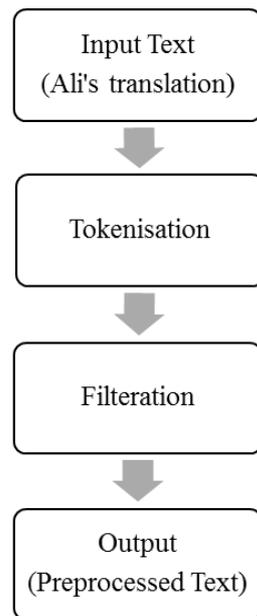


Figure 14: The pre-processing model

In pre-processing (as in Figure 14), the first step was to read in the text via Python as seen in the code below:

```
import nltk
import string
raw = open("engtrans-Ali.txt", 'r').read()
```

Then, *tokenisation*¹¹⁷ was performed followed by the cleaning of the text (e.g. removing: upper-case, punctuation, and stop-words¹¹⁸ such as *the*, *at*, *which*, and *on*). It is, as seen in Figure 13, a form of *filtration*, which produces a *pre-processed text*.

```
#tokenise and remove all upper-case
from nltk.tokenize import sent_tokenize, word_tokenize
tokens =nltk.word_tokenize (raw1)
raw2= [w.lower( ) for w in tokens]

#remove punctuation
punct= set (string.punctuation)
raw3= ''.join (w for w in raw2 if w not in punct)

#removing stop words
from nltk.corpus import stopwords
stop_words=set(stopwords.words("english"))
##add new stop words
stop_words.update(('ever', 'whereby', 'didst', 'wherever', 'whenever', 'shouldst', 'often', 'thereat', 'ot
herwise', 'although', 'oft', 'anywhere', 'whoever', 'wouldst', 'hast', 'manysoever', 'amid', 'specially', 'w
heresoever', 'never', 'wont', 'according', 'whatever', 'anything', 'behind', 'thereon', 'near', 'furthermor
e', 'midst', 'thither', 'towards', 'whilst', 'among', 'hadst', 'towards', 'mayest', 'wherewith', 'wherewith'
, 'wast', 'however', 'henceforth', 'thenceforth', 'thereby', 'couldst', 'could', 'dost', 'throughout',
'therefrom', 'thine',
'whereof', 'wherein', 'shalt', 'without', 'within', 'also', 'yet', 'yea', 'rather', 'whomsoever', 'till', 'th
```

117 *Tokenisation* is defined as: “the task of cutting a string into identifiable linguistic units that constitute a piece of language data” (Bird et al., 2009, p.109). [In Python coding, this is spelled with the [z] orthographic variation].

118 The rationale behind removing stopwords in this research will be discussed in Section 4.1.2.

```
erewith', 'every', 'would', 'whether', 'thus', 'another', 'must', 'therefore', 'beneath', 'forthwith', 'hath', 'may', 'except', 'might', 'doth', 'therein', 'unless', 'whereon', 'thus', 'amongst', 'ah', 'thee', 'thou', 'shall', 'ye', 'thy', 'thyselves', 'whose', 'nay', 'mere', 'perchance', 'till', 'oer', 'betwixt', 'should', 'unto', 'thereof', 'upon', 'either', 'moreover', 'us', 'besides', 'though', 'neither'))  
raw4= []  
for w in raw3:  
    if w not in stop_words:  
        raw4.append(w)
```

After that, going back to Figure 13, the text was stemmed via the Porter Stemmer as seen in the code below:

```
import nltk  
from nltk.stem import PorterStemmer  
ps= PorterStemmer ()  
for w in raw4:  
    print (ps.stem (w))
```

The following is a sample of the output of extracted nature terms from Ali's stemmed text:

```
heaven  
fruit  
stone  
garden  
river
```

In the previous list, we can see that words were stemmed to their most reduced forms. For example, the plural 's' in nature terms *heavens*, *rivers*, *stones*, and *gardens* is removed. However, there was only a 1.9% error rate¹¹⁹ relevant to the issue of reducing the irregular plurals in English such as *women* and *men*, and *children* to their correct singular forms. To resolve this issue, their stems were ascribed manually. Once the stemmed text was produced, the distinctive stem for each of the nature terms (i.e. nouns or verbs relevant to natural phenomena or nature events with no affixes) was examined and identified to be assigned to the synonyms of the natural phenomena that were also disambiguated as will be described in the following section.

2.3.2 Word sense disambiguation

In addition to the stem-based disambiguation of the different forms of natural phenomena in the English text, this same stem for each nature term was allocated to the semantic variations of each of the natural phenomena terms. In this sense, a search for the definition which best suits each term describing a natural phenomenon in the context of the Qur'an was essential, as was disambiguating its other forms via Word sense disambiguation (WSD). WSD is a task in natural language processing (NLP) referring to assigning the appropriate sense selected from a set of pre-defined senses (e.g. definitions and synonyms)

119 The error rate or percentage is calculated by number of natural phenomena words which are not assigned correct stems divided by 154 (the number of nature words in this research), and multiplied by 100.

for a polysomic word (e.g. a nature term in the Qur'an) and according to a context (e.g. their neighbouring words). WSD is necessary for many NLP applications and is believed to help improve their performance, such as information extraction, part of speech tagging, and text categorisation (Menai, 2014, p. 92). In the context of this research, only the synonymous words denoting natural phenomena that belong to the same definition were assigned the same stem and considered for statistical exploration. In addition, *WordNet*, a semantically oriented English dictionary with 155,287 words and 117,659 synonym sets (Bird et al., 2009, p.67), as well as the Lesk Method were used to provide the definitions and synonyms of the English nature terms in Ali's translation of the Qur'an. WordNet was used as an English dictionary to explore the different meanings of each of the nature terms and predict the most popular meaning that is recurrent in the Qur'an. A sample of lines of code for this process to define the word *earth* is shown in the following code.

```
from nltk.corpus import wordnet as wn
for ss in wn.synsets('earth'):
    print(ss, ss.definition())
```

Output:

```
Synset('earth.n.01') the 3rd planet from the sun; the planet we live on
Synset('earth.n.02') the loose soft material that makes up a large part of the land surface
Synset('land.n.04') the solid part of the earth's surface
Synset('earth.n.04') the abode of mortals (as contrasted with Heaven or Hell)
Synset('earth.n.05') once thought to be one of four elements composing the universe (Empedocles)
Synset('worldly_concern.n.01') the concerns of this life as distinguished from heaven and the
afterlife
Synset('ground.n.09') a connection between an electrical device and a large conducting body, such as
the earth (which is taken to be at zero voltage)
Synset('earth.v.01') hide in the earth like a hunted animal
Synset('earth.v.02') connect to the earth
```

Moreover, to disambiguate this word as it occurs in Ali's translation of the Qur'an, the Lesk Algorithm was used. The concept behind this process is explained in the words of its founder Michael Lesk as follows:

The meaning of an English word can vary widely depending on which sense is intended. *Does a fireman feed fires or put them out?* It depends on whether or not he is on a steam locomotive. I am trying to decide automatically which sense of a word is intended (in written English) by using machine-readable dictionaries, and looking for words in the sense definitions that overlap words in the definition of nearby words (Lesk, 1986, p.24, italics in original).

In this regard, this part of the study is an attempt at a solution to the problem of sense discrimination. By utilising the Lesk algorithm via word sense disambiguation provided by NLTK, the correct word sense is predicted by counting overlaps between dictionary definitions of the various senses. An illustration of this NLP application is provided in the lines of the code below:

```
import nltk
```

```
import re
import string
from nltk import word_tokenize
from nltk import sent_tokenize
from nltk.corpus import stopwords
from nltk.corpus import wordnet
from collections import defaultdict, OrderedDict
from nltk.wsd import lesk
##1. Read text
raw = open("engtrans-Ali.txt", 'r').read()
##2. Tokenize into sentences
sent = sent_tokenize(raw)
ambiguous = 'earth'
lesk(sent, ambiguous, 'n')
```

Output:

```
Synset ('worldly_concern.n.01')
```

Input:

```
lesk(sent, ambiguous).definition()
```

Output:

```
'the concerns of this life as distinguished from heaven and the afterlife'
```

As seen above, the corpus comprising of Ali's translation was tokenised into sentences; the target word, the nature term *earth* as a noun, was identified as the ambiguous word in Ali's translation; Lesk algorithm was employed to predict its 'intended' meaning by relying on the overlap of words between the definitions of the target word and words in context in order to determine the sense of the target word (as in Ayetiran and Agbele, 2016, p.165). Hence, one could say that the intended meaning of the word *earth* in the Qur'an is related to the *worldly concern* which denotes *the concerns of this life as distinguished from heaven and the afterlife*. In addition, the error rate of the process of using the above code to assign senses in WSD was 1.3 %. Upon observation, the researcher found that the program initially assigned inaccurate senses to the words *garden* and *light*.¹²⁰ The researcher observed that this was due to the technical issue of not specifying the part of speech of the natural phenomena word form in the code above (e.g., noun as opposed to verb). To solve this issue, she further disambiguated the senses by adding 'n' to the code as follows:

```
lesk(sent, ambiguous, 'n').definition()
```

Once this was done, the error rate was reduced and the words were hopefully assigned the correct senses. In short, this section provides an overview of the computational approach

120 The Lesk algorithm assigned the sense 'work in garden' to the word *garden* and the 'alight from' to the word *light* as two senses which represent verbs.

applied to prepare the list of nature terms from the Qur'an for the statistical profiling that will later be implemented to conduct the quantitative analyses in this thesis.

2.4 A preliminary statistical exploration of nature terms in the Qur'an

This section, which briefly describes an exploration of the raw frequencies¹²¹ and collocations in Ali's translation, aims to discuss the statistical profiling of the nature terms produced in the previous sections. The importance of this section is that it provides an overview of nature as a theme in the Qur'an on three levels of corpus division: the Qur'an as a whole; the Makkan section of the Qur'an; and the Medinan section. It is an exploration that uncovered concepts related to the theme of nature in the Qur'an concerning the several distinctive features of the historical setting of the revelation of the Qur'an. On the other hand, the significance of the statistical exploration in the second part of this section is that it uses the same line of code to provide a list of the 30 most frequent nature terms in the Arabic Qur'an; a list that is used to find the SP of nature in the Qur'an, as will be seen in the methodology of this thesis. Moreover, the overview of the collocations, namely trigrams at this exploratory stage, of nature terms provides background knowledge and a visual presentation of the type of words that are likely to be neighbouring collocates to nodes that denote concepts of nature in the Qur'an.

2.4.1 The frequency distribution for Ali's translation as a bag of words¹²² and the central theme of the Qur'an

Before exploring the frequency distribution of the list of natural phenomena in Ali's translation of the Qur'an, NLTK was employed to find the frequency distribution of all the words in this text as a bag of words (BoW) was implemented. The bag of words, a model used in NLP, is a reduced and simplified representation of a text document from selected parts of the text, based on specific criteria, such as word frequency. In a BoW, a body of text, such as a sentence or a document, is thought of as a bag of words.¹²³ To find the frequencies of words in Ali's translation, this model is followed; it consists of the following tasks as illustrated in the lines of code below:

```
import nltk
import re
from collections import Counter
```

121 In the first part of this section, it reports the preliminary study where nature terms were computed as raw frequencies without the disambiguation that was applied at a later stage to improve the results. See also Alshahrani and Brierley (2017).

122 Based on Zhang et al. (2010).

123 Available from: [<https://ongspxm.github.io/blog/2014/12/bag-of-words-natural-language-processing/>], [Accessed 01 March 2017].

```
from nltk.tokenize import *
import string
tokenizer = WhitespaceTokenizer()
from nltk import word_tokenize
from nltk.corpus import stopwords
#1. Read in text
dataset= open ("engtrans-Ali.txt", 'r').read()
# 2. Pre-process Text
#2.1. Removing punctuation
punct=set(string.punctuation)
dataset1=''.join (w for w in dataset if w not in punct)
#2.2 Tokenization
dataset2 = tokenizer.tokenize(dataset1)
#2.3.Lowercase Only
dataset3= [w.lower() for w in dataset2] ##lower case only
#2.4 Removing stop words
stop_words=set(stopwords.words("english"))
##add new stop words
stop_words.update(('ever', 'whereby', 'didst', 'wherever', 'whenever', 'shouldst', 'often', 'thereat', 'otherw
ise', 'although', 'oft', 'anywhere', 'whoever', 'wouldst', 'hast', 'manysoever', 'amid', 'specially', 'wheresoev
er', 'never', 'wont', 'according', 'whatever', 'anything', 'behind', 'thereon', 'near', 'furthermore', 'midst', '
thither', 'towards', 'whilst', 'among', 'hadst', 'towards', 'mayest', 'wherewith', 'wherewith', 'wast', 'however
', 'henceforth', 'thenceforth', 'thereby', 'couldst', 'could', 'dost', 'throughout', 'therefrom', 'thine',
'whereof', 'wherein', 'shalt', 'without', 'within', 'also', 'yet', 'yea', 'rather', 'whomsoever', 'till', 'therew
ith', 'every', 'would', 'whether', 'thus', 'another', 'must', 'therefore', 'beneath', 'forthwith', 'hath', 'may',
'except', 'might', 'doth', 'therein', 'unless', 'whereon', 'thus', 'amongst', 'ah', 'thee', 'thou', 'shall', 'ye',
'thy', 'thysself', 'whose', 'nay', 'mere', 'perhaps', 'perchance', 'till', 'oer', 'betwixt', 'should', 'unto', 'the
reof', 'upon', 'either', 'moreover', 'us', 'besides', 'though', 'neither'))
dataset4= []
for w in dataset3:
    if w not in stop_words:
        dataset4.append(w)
#3. Count frequencies
Counter(dataset4).most_common(30)
```

The results of the exploration of the 30 most frequent words in Ali's translation of the Qur'an are shown in the following box:

```
[('allah', 2895),
 ('lord', 960),
 ('say', 775),
 ('said', 710),
 ('one', 527),
 ('day', 520),
 ('people', 511),
 ('believe', 470),
 ('earth', 418),
 ('things', 379),
 ('men', 369),
 ('signs', 337),
 ('sent', 325),
 ('truth', 321),
 ('indeed', 320),
 ('verily', 310),
 ('made', 287),
 ('fear', 274),
 ('good', 274),
 ('evil', 266),
 ('faith', 256),
 ('come', 249),
 ('messenger', 246),
 ('give', 246),
 ('see', 239),
 ('penalty', 229),
 ('know', 225),
 ('life', 213),
 ('turn', 207),
 ('man', 206)]
```

To generate a chart of the 30 most frequent words (Figure 15), `matplotlib`¹²⁴ was employed in NLTK:

```
fd = nltk.FreqDist(dataset4)
fd.plot(30,cumulative=False)
```

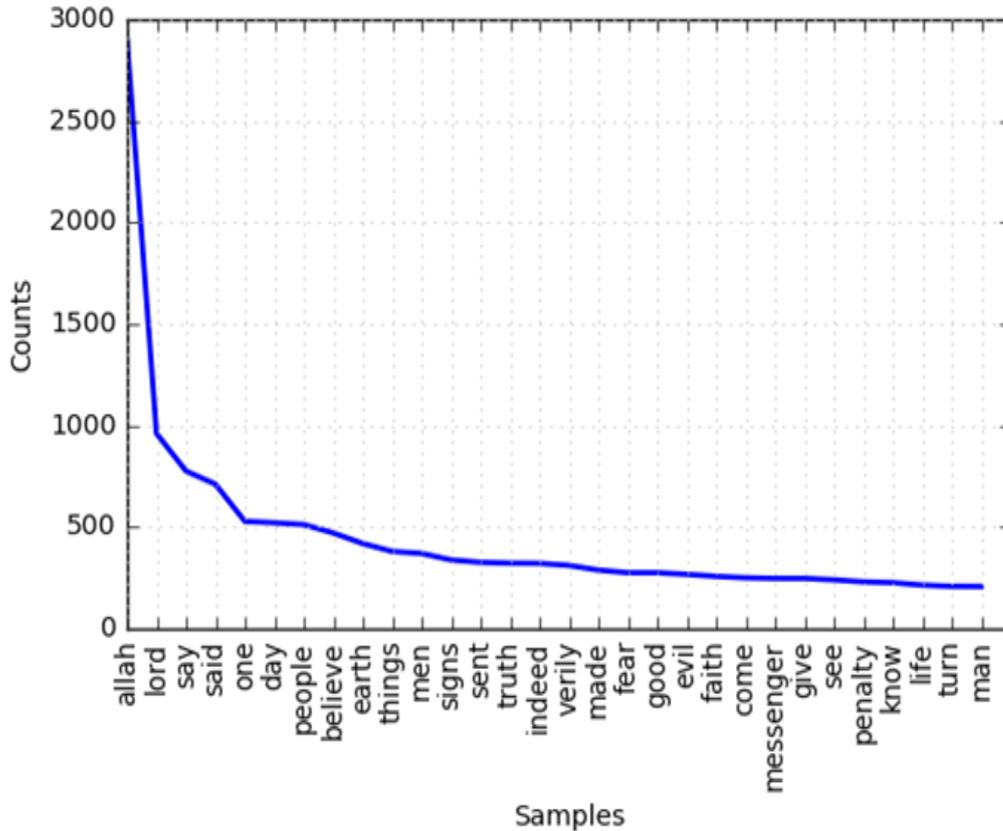


Figure 15: The 30 most frequent words in Ali's translation of the Qur'an

The central theme of the Qur'an, as mentioned earlier, is the declaration of reality and the removal of misunderstanding and misconceptions about the message of Islam.¹²⁵ This theme, as claimed by scholars, is related to other topics that appear in the Qur'an (e.g., nature); these topics emphasise the central theme and convey it to the reader. For example, the topic (theme) of nature in the Qur'an can be traced through the occurrences of terms representing natural phenomena. The counts of these terms ascertain the vitality of this theme and its role in conveying the central theme of the Qur'an. The counts relevant to this argument are the most frequent words in the Qur'an as well as the counts of the nature terms in the Qur'an. For example, the ten most frequent words in Ali's translation of the Qur'an as seen below indicate that these words are related to the central theme of the

¹²⁴ *Matplotlib* is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Available from: [<https://matplotlib.org/index.html>],[Accessed 20 March 2019].

¹²⁵ See also Malik (1997).

Qur'an: 'Allah' being the 'One Lord' and Creator of all, whose Divine book and message to the prophet Muhammed is the Qur'an. In the Qur'an, He addresses 'people' to look at the 'earth' and 'things' as in His creation to 'believe' in His unity. In addition, the verbs 'say' and 'said' are related to the revelation of this truth. Hence, this list of the most frequent words in the Qur'an potentially provide evidence that the Qur'an's central and recurrent theme is the revelation of the message of Islam.

1. Allah
2. lord
3. say
4. said
5. one
6. day
7. people
8. believe
9. earth
10. things

In further support of the argument above, the following subsections will present the frequency distribution of nature terms in Ali's translation, and illustrate their relevance to the central theme of the Qur'an.

2.4.2 The frequency distribution for the list of natural phenomena in the divisions of Ali's translation

This study presents a statistical profiling of nature as a theme in Ali's translation of the Qur'an as a whole, as well as of the Makkah and Medinan sections only. Ali's translation was divided into Makkah and Medina surahs to explore the list of natural phenomena produced in this chapter. The Qur'an was being divided into Makkah, and Medina sections following the historical reality of the revelation itself (as in Alrehaili, 2012).¹²⁶ For the first thirteen years of his prophethood, Prophet Mohammed was in Makkah; the last ten years he was in Medina. Thus, there is a natural division between the earlier Surahs which were revealed to him in Makkah, and the later ones which were revealed to him in Medina. Even though there were debates on how some Medina ayahs were present in Makkah Surahs, one can still make a general distinction in the features between the two sections. It is, therefore,

¹²⁶ See also Alrehaili, S. M. (2012). *The chronology of the texts in the holy Quran according to NLP*. Unpublished MSc Dissertation, School of Computing, University of Leeds.

plausible to divide the Qur'anic corpus into Makkah and Medina Surahs to explore the most prominent features in each. Furthermore, the division is based on the *King Fahd Complex for Printing the Holy Qur'an* in the authorised version of the authentic Arabic Qur'an.

Following this division, the NLP features of the list of natural phenomena were extracted (e.g. *raw frequencies* and *collocations*) via NLTK and associated with the general features of the Qur'anic style in the Makkah and Medina sections. However, prior to conducting the exploration of the terms in the list of natural phenomena, a pre-processing task was performed. The steps of pre-processing the entire text (the same applies to the Makkan and Median texts) along with samples of lines of code used on all in this approach were previously presented in Section 2.3.1. After preparing the data and nature terms were run in Python via NLTK to find their frequencies in the full text and the two divisions as seen in the lines of the following code:

```
import nltk
from nltk import FreqDist
text= open ("engtrans-Ali.txt", 'r').read()
fdist= FreqDist (text)
print fdist ['earth']
print fdist ['water']
print fdist ['man']
print fdist ['cattle']
print fdist ['sun']
print fdist ['mountain']
```

Related to the counts of the most frequent words in the Qur'an are the counts of the most frequent terms of natural phenomena. The ten most frequent natural phenomena in the Qur'an in order of highest to lowest are as follows:

1. day
2. earth
3. men
4. signs
5. life
6. man
7. fire
8. hearts
9. night
10. woman

As seen above, results of the counts of terms in the full text show that the word *day* has the highest count in the list of terms of natural phenomena. The words *brain, leg, land, ox/oxen, locusts, lice, seeds, goats, and frogs*, on the other hand, have the lowest number of occurrences. Similarly, as can be seen in Table 8, results of the Makkan section show that the word *day* has the highest count in the list of terms of natural phenomena. The words with the lowest count, on the other hand, are the words *brain, leg, jinn, plant, ox/oxen, fountain, apes, horses, seas, rocks, hours, goats and frogs*. Results of the counts of terms in the Medinan part show that the word *men* has the highest count in the list of terms of natural phenomena. The words with the lowest count, in contrast, are the words *seed, star, dawn, river, stick, cold and cloud*.

Table 8: The difference in the 20 most frequent nature terms between Makkan and Medinan surahs

<i>Makkan counts</i>		<i>Medinan counts</i>	
1.	day	1.	men
2.	earth	2.	day
3.	signs	3.	signs
4.	men	4.	earth
5.	man	5.	hearts
6.	life	6.	fire/women ¹²⁷
7.	heavens	7.	heavens
8.	fire	8.	life
9.	night	9.	man
10.	sign	10.	hands
11.	soul	11.	light
12.	land	12.	gardens
13.	hearts	13.	death
14.	garden/hell	14.	rivers
15.	water/eyes	15.	mankind
16.	creation	16.	hell
17.	hour	17.	land
18.	mankind	18.	garden
19.	gardens	19.	night/days
20.	hand	20.	mountain

As we can see, the term *day* appears in both lists: the most frequent words in the Qur'an¹²⁸ and the most frequent terms describing natural phenomena. The fact that it appears in both lists shows the significance of nature as a recurrent theme in the Qur'an. One of the roles

¹²⁷ The slash '/' in the table cell represents equal counts. For example, the terms *fire* and *women* are equal in counts in the Medinan section of Ali's translation.

¹²⁸ Results in Section 2.4.1

of nature in the Qur'an is to have people 'ponder' about God's greatness and might in regard to the perfection of His creation. They are *signs* which give proof of a Creator whose creation is perfect. This evidence is most crucial for the central theme of this sacred book. In other words, these *signs*, namely natural phenomena, ascertain the reality which the Qur'an has come to reveal. This is the theme that runs throughout the verses from the beginning to the end, which illustrates the presence of this theme in both divisions of the Qur'an.

In addition, this table shows that there are similarities between the two lists in the counts of the most frequent terms of natural phenomena. The words *day, earth, signs, men, man, life, heavens, fire, night, land, hearts, garden, hell, mankind, and gardens* occur in both sections. This signifies, once more, that the theme of nature runs throughout the Qur'an, whether at the beginning of the revelation in the Makkan section or towards the end in the Medinan section. Furthermore, this sample of results shows that only the Makkan section contains the word *sign* in both singular and plural forms. Although the word *signs* is equally distributed between the two parts, its singular form is contained in the Makkan section only; a fact that makes the word more significant there. This is related to the idea that the beginning of the revelation of truth required signs, which in most cases refer to the creation of the universe and its different components, and at times to miracle stories.

2.4.3 A collocation extraction of nature terms in Ali's translation

In addition to extracting the frequencies of nature terms in this exploratory part of the research, a preliminary exploration of their collocations, namely trigrams via the raw frequency statistical measure in Python, was also conducted. This was also done through the n-grams¹²⁹ feature via NLTK in Python¹³⁰ (see code below). By way of example, Table 9 illustrates a sample of nature terms with their left and right neighbours.

```
from nltk.collocations import*
trigram_measures = nltk.collocations.TrigramAssocMeasures()
tokens=dataset4
finder = TrigramCollocationFinder.from_words(tokens)
scored = finder.score_ngrams(trigram_measures.raw_freq)
set(trigram for trigram, score in scored) == set(nltk.trigrams(tokens))
print (set(nltk.trigrams(tokens)))
```

129 An *n-gram* is a sequence of *N* n-gram words: a 2-gram (or bigram) is a two-word sequence of words like "please turn", "turn your", or "your homework", and a 3-gram (or trigram) is a three-word sequence of words like "please turn your", or "turn your homework" (Jurafsky and Martin, 2009, p. 2). See also: [<https://web.stanford.edu/~jurafsky/slp3/4.pdf>]

130 Available from: [<http://www.nltk.org/howto/collocations.html>], [Accessed March 2017]. See code above.

Table 9: A sample of trigrams of natural phenomena terms in Ali's translation

<i>Left neighbour</i>	<i>Natural phenomenon term</i>	<i>Right neighbour</i>
manageable	earth	made
shakes	earth	swallowed
gathered	earth	multiplied
moved	earth	trumpet
growing	earth	produced
raise	earth	return
carpet	earth	made
mountains	earth	day
wide	earth	made
fragments	earth	split
cleft	heaven	dim
earth	heaven	creator
rolled	heaven	handful
stars	moon	sun
light	moon	glory
law	moon	sun
measured	moon	all-knowing
prostrate	moon	sun
cleft	moon	nigh
moon	sun	subjected
moon	sun	day
moon	sun	prostrate
glorious	sun	round
scatter	wind	warning
sent	wind	penalty
dust	wind	scattered
sky	rain	sends
stirred	rain	send
even	rain	sends
calamity	rain	give
cloud	rain	bring
saved	sea	divided
rejected	sea	drowned
Pharaoh	sea	across
command	sea	sail
subject	sea	made
call	sea	seizes
drink	water	see
burning	water	boiling
Noah	water	penalty
matter	water	flowing

<i>Left neighbour</i>	<i>Natural phenomenon term</i>	<i>Right neighbour</i>
sweet	water	provided
fruits	water	springs
abundance	water	clouds
thickest	mountain	seest
standing	mountain	earth
heap	mountain	commotion
scattered	mountain	asunder
standing	mountain	made
vanish	mountain	doors

This NLP task provides a preview of the kinds of words that co-occur with nature terms in the Qur'an and this case Ali's translation. These are a part of this exploratory overview of natural phenomena in the Qur'an and provide a visual representation of the collocational behaviour of such terms as they reoccur in the Qur'an. Furthermore, the same collocation finder provided by NLTK will be implemented again later on in this research, but with a focus on finding bigrams, rather than trigrams, to have an even closer look at the neighbouring words of the natural phenomena in the Qur'an and its translations.

2.4.4 The 30 most frequent natural phenomena terms in the Arabic Qur'an

This section gives an account of the NLP application of finding the raw frequencies of nature terms in the Arabic Qur'an. It is the final exploratory experiment in this chapter conducted in the preliminary stage of this research to provide the 30 most frequent natural phenomena in the Qur'an; a list that is used to explore the SP of nature as a theme in the Qur'an and will later on be used to evaluate five translations of the Qur'an for their congruency with the Arabic representation of the nature in the Qur'an. To compute the raw frequencies of nature terms in the Qur'an via NLTK in Python, the same order of tasks of pre-processing and processing text mentioned in the previous sections was applied. The 30 most frequent¹³¹ natural phenomena terms in the Qur'an are shown in the following table.

Table 10: The 30 most frequent natural phenomena terms in the Qur'an

<i>Arabic Words</i>	<i>Nature Term</i>	<i>Frequency</i>
ناس-أهل-قوم-أقوام	people	534
يوم-أيام-يومين	day	481

¹³¹ Upon examining the list of the 30 frequent natural phenomena in the Qur'an, it was noticed that several words such *devil*, *angel*, and perhaps *death* would by virtue of dictionary definition either be positive or negative. Therefore, they were preserved in the list and other words were added to have a wider exploration of the occurrences of nature terms in the Qur'an and a more thorough evaluation of its selected translations.

<i>Arabic Words</i>	<i>Nature Term</i>	<i>Frequency</i>
أرض-خسف	earth	462
سماء-سماوات	sky/s/heavens	324
نفس-أنفس-نفوس	soul/s	295
القيامة-البعث	resurrection	276
النار-جهنم	hell	255
حياة-الدنيا	life	255
رجل-رجال-إنسان	man/human/men	151
جنة-جنات	the garden/s	144
قلب-قلوب	heart/s	142
يد-يمين-شمال	hand/s	113
إنس-بشر	mankind	107
شيطان- شياطين	devil/s	103
ملك-ملائكة-الصافات-الزجرات-التاليات- المقسمات-النازعات-الناشطات-السابحات-		
السابقات-المديرات	angel/s	95
ليلة-ليل-ليال	night/s	93
بيوت-بيت-مساكن	homes/houses	77
موت	death	71
وجه-وجوه	face/s	69
ماء	water/rain	61
العالمين	worlds	61
قرية-قرى	village/town	55
جبل-جبال-رواسي-أعلام	mountain/s	53
أنهار	river/s	53
نساء-امرأة	woman/en	49
نور-ضياء-مصابيح	light	47
صدر-صدور	chest/s/breast/s	44
بحر-بحرين-بحار	sea/s	39
ثمره-ثمرات-فاكهة-فواكه	fruit	35
شمس-سراجا	sun	34
أنعام	cattle	32
جوار-سفينة-فلك	ship	32
شجر-شجرة	tree /s	29
جن-جان	jinn	27
قمر	moon	27
سنة-سنيين	year/s	35
عيون-عينا-سلسيل	springs	20
ريح-رياح-عاصفات-ناشرات-فارقات-ملقيات	wind /s	23
أغرق-غرق	drowning	23
ظلمات- غاسق	darkness	23
نخيل-نخل	date-palm/s	20
نجم-نجوم	star/s	13

To reiterate, the significance of this list lies in the fact that it is the final list used to explore the theme of nature in the Qur'an via collocation to find SP. Also, it is the basis on which the evaluation of the five chosen English renderings of the Qur'an for their representation of nature as a theme is based.

2.5 Conclusion

This chapter presented an overview of several experimental studies that the researcher conducted in the preliminary stage of this research to produce a dictionary of words that represent the theme of nature in the Qur'an. Firstly, it provided an overview of nature as a theme in the Qur'an, as discussed in previous literature. It also examined lists of nature terms in the Qur'an that had been compiled in the past. The chapter went on to describe the exploratory NLP tasks and applications on both Ali's translation and the Qur'an. Also, it most importantly described the tasks and applications applied on the Qur'an to prepare a list of words related to natural phenomena. This produced list will later be employed to find the SP of nature in the Qur'an and evaluate and rank the selected translations based on their congruency with the representation of nature in the Qur'an as a source text. The approach used here can also be useful in exploring other concepts in the Qur'an. Moreover, it can be considered a suggested platform for building new and more comprehensive conceptual ontologies of the Qur'an, which can, in turn, facilitate a clearer understanding of the Holy Book.

Chapter 3 Literature Review:

Determining the Theoretical Framework and Methodology

Introduction

This chapter will provide an overview of the studies that have determined the choice of both the theoretical framework and methodology of this research. It consists of four parts that address the following specific issues: the defining components of the theoretical background of this research; the corpus-based studies on nature in the Qur'an¹³² and others on SP in the Qur'an together with the available NLP projects designed for the Qur'an and computational tools developed for linguistic research in the Qur'an (e.g., the *Qur'anic Arabic Corpus*, *Qur'any*, etc.); and finally, the general methods of evaluating translation followed by a conclusion to the chapter. To elaborate, the first section, which helped to formulate the theoretical framework of this research, provides an overview of the debate in the previous literature on defining corpus linguistics and its different applications, with a detailed discussion of collocation, lexico-grammar (LG), and SP. The second section, which aided in choosing the appropriate methodology for the first stage of this research (i.e., exploring SP of nature in the Qur'an and its translations), gives an account of the previous computational and non-computational corpus-based studies on nature and others on SP as a collocational phenomenon. For example, it gives an account of the SP analysis as a tool for accuracy in translation with the recently developed NLP tools that were designed specifically for the analysis of the language of the Qur'an. Furthermore, the third section, which informed the chosen method for evaluating the selected translations in the second stage of this research (i.e. evaluating the Qur'an translations regarding their congruency with the representation of SP of nature in the Arabic Qur'an), discusses various general methods of evaluating a translation. Finally, the fourth section of this chapter provides a conclusion to this literature review with a focus on its importance in determining both the theoretical framework and methodology of this corpus-based research.

For an introduction to corpus linguistics, the following sources were mainly consulted: Biber *et al.* (1998); McEnery and Wilson (2001); Meyer (2002); Hunston (2006); McEnery and Hardie (2012); and Kennedy (2014). In addition, for views on the usefulness of corpora in translation studies, Tymoczko (1998) and Olohan (2004) were two useful resources.

¹³² The studies which were covered in the previous chapter were only descriptive commentaries on the presence of natural phenomena in the Qur'an and they were not computational or of a corpus linguistic nature.

3.1 Defining ‘corpus linguistics’, ‘collocation’, ‘lexico-grammar’, and ‘SP.’

The first aim of this section is to provide the different views on the definition of corpus linguistics as the main methodology of this corpus-based thesis that is combined with an NLP method and translation studies application. It then highlights the adopted definitions of the three main components of the adopted theoretical framework: collocation, lexico-grammar, and SP. It discusses the different criteria for collocation identification and approaches to the study of collocation to indicate the ones chosen for the purpose of this research. Moreover, it defines lexico-grammar as an integral part of interpreting collocation. Finally, it discusses viewpoints on the relationship between SP and collocation as units of meaning of the lexical item primarily to introduce the notion of SP, which is adopted in this research.

Before reporting the different views on the definition of corpus linguistics, it is essential to define the notion of a ‘corpus’, because, as the name suggests, it is the basis for a form of the theoretical and empirical linguistics known as corpus linguistics that will be discussed shortly. The word *corpus*, from the Latin word ‘corpus’ meaning body and with plural *corpora*, is a collection of texts, especially if complete and self-contained such as the corpus of the Anglo-Saxon verse¹³³ [e.g., the Qur’an and its translations in this research]. In two related views, it is a collection of: “naturally occurring language text, chosen to characterise a state or variety of a language” (Sinclair, 1991, p.171), and “(1) machine-readable (2) authentic texts (including transcripts of spoken data) which is (3) sampled to be (4) representative of a particular language or language variety” (McEnery et al., 2006, p.5). After defining what a corpus is, the following lines will review the different viewpoints on the definition of corpus linguistics.

1- The debate on what corpus linguistics is

There is a multiplicity of views on the definition of corpus linguistics from some of the most influential corpus linguists (such as: Sinclair, 1991; Oostdijk and Haan, 1994; MacCarthy, 2001; Tognini-Bonelli, 2001; Aarts, 2002; Hunston, 2002; Meyer, 2002; Butler, 2004; McEnery and Xiao, 2004; Teubert and Čermáková, 2004a; Römer, 2005; Teubert, 2005; Hunston 2006; McEnery and Gabrielatos, 2006; Thompson and Hunston,

133 See also: *Concise Oxford Companion to the English Language*: Oxford University Press. Retrieved 24 Nov. 2018, from [https://www.oxfordreference.com/view/10.1093/acref/9780192800619.001.0001/acref-9780192800619-e-319].

2006; Teubert et al., 2007; Lüdeling and Kytö, 2008; Gries, 2009; 2010b; O'Keeffe and McCarthy, 2010; McEnery and Hardie, 2012; Kennedy, 2014; Mahlberg, 2014b; Rabadi, 2014; Jones and Waller, 2015). This subsection looks at how *corpus linguistics* is defined and described in a variety of relevant sources.

McEnery and Gabrielatos (2006) and Taylor (2008), who provide overviews of the way practitioners have described corpus linguistics, stress that there is no general agreement as to what exactly constitutes corpus linguistics. For example, in the introduction of her paper, Taylor (2008) states that corpus linguistics is defined as “a tool, a method, a methodology, a methodological approach, a discipline, a theory, a theoretical approach, a paradigm, or a combination of these” (p.180).¹³⁴ Aarts (1984), who is said to be one of the founding fathers of corpus linguistics and often identified as the source of the term *corpus linguistics*,¹³⁵ wrote the first book known to be dedicated to this subject¹³⁶ entitled, *Corpus linguistics: Recent Developments in the Use of Computer Corpora in English Language Research*.¹³⁷ In this regard, McEnery and Gabrielatos (2006, p.3) point out that the more specific term *corpus linguistics* did not come into common usage until the early 1980s when it was coined in Aarts and Meijs (1984). From that point on, scholars have discussed the possible parameters of this field of study and wrestled with issues of how to define it and whether this approach to considering language constituted a distinct theory of language or rather a method of language analysis.

Aarts (2002) and Teubert (2005), among others, define corpus linguistics as a discipline with a theory. For example, Teubert (2005) emphasises the theoretical conceptualisation and describes corpus linguistics as “a theoretical approach to the study of language” (2005, p.2). He also comments on the diversity of methods and states that:

[c]orpus linguistics is not in itself a method: many different methods are used in the processing and analysing corpus data. It is rather an insistence on working only with real language data taken from the discourse in a principled way and compiled into a corpus (2005, p.4).

Similarly, Stubbs (1993), who once rejected the definition of corpus linguistics as only a methodology and claims it is “limited”, comments that “[i]n this vision of the subject, a corpus is not merely a tool of linguistic analysis but an important concept in linguistic

134 Available from: [http://sro.sussex.ac.uk/id/eprint/53389/1/what_is_corpus_linguistics.pdf], [Accessed, 18 July 2019].

135 Other important figures of corpus linguistics include : Meijs (1987) and Kytö et al. (1988).

136 It has been also claimed that the term had in fact been used previously, for example, in Aarts and van den Heuvel (1982) (as cited in Taylor, 2008, p.179).

137 See also Aarts and Meijs (1984) and Aart and Meijs (1986).

theory” (1993, pp.23-4). In his later work, Stubbs (2001) compares corpus linguistics to a science, when he writes:

[g]eologists are interested in processes which are not directly observable because they take place over vast periods of time [...] Corpus linguists are interested in processes which are not directly observable because they are instantiated across the language use of many different speakers and writers (p. 243).

On the contrary, it is argued by many that corpus linguistics is not a theory in itself but rather a methodology (e.g., e.g., Biber et al., 1998, pp. 3-4; McEnery and Wilson, 2001, p.2; Meyer, 2002, xi; Kennedy, 2014, p. 7). Others, such as Leech (1992, p.106), take this definition to another level, and argue that corpus linguistics is more than a mere methodology, but rather “a new research enterprise” and “a new philosophical approach”. In this regard, Leech (1992), who describes “computer corpus linguistics” as new a “paradigm” synonymous with corpus linguistics, writes:

[c]omputer corpus linguistics defines not just a newly emerging methodology for studying language, but a new research enterprise, and in fact, a new philosophical approach to the subject. (p. 106).

Like Leech, Taylor (2008) elaborates on this definition by stressing the importance attached to the scientific side in corpus linguistics as a methodology,¹³⁸ and McaCarthy (2001) portrays corpus linguistics as representing “cutting edge change in terms of scientific techniques and methods” (p.125). Corpus linguistics is also defined as a methodology in McEnery and Wilson (2001,p. 2), and is described by Meyer (2001, xi) as “an approach or a methodology for the study of language use”. Similarly, Hunston (2006) claims that corpus linguistics is a methodology rather than a theory of language description and that it involves the following:

- Looking at naturally occurring language;
- Looking at relatively large amounts of such language;
- Observing relative frequencies, either in raw form or mediated through statistical operations;
- Observing patterns of association, either between a feature and a text type or between a group of words (Hunston, 2006, p.244).

138 Taylor (2008) presents a paper which explores the definition of corpus linguistics via a compiled corpora about corpus linguistics, with the following keywords which resemble applied sciences terms: *repetition, empirical, statistical, methodology, data, quantitative and qualitative*.

According to Hunston (2006), at the basic level, a corpus analysis software performs one of the following: searches the corpus for a given target item; counts the number of instances of the target item in the corpus and calculates relative frequencies and displays instances of the target item so that the corpus user can carry out further investigation. Once a certain target is found, it is examined for its meaning. Meaning in relation to a corpus methodology is “not atomistic, residing in words, but prosodic, belonging to variable units of meaning and always located in texts” (Tognini-Bonelli, 2001, p.89). Thus, any corpus, including a “raw” corpus (i.e., a corpus that consists only of text, with no further information added), may be searched for instances of a single word (e.g., *day*). It also allows a single search to find sets of words (e.g., *day, month, and year*) and strings of words (e.g., *the next day*). The software can count the occurrence of categories and usually compare their frequency in corpora of different kinds. To summarise, corpus linguistics is a methodology or group of methodologies which uses computer software to elicit meanings out of the text (e.g., via collocation).

Furthermore, advocates of this definition of corpus linguistics view it as a methodology which can be used in the application of a variety of theories (as in Mahlberg, 2005, p. 371; Taylor, 2008, p.31; Taylor and Marchi, 2018, p.2). By way of example, Mahlberg (2005) gives a brief account of some of the corpus linguistic methodologies that possess a “theoretical status”¹³⁹ or make theoretical claims, such as Hunston and Francis (2000), Sinclair (2004a), and Hoey (2005). Based on her discussion, the key features of this approach can be summarised as follows:

1. Language is a social phenomenon and meaning can thus be viewed as use.
2. Patterns of language use become visible through the use of corpora.
3. Corpus evidence illustrates that meaning and form are associated (as in Mahlberg, 2006, pp.370-1).

Likewise, Gries (2009, p.191) claims that corpus linguistics is favoured as “a major methodological paradigm in applied and theoretical linguistics”. He also describes it as a “distributional discipline” by nature, which contains “nothing but distributional frequency data”, such as:

- *frequencies of occurrence* of linguistic elements, which can be studied from two different perspectives:

139 The term she says was used by Tognini-Bonelli (2001).

1. how frequent are morphemes or words or patterns/constructions in (parts of) a corpus?
This information can be provided in various different forms of frequency lists;
 2. how evenly are morphemes or words or patterns/constructions distributed across (parts of) a corpus? This information can be provided in the form of various dispersion statistics;
- *frequencies of co-occurrence*: how often do linguistic elements such as morphemes, words, patterns/construction co-occur with another linguistic element from this set or a position in a text (Gries, 2010c, p.5).

In a definition which overtly combines both method and theory, Thompson and Hunston (2006, p.8) state that: “at its most basic, corpus linguistics is a methodology that can be aligned to any theoretical approach to language”. Their theoretical approach adopts the view that meaning is not located in single words, but in “units of meaning”, which reverberates with Sinclair’s terminology (theory such as the Extended Lexical Unit).¹⁴⁰ In this sense, the communicative discourse unfolds greatly as a series of semi-fixed phrases which are explored and analysed (method) (Sinclair, 2004c, pp.11-2).

In this research, the adopted definition of corpus linguistics combines the definitions by McEnery and Hardie (2012), Sinclair (2004c), Thompson and Hunston (2006), and Gries (2010c). For example, McEnery and Hardie (2012) view corpus linguistics as a methodology of “dealing with some set of machine-readable texts which is deemed an appropriate basis on which to study a specific set of research questions” (p. 1). In addition, Sinclair (2004c, p.10) provides a brief definition of corpus linguistics which merges two lines of thought when he says, “theory derives from the pattern; pattern reflects theory” at both lexical and grammatical levels. Moreover, in accordance with Gries’s view of corpus linguistics as a distributional approach to studying language (2010c), this research aims at exploring SP of nature in the Qur’an via the Natural Language Toolkit (NLTK) in Python; the software used to explore the frequencies and collocation extraction of natural phenomena in six texts: the original Arabic texts and five of its acclaimed English translations. Using NLTK, this thesis firstly answers questions about the presence of SP in nature in the Qur’an with its role in the representation of meaning via the analysis of collocation and lexico-grammatical patterns. Secondly, it compares this representation of meaning to five English translations of the Qur’an. Moreover, the *Qur’anic Arabic Corpus*,

140 Sinclair’s theory of the Extended Lexical Unit (2004a).

as well as *Sketch Engine*, are utilised in applying the ontology of Qur'anic concepts to nature terms as well as their concordances. In order to provide a rigorous theoretical underpinning to the analyses, notions from linguistic theory were adopted for the analysis of collocation and SP (e.g., *the Extended Lexical Unit* [collocation] and *Lexico-Grammar*). Finally, the software tool *GraphColl*¹⁴¹ was used to produce data visualisation in this research, and to provide a sophisticated visual representation of collocations of natural phenomena in the Qur'an.

2- *Corpus-based application in translation studies*

Hunston (2006) discusses two types of applications of corpus linguistics: *language teaching* and *translation*. She claims that they are “two of the more frequently encountered applications of corpus linguistics” (p.245). She calls the type of corpora in translation studies “parallel corpora”, where one corpus consists of the source text and the other of the translations of the source text. In this context of the corpus-based application, the examination of parallel corpora emphasises that what translators translate is not the word but a larger unit (Teubert and Čermáková, 2004b, pp. 123-4). This research employs the use of parallel corpora, which is contrasted by definition to comparable corpora.¹⁴² Kenny (2006), who worked under the supervision of Mona Baker to compile a parallel corpus of approximately one million words from German literary fiction and its translation into English, claims that parallel corpora are needed for more empirical research in corpus linguistics (p.2). Her study aimed to extract unusual or idiosyncratic collocations in the German originals and assess the creativity of their translations in English. It is a project somewhat like this research in that it presented an exploration of what happens in translation to vocabulary items with negative or positive, or neutral semantic prosodies. The following subsection discusses previous studies of collocation as a prominent linguistic feature in corpus linguistic studies.

141 *GraphColl* is a new corpus linguistic tool developed by the authors (Brezina et al. , 2015) in LancsBox as a way of helping analysts to interpret collocational networks.

142 *Comparable corpora* are corpora in two or more languages with the same or similar composition. All corpora have an explicit or implicit composition (Teubert, 1996, p. 245; McEnery and Xiao, 2007, p.131). Although they are similar to parallel corpora in the fact that they are collections of electronic texts which are closely related to each other, they are different from the latter in the type of the link between these texts. Parallel corpora consist of a set of texts in language A and their translations in language B (e.g., the works of Dickens and their translations into French). In other words, the relationship lies in shared meaning. By contrast, what links the collections of texts in comparable corpora is that they have been put together according to the same type of criteria (texts of a certain size, on a set topic, from a given period, etc.) (Kenning 2010, p.492) .

3.1.1 Collocation: “knowing a word by the company it keeps”

The corpus linguistic research on *collocation* or words and their associations is vast, and the analysis of collocation is built on the Firthian notion “knowing a word by the company it keeps”; a proverb expressed similarly in the two Arabic sayings: المرء يعرف بأقرانه and المرء بخليته (both typically translated into English as “a man is known by the company he keeps”).¹⁴³ This concept has led to many insights on different word contexts and in various languages and has been explored in terms of its secondary implications, such as pedagogical studies or in applied translation studies. However, before considering its relevance to this research, it is important to provide an overview of the different definitions of collocation and to highlight the approach best suited for this study. This is especially important because there is no single universally accepted definition of collocation; it is differently defined and understood per the requirements of particular research questions or applications. The task of opting for a definition is especially important for this research because it centrally involves extracting collocations of natural phenomena terms as well as examining their collocate sets to unveil their semantic prosodies. Hence, it is necessary to define collocations and investigate the criteria for identifying them as well as the methodological approaches used to find them based on previous corpus-based studies. Put simply; this section will describe how collocations are conceptualised and operationalised as seen through the lens of the previous literature on collocation.

1- The definition of collocation

The concept of *collocation* was first defined by Firth (1957, p.11) as “one of the levels of meaning” (as cited in Gabrielatos, 1994, p.1), this definition later being refined in 1986 by Benson and Ilson as “an arbitrary and recurrent word combination” (1986, p. 23). Taking a more rigorous theoretical approach, Halliday (1961) and Sinclair (1987b) introduce the notion that patterns of collocation can form the basis for lexical analysis of a language as an alternative to, and independent of, grammatical analysis. Both Sinclair and Halliday regard collocation and grammatical set as complementary, neither being subsumed under the other. These statements about collocation illustrate that collocation has been defined differently over time. The concept of collocation in the previous research is arguably viewed in accordance with at least two defining perspectives: the theoretical or

143 See also Al-Timen (2015, p. 14017).

phraseological and the empirical viewpoints. For instance, on the phraseological level, Halliday and Hasan (1976, p.150) define collocation as a “sub-type of phraseme”, where *phraseme* (also called a set phrase or set expression) is “a phraseological expression” which represents a multiword phrase- that is, “a linguistic expression formed by several (at least two) lexemes syntactically linked in a regular way” (Mel’Čuk, 2012, p.31). For an illustration of the types of phrasemes, which include collocations, see Figure 16.

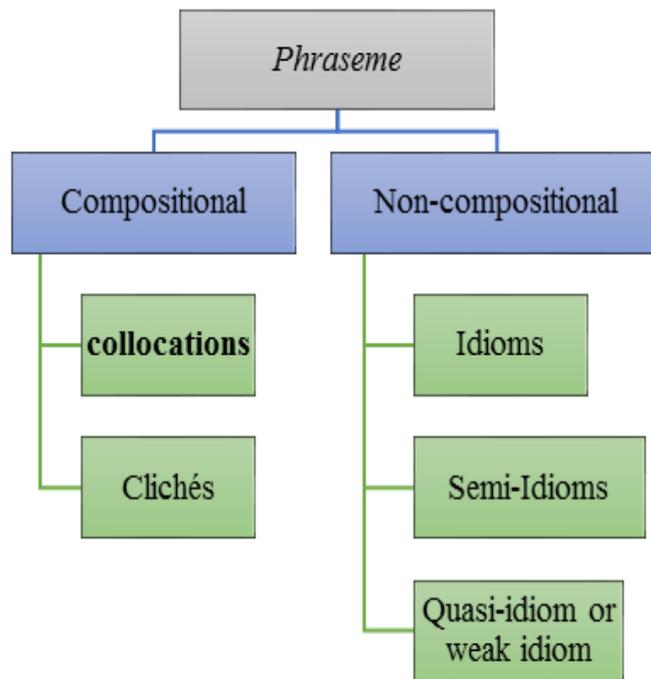


Figure 16: The types of the phraseme (figure based on the classification of Mel’Čuk, 2012, p.32, pp. 31-56)

Mel’Čuk (2012), who agrees with this classification, claims that phrasemes can be compositional and non-compositional; collocations belong to the compositional type. Mel’Čuk (2012) also states that a collocation is “binary” – it consists of two major elements, which are:

[a] base, lexical expression chosen freely by the speaker and a collocate, lexical expression chosen as a function of the base to express a given meaning bearing on the base (p.38).

Moreover, collocation is semantically compositional, since its meaning is divisible into two parts such that the first one corresponds to the base [referred to as a *node* in this research] and the second to the collocate. An example of a phraseological collocation is the expression *strong tea*. While the same meaning could arguably be conveyed by the roughly

equivalent *powerful tea*, this expression is considered excessive and awkward by English speakers.

In contrast, based on the empirical view that collocation is a corpus linguistic technique, Evert (2008) sees collocations as empirical phenomena, which constitute “observable facts about language, i.e. primary data”. In this regard, he writes:

From a strictly data-driven perspective, they [i.e., collocations] can be interpreted as empirical predictions about the neighbourhood of a word. For instance, a verb accompanying the noun *kiss* is likely to be either *give*, *drop*, *steal*, or *plant* (Evert, 2008, p.7, my brackets and Italics in the original).

This is in harmony with Firth’s definition of collocation (1957), which states that the collocations of a word are “statements of the habitual or customary places of that word” (p.181). Accordingly, it can be said to reflect the tendency or habit of words to co-occur together. Examples of collocations include noun phrases like *weapons of mass destruction*, phrasal verbs like *to make up*, and other stock phrases like *the rich and powerful*.

During the 1990s, with the revival of corpus linguistics, based on computerised corpora, Sinclair (1966; 1991; 1998; and 2004a) and many of his followers continued to refer back to Firth’s work (Léon, 2007). In addition, Renouf and Sinclair (1991, p.128) present another definition of collocation relevant to the digital historical period of collocation study,¹⁴⁴ which states that collocation is “a sequence of words that co-occur more often than would be expected by chance”. Hence, to generalise from recurrent word pairs in a particular corpus, excluding word pairs whose recurrence may be an accident of the sampling process, researchers introduce the concept of statistical association (Sinclair, 1966, p. 418). This definition is adopted in this research allowing collocations in the data to be computed and verified mathematically. In an explanation of this linguistic phenomenon, Sinclair (1991) writes:

Collocation is the occurrence of two or more words within a short space of one another. The usual measure of proximity is a maximum of four words intervening. Collocations can be dramatic and interesting because unexpected, or they can be important in the lexical structure

144 Louw (2007b) distinguishes between two historical periods of the study of collocations: the “analogue period” and the “digital period” (pp.91-2). In the analogue period, investigators like Palmer (1974, p.94) had no easy way of verifying the frequency of collocates in the language, because computers were insufficiently powerful to perform this task and corpora were too small for results to be regarded as trustworthy. In contrast, the digital period, which Louw claims began with the advent of the first edition of Sinclair’s *COBUILD English Language Dictionary*, is reason that a notion such as semantic prosody exists (Louw, 2007b, p.92).

of language because of being **frequently** repeated. Each citation or **concordance** line exemplifies a particular word or phrase. This word or phrase is called the **node**. It is normally presented with other words to the left and the right, and these are called **collocates**. The collocates can be counted and this measurement is called the **span**... attention is concentrated on lexical co-occurrence independently of grammatical pattern or positional relationship. (Sinclair, 1991, p. 170, emphasis added).¹⁴⁵

This research also adopts the related notion of the ‘Extended Lexical Unit’ (ELU) (Sinclair, 1996, pp. 87-8; Sinclair, 1998, pp.14-15; Sinclair, 2004a, p.141). This is based on the view that a given “lexical item” is characterised in terms of “five categories of co-selection”, this co-selection resulting in an ELU. Two of the five categories are obligatory: the core and the SP. The other three are optional: collocation (the co-occurrence of words with one another), colligation (the co-occurrence of words with grammatical choices) and semantic preference (the restriction of regular co-occurrence to items which share a semantic feature). Sinclair also claims that SP, “the unit’s attitudinal and pragmatic function”, is the reason for the speaker/writer to make use of that unit as a particular linguistic choice. In line with this, he states that the optional categories of co-selection are variable – their variations signifying different shades of meaning of a lexical unit in language in use. Sinclair argues that each of the five categories of co-selection should be empirically examined both syntagmatically and paradigmatically. He believes that the syntagmatic dimension of collocation is the most straightforward and obvious relationship because all the elements are present in the ELU. The paradigmatic axis is less apparent, because it involves elements which might have been, but are not, present in the text.

Also, Sinclair (1998) defines what he calls “mutual collocation” as the feature where the words involved collocate with a different third word, usually with the same item or items, and illustrates it with the following example: “whereas *manual* and *restoration* are significant collocates of *work*, they themselves do not co-occur significantly” (p.14). Furthermore, Sinclair (2004a) elaborates on his definition of ELU which includes the previously mentioned terms “node” (the item whose collocations are studied), “collocate” (the co-occurring item), and “span” (the number of lexical items on each side of the node that are examined). He distinguishes “meaning by collocation” from both the “conceptual

145 Available from: [https://link.springer.com/chapter/10.1057/9780230624856_8], [Accessed 18 July 2019].

or idea approach to the meaning of the word and contextual meaning” (Sinclair, 2004a, p.141).

To conclude, the definition adopted in this research derives from the Firthian sense and is one that requires a syntagmatic and quantitative approach to collocation which is accepted by many corpus linguists (e.g., Greenbaum, 1974; Halliday and Hasan, 1976; Hoey, 1991; Sinclair, 1991; Stubbs, 1996; Partington, 1998; McEnery and Wilson, 2001; Hunston, 2002). These definitions state, as seen above, that collocations of a given word are “statements of the habitual or customary places of that word” (Firth, 1957, p.196). Hence, this research examines the units of meaning of collocational frameworks containing a word representing a concept related to nature and its surroundings. It then labels each with the evaluative prosodies and the discourse prosodies relevant to the theme of nature in the Qur’an. It adopts an established theoretical framework that investigates recurrent word combinations (e.g., recurrent patterns of natural phenomena in the Qur’an), on the basis that the meaning of a word can be determined at least in part by its typical collocate sets (e.g., grammatical patterns).

2- *Criteria for operationalising collocation identification*

For at least the past five decades, research has focused on three criteria for identifying collocations (as in Gries, 2013; Brezina et al., 2015). These are: (i) *distance*, (ii) *frequency*, and (iii) *exclusivity* and are explained as follows:

The distance specifies the span around a node word (the word we are interested in) where we look for collocates. This span is called the ‘collocation window’. The **distance** of the collocate from the node can be as little as one word if we interested, for instance, in the adjectives immediately preceding a noun in English, or as much as a span of four or five words on each side of the node, if we are interested in more general associations. The second criterion, **frequency** of use, is an important indicator of the typicality of the word association. For instance, the noun *love* frequently occurs with the preposition *in* and therefore *in love* is an important ‘chunk’ in the English language. However, it can also appear in front of many other nouns, such as *case*, *fact*, or *school*. Consequently, the relationship between *love* and *in* is not **exclusive**. On the other hand, *love* is much more strongly and exclusively connected with the noun *affair*; when the word *affair* appears in the text, there is a large probability that the preceding word is *love* (Brezina et al., 2015, p.140, emphasis added, italics in original).¹⁴⁶

146 Available from: [<https://benjamins.com/catalog/ijcl.20.2.01bre/fulltext/ijcl.20.2.01bre.pdf>], [Accessed 18 July 2019]

Gries (2013) suggests that, in addition, there are three more criteria that should be considered: (iv) *directionality*, (v) *dispersion* and (vi) *type-token distribution* among collocates (Gries, 2013, p.159; cf. Brezina et. al, 2015, pp. 140-41). Directionality refers to the fact that the strength of the attraction between two words is seldom symmetrical.¹⁴⁷ For instance, the word *affair* has a stronger relationship with the word *love* than *love* with the word *affair* because *love* co-occurs with other words other than *affair* more often than vice versa. Dispersion is the distribution of the node and the collocates in the corpus. Gries argues that although the most common statistics in corpus linguistics are frequencies of occurrence and frequencies of co-occurrence of two or more linguistic variables, these frequencies in isolation can be misleading. He gives an example from Leech et al. (2001)¹⁴⁸ to show how raw frequencies can be misleading indicators of the overall importance of words in a corpus (Gries 2008, p.404). His example shows that the words *HIV*, *keeper*, and *lively* are about equally frequent in the British National Corpus (BNC), which would usually be interpreted as an indicator of their overall importance. A look at how these words are distributed in the corpus, however, suggests a very different result. While *lively* and *keeper* both occur in 97 of 100 equally-sized corpus parts, *HIV* occurs in only 62, which indicates that *HIV* is much more specialised. This assessment, he says, is supported when Leech et al. compute a more refined measure of dispersion for *lively*, *keeper*, and *HIV* which is: 0.92, 0.87, and 0.56 respectively. The third criterion which Gries (2013) claims to be significant is the *type-token distribution*, which considers not only the strength of a given collocational relationship but also the level of competition for the slot/s around the node word from other collocate types. To illustrate, he says that in the BNC, there are about 13 thousand different collocate types, which compete with *affair* for a slot near the word *love*. In addition to the criteria mentioned above, Brezina et al. (2015) add a seventh feature, which is the *connectivity between individual collocates*. In this regard, they state:

Collocations of words do not occur in isolation but are part of a complex network of semantic relationships which ultimately reveals their meaning and the semantic structure of a text or corpus (p.141).

147 Brezina et al. (2015, p.140) explain this as follows: an example of a symmetrical collocation window is a semantic prosody of three words on the left and three words on the right. An example of an asymmetrical collocation window is when there are zero words on the left and three words on the right and these produce asymmetrical results with any collocation measure.

148 See also Leech, G. N., Rayson, P. & Wilson, A. (2001). *Word frequencies in written and spoken English: based on the British National Corpus*. London: Longman

They argue, as seen above that collocates should not be considered in isolation but rather as part of a larger collocation network. The origin of such a view is seen in work of Phillips (1983;1985;1989) and later investigated in a number of studies using both general and specialised corpora (e.g., Alonso et al. 2011; McEnery 2006; Williams 1998).¹⁴⁹

Phillips (1989), who preferred to call collocation networks “lexical networks”, claims that they can be employed to operationalise the psychological notion of the “aboutness” of a text. He argues that these networks constitute a distinct level of linguistic analysis, which cannot be explained by traditional linguistic theory; hence, requiring a deeper understanding of lexical processes and their interconnections via collocation networks (p.52). In addition, Williams (1998) examines the lexical structure of research articles on plant biology, and proposes a stepwise method, which starts with a single initial node and its collocates, and from there gradually builds a complex collocation network by considering each of the collocates as a new node and adding a network of collocates around each such node. He compiles his initial nodes from the first 50 lexical words in the frequency list based on his text or corpus, and explores collocations using the Mutual Information (MI) score with several cut-off points (4, 5 and 6) (Williams, 1998, pp.155-7). His work was found to have merit in that it can be replicated, yet his choice of span is not clear (Brezina et al., 2015, p.143). Another study which applied the use of collocation networks as one of many tools for exploring discourse related to swearing in English is McEnery (2006). His work is more advanced than that of Williams (1998) in that it builds directional collocation networks (direction marked by an arrow) starting with the nodes (compiled via the keyword procedure). The association measure employed by McEnery (2006) is the squared version of MI (MI²), with a cut-off point of 3 and a span of +/- 5 words around the node. Moreover, based on Phillip’s view on collocation networks, the previous explorations via collocation networks, and for a wider range of analyses with improved visual representation,¹⁵⁰ Brezina et al. (2015) introduce *GraphColl*: a new tool developed by the authors which builds collocation networks from user-defined corpora. To conclude, the principle of connectivity as seen in collocation networks and the relatively

149 See also (Brezina et al., 2015)

150 The authors summarise the rationale behind suggesting the use of *GraphColl* as opposed to replicating Phillip’s work in three points: (1) although Phillips clearly intends to identify syntagmatic lexical sets (i.e., collocations), his cluster analysis revealed the paradigmatic rather than syntagmatic relationships between words (i.e., collocate sets). (2) it is unclear how he represented his data, and whether the output of a cluster analysis (a dendrogram) can be changed into a digraph and how directionality is demonstrated in these visual representations. (3) the software used by Phillips is not available anymore (Brezina et al., 2015, p.142-3).

novel tool for collocation identification in corpus linguistics are important in this thesis for the study of SP in relation to interconnected discourse prosodies of the Qur'an. Collocation networks are visually represented using the recently developed data visualisation tool LancsBox (Version 4.0)¹⁵¹ to compare the Arabic text of the Qur'an with five of its English translations.

To operationalise at least two of these criteria in exploring collocations, most corpus linguists agree that it is plausible to statistically distinguish the frequent and exclusive collocations in any text corpora (e.g., McEnery and Wilson, 2001; Evert, 2004; 2005; 2008; McEnery and Hardie, 2012; Brezina et al., 2015; Brezina, 2018). In this regard, McEnery and Wilson (2001) write:

Given a text corpus, it is possible to determine empirically which pairs of words have a substantial amount of 'glue' between them and which are, hence, likely to constitute significant collocations in that variety rather than chance pairings (p. 86).

To do so, these scholars who follow empirical models of exploring collocations stress the importance of applying statistical measures which will guarantee that the 'glue' [association or affinity] between the node and collocate is strong and that a collocation is sound. They claim that this entails finding the following measurements:

- a) *OII*-known directly as $f(n, c)$ [frequency of collocation]
- b) *RI*-calculated from $f(n)$ multiplied by window size [the part of the corpus which has the collocations]
- c) *CI*- known directly as $f(c)$ [frequency of collocate]
- d) *N*-known directly as the size of the corpus size [number of tokens in the corpus]
- e) *Collocation window size* [number of words around the node, e.g., four words to the right and four to the left make a collocation window size which is 8] (based on Evert, 2008, pp.11-7; Brezina et al., 2015, pp.144-5; Brezina, 2018, pp. 69-71, my brackets)

¹⁵¹ *LancsBox* is a new-generation software package for the analysis of language data and corpora developed at Lancaster University Available from: [<http://corpora.lancs.ac.uk/lancsbox/>].

In addition, they claim that these measurements can be represented in a contingency table as in Table 11, which displays the representation of the observed frequencies for each of the explored collocations in the corpus.

Table 11: A contingency table for the observed frequencies (table taken from Brezina, 2018, p.70 and based on the notation in Evert 2008)¹⁵²

	<i>Collocate present</i>	<i>Collocate absent</i>	<i>Totals</i>
Node present	O_{11}	O_{12}	$R_1 \times \text{window size}$
Node absent	O_{21}	O_{22}	R_2
Totals	C_1	C_2	N

Expected frequencies can also be calculated and compared to the observed frequencies as in the following equation from Brezina (2018, p. 69):

$$\text{Expected frequency} = \frac{\text{node frequency} \times \text{collocate frequency} \times \text{window size}}{\text{number tokens in the text or corpus}}$$

These frequencies can also be displayed in a contingency table of expected frequencies for each of the collocations, such as the following:

Table 12: A contingency table for the expected frequencies (table taken from Brezina, 2018, p.71 and based on the notation in Evert 2008)

	<i>Collocate present</i>	<i>Collocate absent</i>	<i>Totals</i>
Node present	$E_{11} = \frac{R_1 \times C_1}{N}$	$E_{12} = \frac{R_1 \times C_2}{N}$	R_1
Node absent	$E_{21} = \frac{R_2 \times C_1}{N}$	$E_{22} = \frac{R_2 \times C_2}{N}$	R_2
Totals	C_1	C_2	N

The calculation of the Expected frequency E , or what Brezina et al. (2015, p.144) and Brezina (2018, p.71) define as the “random occurrence baseline”, is based on a null hypothesis that the node and collocate are independent and serves as a reference point for the interpretation of the observed frequency O of the word pair. Accordingly, the pair is only considered collocational if the observed frequency is substantially greater than the expected frequency (Evert, 2008, p.17).

In addition, as the previous literature on exploring collocation shows, the use of an association measure is necessary to automatically perform the frequency calculations above

¹⁵² Both tables illustrate the observed and expected frequencies of the collocation (love, affair) in the original texts.

and rank word pairs to filter the best collocations based on their association scores. This is ascertained by Brezina et al. (2015), who write:

For the majority of association measures, the statistical procedure for identification of collocates involves two steps: (i) establishing a random co-occurrence baseline (expected frequencies), (ii) comparing observed frequencies with the random co-occurrence baseline. All the widely used association measures are therefore based on comparison of (some of) the values in two tables: the contingency table with observed frequencies [e.g., Table 11] and the contingency table with the expected frequencies [e.g., Table 12] (p.144, my brackets).

Accordingly, association measures are the means of highlighting collocations along the two abovementioned dimensions of *frequency* and *exclusivity*. Frequency refers to the number of instances in which a node and collocate occur together in a corpus, and exclusivity to a specific aspect of the collocation relationship where words occur predominantly in each other's company (as in Brezina 2018, p.71). Moreover, Brezina et al. (2015, p.144) discuss the applications of collocation measures on the different aspects of the collocational relationship between words. For example, they say that MI score highlights rare and unique combinations, often terms or compounds which may stand out but are not necessarily representative through frequency, while Log Dice identifies combinations of words which appear both uniquely in each other's company and also frequently in discourse. Also, Delta P, which takes directionality of collocation into consideration, produces two values for each collocational relationship: i) strength of attraction between the node and collocate and ii) strength of attraction between the collocate and the node.¹⁵³ Log likelihood, which is implemented in this research, provides two opposite hypotheses as seen below:

• **Hypothesis 1.** $P(w^2|w^1) = p = P(w^2|\neg w^1)$

• **Hypothesis 2.** $P(w^2|w^1) = p_1 \neq p_2 = P(w^2|\neg w^1)$

Hypothesis 1 is a formalization of independence (the occurrence of w^2 is independent of the previous occurrence of w^1), Hypothesis 2 is a formalization of dependence which is good evidence for an interesting collocation. It claims that the co-occurrences of both words are dependent on each other. Moreover, the Log likelihood score tells us how much more

¹⁵³ See also Gries (2013)

likely a collocation would occur under one of the hypotheses than the other (as stated in Manning and Schütze, 1999, pp.172-4).

Finally, Brezina et al. (2015) introduce what they call “a collocation parameters notation (CPN)” and claim that it is essential for the sake of replicating, comparing, and reporting results; because it captures all the important parameters which can affect collocate identification. This can be seen in the example shown in the table below:

Table 13: An example of a collocation parameters notation (CPN) (table adopted from Brezina et al., 2015, p.157; Brezina, 2018, p.75)¹⁵⁴

<i>Statistic ID</i> ¹⁵⁵	<i>Statistic name</i> ¹⁵⁶	<i>Statistic cut-off value</i>	<i>L and R span</i>	<i>Minimum collocate freq. (C)</i>	<i>Minimum collocation freq. (NC)</i>	<i>Filter</i>
4a	MI2	3	5L-5R	5	1	function words removed or no filter applied

} Examples

Accordingly, the identification of collocation extraction in any corpus-based methodology is represented via CPN as in the following example taken from the previous table:

CPN = MI2(3), L5-R5, C5-NC1; function words removed (example)

Similarly, the method of extracting collocations of nature terms in the Qur’an and its translations is identified each collocation is individually named with a “frequency signature” via the values found in accessed via Brezina et al.’s (2015) CPN, as will be seen in the following chapter. Moreover, accordance with Evert’s model of collocation extraction (2008). Details of this model will also be discussed in the following chapter. The following subsection will depart from the criteria of identifying collocations to provide a brief overview of the general approaches that scholars adopt in finding collocations depending on their perspective of the phenomenon of collocation.

¹⁵⁴ In Brezina et al. (2015, p.157), this table referred to McEnery’s (2006) settings of identification of collocations of swearing in English; in the *Minimum collocation freq. (NC)* column it has the phrase: N/A [no minimum collocation cut-off point] and in the *Filter* column: “function words removed; the strongest collocates considered”.

¹⁵⁵ In their Appendix 1, which lists the default association measures implemented by *GraphColl*, Brezina et al. (2015, pp168-72) give each association measure a unique identifier called a *Statistic ID*. They are also available in Evert (2004, Section 3; 2008).

¹⁵⁶ The name of the employed *association measure*.

3- Approaches to the study of collocation

Over the past years, two general definitional-type approaches to collocation have emerged (Herbst, 1996, p.379; Nesselhauf, 2004a, p. 11; 2004b, p. 1).¹⁵⁷ One considers collocation as “the co-occurrence of words at a certain distance and distinguishes between frequent and infrequent co-occurrences” (Mohammed, 2012, p. 23). This view has been referred to as the *statistically oriented approach* (Herbst, 1996, p.380); the *frequency-based approach* (Nesselhauf, 2004a, p.12; 2004b, p. 1), or the *distributional approach* (Evert, 2005, p.15). In the second approach, collocations are described as “one particular type of word combination, or as a phraseological unit which is semantically opaque with a degree of fixedness” (as in Mohammed, 2012, p. 24). This view has been called the *significance-oriented approach* (Herbst, 1996, p.380), the *phraseological approach* (Nesselhauf, 2004a, p.12; 2004b, p. 1), or the *intensional approach* (Evert, 2005, p. 16). The distributional approach, which is adopted in this research, is closest to Harris’ views on distributional semantics and Firth’s (1951;1957) definition of collocation. For example, according to Harris (1951):

In both the phonologic and morphologic analyses, the linguist first faces the problem of setting up relevant elements. To be relevant these elements must be set up on a distributional basis: *x* and *y* are included in the same element *A* if the distribution of *x* relative to the other elements *B*, *C*, etc. is in some sense the same as the distribution of *y*. Since this assumes that the other elements *B*, *C*, etc., are recognized at the time when the definition of *A* is determined, this operation can be carried out without some arbitrary point of departure only if it is carried out for all the elements simultaneously. The elements are thus determined relatively to each other, and on the basis of the distributional relations among them (Harris, 1951, p.7, italics in original).

As seen above, Harris’ distributional approach (1951), which he first presents as a procedure for phonemic analysis then turns into a general methodology employed to every linguistic level,¹⁵⁸ is based on the view that the semantic similarity between two words (e.g., the variables *x* and *y*) is, in fact, a function of the degree of the similarity of their “linguistic environments” [i.e. of the degree to which they can occur in similar contexts]. In this sense, he claims that similarity in distributions should be taken as an account for meaning itself, and therefore be used to build paradigmatic classes [e.g., collocate sets for

¹⁵⁷ See also (Mohammed, 2012, pp.23-25)

¹⁵⁸ See also (Lenci, 2008, pp. 3-4).

near-synonyms] out of distributionally semantic similar linguistic expressions (Harris, 1951, p.157). In this regard, one can say that Harris, like his later contemporary Firth (1951), views the distributional approach to semantic analysis as a solid empirical foundation; a view which is adopted in this research along with Firth's empirical view of collocation. Although Firth introduced the term collocation in 1951, it was not until 1957 that he defined this term which he called, "collocability" (Firth, 1957, p.194). He defined a collocation as a frequently recurrent word and argued that the meaning and usage of a word can be described mostly by its typical collocates: "You shall know a word by the company it keeps" (p. 179). Collocations in this Firthian sense can also be interpreted as empirical statements about the predictability of word combinations: they quantify the "mutual expectancy" between words and the statistical influence a word exerts on its neighbourhood (Firth 1957, p.181).

There are, in fact, many different approaches to recurrent lexical combinations and the inferred meanings of words, Lehrer (1974) mentions six approaches to collocation proposed by linguists (Lehrer, 1974 as cited in Crystal, 2011, pp. 173-83), in relation to the *collocational* or *selectional restrictions* which specify the conditions governing the permitted combinations of lexical items within in a given grammatical context. These restrictions are only concerned with the lexical level (they are not concerned with grammar) and are as follows: *lexical approach*; *semantic approach*; *transfer-features approach*; *lexico-semantic approach*; *semantic prosody approach*; and *metaphorical extension approach*. These are discussed in this review merely to illustrate the different trends in exploring collocations. In the following, each will be briefly defined. In this study, the lexical and the SP approaches to collocation will be adopted in conjunction with the distributional approach, as will be discussed below.

According to the *lexical approach* taken by Firth and his followers, "one level of or aspect of the meaning of the word is determined by its collocational environment" (Firth, 1957, p.195) (e.g., the syntagmatic level of language). In this regard, Firth writes:

Meaning by collocation is an abstraction at the syntagmatic level and is not directly concerned with the conceptual or idea approach to the meaning of words. One of the meanings of the word *night* is its collocability with *dark*, and of *dark*, of course, collocation with *night* (Firth 1957, p.196, my italics).

On the other hand, the *semantic approach* claims that co-occurrence restrictions are the result of the meanings of the lexical items themselves (rather than being a function of their collocates), i.e. collocations are reflections of the semantics of the lexical items which constitute them. Hence, these co-occurrence restrictions are to be stated in terms of what semantic features may occur together:

Lexical items such as *the vase broke* may be accounted for by saying that the verb *break* requires an object with the feature (+fragile), so that sequences such as *the book broke* is to be rejected as unlexical simply because *book* does not have this feature, namely (+fragile) (Lehrer, 1974, p. 176).

Another proposed approach for the study of collocation is the *transfer-features approach*. The difference between the previous approaches and this approach can be illustrated by the example *Susan broke the bird*. It has been mentioned above that the verb ‘break’ requires an object (or subject in the ergative) with the feature (+fragile). This makes this sentence an ‘unlexical construction’; however, this approach would justify the lexicality of the meaning of this sentence by saying that the feature (+fragile) is transferred to *bird* and one semantic implication here is that the *bird* mentioned in the example is fragile.

To resolve the conflicting views on the approaches to collocational semantics, the *lexico-semantic approach* was proposed by Lehrer (1974, p.183). In this approach, both the lexical item with its semantic features and the collocational environment (e.g., collocates) are taken into consideration in the study of collocation. To extend this approach to the study of the semantic features (understood as forms of connotation) of both the lexical item and the collocates, one can employ the *semantic prosody* (SP) approach. Put simply, SP (as it will be defined in 3.1.3), means the *good* or *bad* connotations a combination of lexical items conveys to the speaker and/or reader (Rundell, 2000, p.1).¹⁵⁹

Finally, there is the *metaphorical extension approach* which sheds light on the frequent metaphorical use of lexical items. It claims that lexical items which undergo metaphorical extensions (while retaining their central or original meaning) tend to enter into a range of collocations compatible with the number of shades of meaning they acquire (as in Dickins et al., 2017, pp.204-6). A study which applies this approach in the analysis of the English

¹⁵⁹ See also [<http://old.hltmag.co.uk/jul00/idea.htm>]

collocation *break an appointment* is Poulsen (2005). Poulsen uses the the British National Corpus and generates 1000 concordance lines with *break*¹⁶⁰ and 908 concordances with *appointment*, where she only finds five examples of *break an appointment*. In her analysis of the verb *break*, she found that it is mainly followed by an object which denotes the container. The phrasal verb *break open* also focuses on the container, specifying that the “focus is on getting access to something inside it”, while *break into* and *break out* respectively offer “the contrasting perspectives of entering a container from the outside” and “taking something out from the inside of a container”. She also mentions *breaks eggs*, which she says refers not only to “the breaking of a container, the egg shell, but also implies getting at the contents”. Finally, she remarks on the metaphorical extension of the verb *break* in the collocation *break news*. She argues that in this metaphorical sense, “there also seems to be a merging of container and contents”. The example she gives is the following sentence: *She did not want to have the news broken to her carefully*, and claims that the central meaning of the verb *break* here extends to include the metaphorical sense (p.169). Other lexical items which do not extend their central meaning, on the other hand, will be confined to a highly restricted set of collocations (Lyons 1977, pp. 262-63).

To recap, based on the concept of collocation reviewed here, this research adopts the empirical Firthian notion of collocation (1957) together with some aspects of the collocation criteria of both Gries (2013) and Brezina et al. (2015). This is especially important to this research because the textual analysis and comparison of the texts are based on this understanding of identifying collocation through the chosen criteria of connectivity, distance, frequency, and exclusivity. The following section will provide an overview of the relationship between collocation and SP as it appears in the literature. It mainly focuses on Sinclair’s and Stubb’s views and insights on the units of lexical meaning and the notion that SP is a close relative to collocation and that the two always co-occur.

4- *Collocation and SP as a collocational phenomenon*

As mentioned above, the term *extended unit of meaning* was introduced by Sinclair in 1996. Sinclair was a student of Firth’s and, like Firth, he adopts a contextual/functional approach to meaning. Based on this approach, he argues that corpora can be used to provide evidence for new and revolutionary semantic insights. His model consists of four types of co-

160 She analyses the different forms of *break*, which are *break*, *breaks*, *broke*, *broken* and *breaking* (See also Poulsen, 2005, p.13)

occurrence relations in extended lexico-semantic units (cf. Stubbs, 2001, p.64):¹⁶¹ *collocation, colligation, semantic preference, and SP*. In this regard, Sinclair (1996, p. 94) writes:

So strong are the co-occurrence tendencies of words (collocation), word classes (colligation), meanings (semantic preference) and attitudes (semantic prosody) that we must widen our horizons and expect units of meaning to be much more extensive and varied than is seen in a single word (as cited in Zethsen 2006, p.280, Zethsen's parentheses).¹⁶²

Similarly, Stubbs says that in the contextual/functional approach to meaning and corpora, it is essential to study the meaning of a lexical item via the extended unit of meaning (e.g., collocation and SP) (Stubbs, 2001, p.22). It should be noted, however, that Sinclair's extended units of meaning, which are adopted in this research, are not associated with completely fixed expressions such as idioms. Instead, they are a system (or framework, or matrix) in which there is room for variation. Another observation Stubbs (2001, p.63) makes about Sinclair's model is that it contains two closely related central ideas:

1. Meaning is typically dispersed over several word-forms which habitually co-occur in text.
2. These co-occurring word-forms 'share' semantic features.

These two points reflect his central insight that collocation and SP are inseparable. The first point refers to the view that the 'lexical item', which is the primary carrier of meaning, is said to be a collocation when it habitually co-occurs with words in the text. The second argues that these collocations have similar features (e.g., connotative colourings are shared to form SP).

Finally, drawing on the work of Sinclair and Stubbs, this research adopts the position that meaning, and in particular *associative* or *evaluative meaning* and *attitudinal meaning* (two integral parts of the connotative meaning), cannot be limited to the lexeme. It claims that meaning is to be explored based on the extended unit of meaning. Furthermore, within a model of extended units of meaning, it is at the level of SP that evaluation can take place. SP is also an indispensable tool for eliciting speaker attitude and making qualified guesses at likely listener interpretation.

¹⁶¹ In his later work, Sinclair (2004a) argues for five co-selection categories (p. 141).

¹⁶² Available from: [<https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/viewFile/218/142>], [Accessed 19 July 2019]

5- Collocation and the lexical item

Sinclair (2004a, p.141) proposes that a given ‘lexical item’ in the Extended Lexical Unit (ELU) is characterised in terms of ‘five categories of co-selection’. Two of these are obligatory: the *core* and the *semantic prosody*. The other three are optional: *collocation* (the co-occurrence of words), *colligation* (the co-occurrence of words with grammatical choices) and *semantic preference* (the restriction of regular co-occurrence to items which share a semantic feature) (See Figure 17).

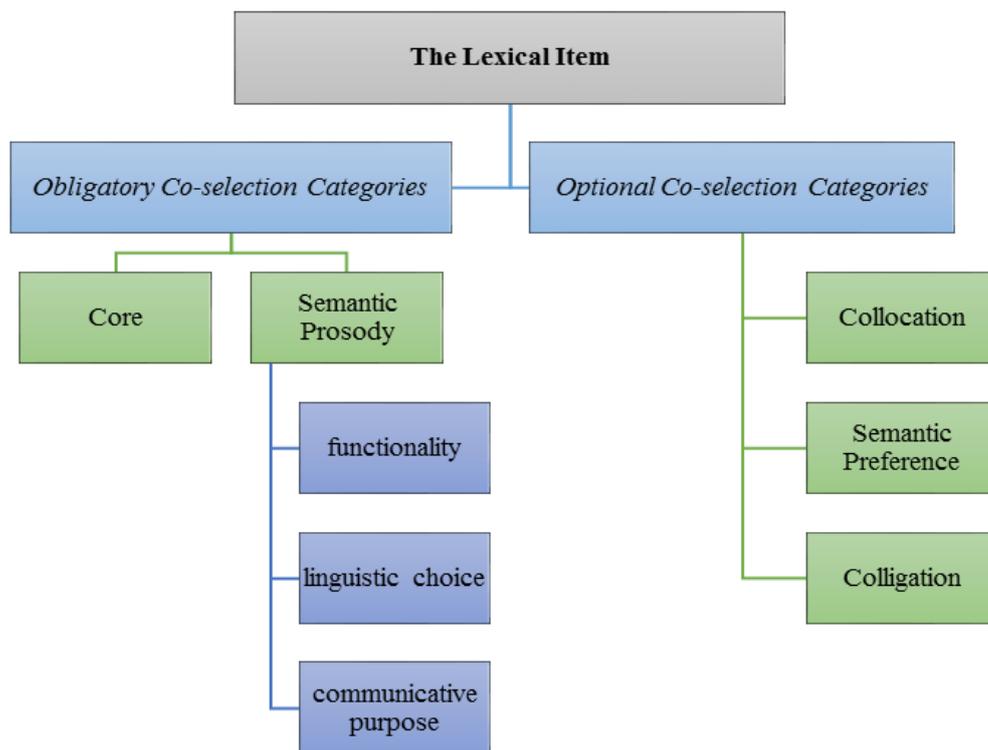


Figure 17: The Extended Lexical Unit (ELU)

The *core* word is either a word or a “combination of specific words” that gets together simultaneously to operate effectively as one meaning, for example, ‘of course’, ‘in fact’, etc. The *semantic prosody*, then, describes how these core words can be interpreted with a particular connotation through frequent association or occurrences with particular words (collocates in collocations) that have other meanings (See Section 3.1.3 for the discussion on the different features of semantic prosody). Collocations are composed of three components within the definition of ELU:

Node: the item whose collocations are studied;

Collocate: the co-occurring item;

Span: the number of lexical items on each side of the node examined (Sinclair, 2004a, p.141).

Colligation is another concept first put forward by Firth (1957) and defined by Hoey (2000, p. 234) as “the grammatical company a word keeps and the positions it prefers”. Thus, collocation is “the lexical company a word keeps”, whereas colligation is “the grammatical company the word keeps and the positions it prefers”.¹⁶³ Hoey (1997, p.4) classifies colligation into two sub-classes: “colligation in terms of textual position” and “colligation in terms of grammatical context”. Colligation in terms of the textual position involves the fact that a lexical item may tend to occur in a certain textual position; for instance, at the beginning or end of a text. Colligation in terms of grammatical context involves the fact that a lexical item tends to co-occur with “a particular grammatical category of items”, for instance, after modifying adjectives, quantifiers or before noun phrases. When a word has multiple meanings, each meaning will be associated with a different grammatical context, “with sense and a specific grammatical context in a direct relationship”.

Hunston (2007, p.266) defines *semantic preference*, also known as an attitudinal preference, as “the frequent co-occurrence of a lexical item with items expressing a particular evaluative meaning”. Thus, semantic preference describes the nature of the lexical items (collocates) that co-occur with the node word/s sharing similar semantic features. It is also regarded as the final steppingstone that makes it possible to progress from the concrete realities of collocation to the abstract perception of semantic prosody (For more on the interaction between semantic prosody and semantic preference see Section 3.1.3).

3.1.2 Lexico-grammar and collocation

In order to coherently investigate collocation in this research, a theory of language had to be adopted in which collocation is an adequately integrated element. Several linguists have proposed an integrated view of collocations: Sinclair on the Idiom Principle (1996); Biber and Conrad (1999) on Lexical Bundles; Hunston and Francis (2000) on Pattern Grammar; Stefanowitsch and Gries (2003) on Collostructions; Hoey’s Lexical Priming (2004; 2005; Hoey et al., 2007); Halliday and Matthiessen (2004) on Functional Lexico-Grammar; and Goldberg (2006) on Construction Grammar. All these theories are arguably similar in that they all agree that form and meaning are inseparable and that the unit of meaning in language is not the word in isolation but a construction or phrasal unit (at different levels

¹⁶³ This statement is based on both Firth and Hoey’s words.

of complexity).¹⁶⁴ The fact that the core claims of these approaches as to the interrelatedness of vocabulary and syntax and their inductively data-driven conclusions are broadly similar would tend, according to Hunston (2008, p.292), “to increase confidence” in all of them. This made it much less problematic to choose from among these approaches, one that can account for collocations of natural phenomena in the Qur’an. The theory which the researcher decided to principally rely on in the analysis of bigrams of natural phenomena in the Qur’an was the functional Lexico-Grammar (Halliday and Matthiessen, 2004). Of all the approaches considered above, this theory best suits the definition of SP adopted in this research, which focuses on both the semantic associations and pragmatic functions of SP which are uncovered through collocation extraction (quantitative analysis) and collocation interpretation (qualitative analysis and quantitative analyses in calculating SP).

Halliday first introduced the term “Lexico-Grammar” (LG), also called ‘lexical grammar’ (i.e., lexicon plus grammar), in 1961. He argues that the meanings conveyed by collocations are closely related to the patterns of lexico-grammar, the term ‘Lexico-Grammar’ being used in systemic functional linguistics (SFL) to emphasise the interdependence of and continuity between vocabulary (lexis) and syntax (grammar) (Halliday 1961; Halliday and Matthiessen, 2004). In this regard, Halliday and Matthiessen (2004, p.43) stress that lexis and grammar are “two poles of a single cline, or continuum” (See Figure 18).

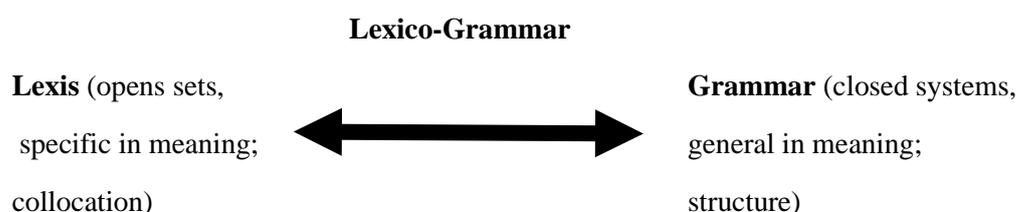


Figure 18: Halliday and Matthiessen’s lexico-grammar cline (2004, p.43)

Lexico-grammar is constantly “at work” creating meaning in the form of text. Insight is gained into this process by placing features of the wording, both lexical items and grammatical categories, in their syntagmatic (the company they keep, i.e. collocation) and paradigmatic contexts (the options that are open for them) in the system of the language. As the text unfolds, patterns emerge, some of which acquire added value by resonating with

¹⁶⁴ See also Hunston (2008, p. 292).

other patterns [collocations] in the text or the context of the situation [evaluative prosody]; a viewpoint in harmony with Partington et al. (2013, pp.8-9). The text itself is an instance; the resonance is possible because behind it lies the potential that informs every choice made by the speaker or writer [discourse prosody], and in terms of which these choices are interpreted by listeners and readers (Halliday and Matthiessen, 2004, p.43).

Moreover, Hunston and Francis (1991, p.107) define lexical grammar (i.e., lexico-grammar) as “the analysis of the behaviour of particular words in terms of their grammatical context”.¹⁶⁵ They claim that corpus-based analysis supports the concept of a lexical grammar based on the view that it is difficult to make a strict distinction between lexicon and grammar. This is because lexical items must be characterised in terms of their distributions in grammatical patterns and that many grammatical patterns are, as with lexical items, specific and therefore must be learned. Similarly, Halliday and Matthiessen (2004, p.43) claim that LG is the heart of the language and the abstract level of coding, and Sinclair (2004a, p. 177) describes it as a kind of grammar “laced or perhaps spiked with some lexis”. It is not a separate system or “module” separate from semantics but is instead an underlying component of the meaning-making system of a language. It has evolved to convey expressive and communicative functions (as in Gledhill, 2011, p.73). Finally, Benson et al. (1986, p.160) classify lexico-grammatical patterns in English into two classes of collocation, which can be said to reflect the relationship between grammar and vocabulary. They are lexical and grammatical patterns of collocation (which themselves have 7 and 8 sub-types respectively). Lexical collocations are made up verbs, adjectives, nouns, and adverbs in different possible combinations, while grammatical collocations involve verbs, adjectives or nouns combined with a preposition or a grammatical structure. Table 14 illustrates the different English lexico-grammatical patterns across collocations, with examples for each combination.

Table 14: English lexico-grammatical patterns across collocations (table taken from Benson et al., 1986, p.160)

<i>Pattern</i>	<i>Example/s</i>
V + N/P (or prepositional phrase)	compose music; set an alarm
V + Adv	walk heavily; argue heatedly
V + N	make a decision; take place/part
Adj + N	strong/weak tea
N + V	bombs explode; alarms go off
N1 + of + N2	a pride of lions; a bunch of keys

¹⁶⁵ Available from: [<https://epdf.pub/a-glossary-of-corpus-linguistics-glossaries-in-linguistics.html>], [Accessed 04 May 2018].

<i>Pattern</i>	<i>Example/s</i>
Adv + Adj	quite safe; deeply absorbed
N + Prep	ability in/at; kind of; changes in
N + to + Inf	an attempt to do it; years to come
N + that-clause	He took an oath that he would do his duty. We reach an agreement that ...; the fact that
Prep + N	on purpose; in fact,
Adj + Prep	tired of; bored with; angry with/at
Adj + to + Inf	ready to go; easy to learn; likely to be
Adj + that-clause	She was afraid that she would fail the exam; he was delighted that...
V + Prep	I believe in...
V + direct O + to + indirect O =V + indirect O + direct O	She sent the book to him. = She sent him the book.
V + direct O + to + indirect O (no movement for dative)	They mentioned the book to her.
V + direct O + for + indirect O =V + indirect O + direct O	She bought a shirt for her husband. =She bought her husband a shirt.
V + Prep + O	They came by train;
V + O + Prep + O	We invited them to the meeting.
V + to Inf	She continued to write.
V + bare Inf	Mary had better go.
V + V-ing	They enjoy watching TV.
V + O + to Inf	We forced them to leave.
V + O + bare Inf	She heard them leave.
V + O + V-ing	He felt his heart beating.
V + a possessive and V-ing	I cannot imagine their stealing apples.
V + that clause (rather uncommon)	The doctor suggests to me that I take vitamins.
V + O + to be + C	We consider her to be well-trained.
V + O + C	She dyed her hair red.
V (+ O1) + O2	The teachers asked (the students) questions. It took/cost (us) ten minutes/cents.
V + O + Adverbial	You carry yourself well (/like a soldier).
V (+O) + wh-clause/wh-phrase	She asked (us) why we had come.
It + V + O + to Inf	It surprised me to learn of her decision.
It + V + O + that-clause	It surprised me that our offer was rejected.
V + C (Adj or N)	He was a teacher.
V + C (Adj)	The food tastes good.

Lexico-grammar and collocation

In terms of Halliday and Matthiessen's lexico-grammar cline (2004) in Figure 18, collocation falls within the range of the lexical item as it interacts with grammatical patterns to produce meaning. Collocation which has meaning (as in Firth 1951) belongs to one of the poles of the lexico-grammatical cline (Figure 18) and falls under the open sets of lexical items (per Sinclair's definition 2004b). Meaning emerges from the habitual co-occurrences (collocations of lexical items) and sets (grammatical patterns) which govern the structures of collocations. Hence, the analysis of collocations in this research starts by finding the collocations of lexical items representing nature terms in the Qur'an and then moves on to link them to their grammatical patterns. The resulting lexico-grammatical patterns, when examined in their original context, reflect the meaning [specifically in this case, the SP of nature terms in the Qur'an].

The process of analysis of the LG patterns¹⁶⁶ to describe the SP of natural phenomena in this research entails the following: statistically examining them as lexico-grammatical features and categorising them into groups of pragmatic functions (discourse prosodies) and evaluative prosodies of nature in the Qur'an. From a theoretical point of view, this type of work resembles that of Partington et al. (2013), whose findings stress the presence of a significant link between SP and lexical grammar in corpus-based research. Their domain of research focuses on what they call, "semantic-or evaluative- prosody"; and SP in their research is considered a "notion devised for the first time within corpus-based lexical grammar". In this regard, they claim that:

lexical items are very largely co-selected by speakers in batches rather than singly and that therefore meanings in utterances, including and especially evaluative meanings, are prosodic-spreading over stretches of language-rather than atomistic in nature. (Partington et al., 2013, pp. 8-9)

Hence, Partington et al. (2013) provide a precedent in linking SP to lexical grammar. However, their work seems to only focus on evaluative semantic prosody, which they claim is synonymous with SP; they do not explicitly refer to the communicative side of SP (discourse prosody). Moreover, their work does not utilise this link to explore collocations via a mixed approach of finding collocations and SP.

On the practical level, the quantitative part of this analysis adopts Brezina's model for analysing LG features (2018). According to Brezina (2018), there are two approaches to particularly analysing lexico-grammar in corpora:¹⁶⁷

The first approach uses the 'Whole corpus' research design¹⁶⁸ and compares the frequencies of a linguistic variable (and its variants) in broadly defined subcorpora. The second approach employs the 'Linguistic feature' research design and carefully defines the contexts in which a particular variable can occur (i.e. its lexico-grammatical frame) and analyses factors which contribute to the occurrence of one variant of the variables as opposed to another (Brezina, 2018, p.102, Brezina's parentheses).

166 What is meant by 'LG patterns' here is the list of syntactic representation of bigrams of nature in the Qur'an (e.g., N+N, N+V, ADJ+N, etc.).

167 The term *lexico-grammar* is used for the features along the cline between lexis and grammar. These are explored using corpus techniques (i.e., statistical analysis of collocations of natural phenomena).

168 According to Brezina (2018), when analysing corpora, there are different research designs. The choice of which one of these to adopt has significant implications for the specific statistical procedures to be used with the data. In this regard, generally speaking and not entirely relevant to LG analysis, three types of research design can be distinguished: (1) whole corpus design, (2) linguistic feature design (3) and individual text/speaker design (p. 21).

This research employs the linguistic feature design to look at the general distribution of lexico-grammatical features in the Qur'an as the ST and its variant representations in the five TTs. Results of these frequencies will be represented in a stacked bar chart (as in Brezina, 2018, p. 104), which provides "a graphical representation of frequencies of a linguistic variable with multiple variants". To do so, the LG patterns of nature in the Qur'an are first determined, and their frequencies are calculated in the first part of this research. Then, they are compared to their counterparts in the English translations of the Qur'an, and visually represented using stacked bar charts as in Brezina's (2018). It would be possible to use these statistical results of the association measure (e.g. the Log Likelihood ratio) then to answer questions about the closest of the translations to the lexico-grammatical patterns of the Qur'an. However, in this research, the results will be used solely to portray the functional side of these patterns in revealing the pragmatic functions (i.e., discourse prosodies) of natural phenomena in the Qur'an, upon which comparison and/or evaluation of different translations is premised.

To conclude, lexico-grammar, as a component of the theoretical framework used in this research, has been adopted to account for collocation both quantitatively and qualitatively in the analysis of the bigrams of natural phenomena in the Qur'an.

3.1.3 Semantic prosody: a collocational phenomenon and cohesive device

The evolution of a definition for semantic prosody can be inferred from the chronological reviews of linguists' work on this feature of language (e.g., Stewart, 2010; Zhang, 2010; Cheng, 2013). When describing a word form that is likely to be followed by something positive or negative, Sinclair (1987b) referred to "good/positive" or "bad/negative" semantic profiles. For example, he claimed that there is a specific lexico-grammatical environment, "semantic environment" (1987, p.112), of the phrasal verb *set in*, namely its subjects, referring to some state of affairs, such as *rot*, *decay*, *despair*, or *bitterness*. Therefore, he described this phrasal verb as having a bad semantic profile (pp. 155-6). The co-occurrence of the most frequent collocates of *set in* thus created a negative default value, which explained both semantic associations of words and speaker's attitudes about their choice of words (Sinclair 1987b as cited in Cheng, 2013, p.1).

As a linguistic term, however, the notion of SP was first introduced by Bill Louw in 1993 (cf. Stewart 2010; Zhang 2010). Since then, it has become one of the most important concepts in corpus linguistics and has received significant attention both from Louw

himself and from other corpus linguists (e.g., Stubbs, 1995; Sinclair, 1996; Partington, 1998; Hunston and Francis, 2000; Louw, 2000; Hunston, 2007; Louw and Chateau, 2010). Louw borrows the word 'prosody' in semantic prosody from Firth (1957), who used it to refer to a phonological colouring which spreads beyond semantic boundaries. To give an example, the word *animal* has strong nasal prosody, the vowel sound of the letter 'a' being endowed with a nasal quality through assimilation, simply because 'a' is closely adjacent to the nasal sound of *n*. Another example Louw discusses is the following:

The exact realisation of the phoneme /k/, for example, is dependent upon the sounds adjacent to it. The /k/ of *kangaroo* is not the same as the /k/ of *keep*, because during the realisation of the consonant the mouth is already making provision for the production of the next sound. Thus the /k/ of *kangaroo* prepares for the production of /æ/ rather than /i:/ or any other sound, by a process of phonological colouring (Louw, 1993, p. 158).

Louw claims that, in the same way, lexical items share this particular phenomenon of prosody in lexical patterning. For example, he says, that an expression such as "symptomatic of" prepares the listener/reader to produce what follows; in this case something undesirable (e.g., management inadequacies, numerous disorders). In this regard, just as the realisation of a phoneme is influenced by the sounds which precede it as well as those which follow, the meanings of a keyword, by analogy, is influenced not only to words that appear before or after it but more generally by that keyword's close surroundings (Louw, 1993, p. 170). Inspired by the Firthian notion of prosody, Louw coins the term *semantic prosody* and gives it an initial definition: "a consistent aura of meaning with which a form is imbued diachronically by its collocates" (Louw, 1993, p.157). Later on, Louw further defines SP as "a form of (collocational) meaning which is established through the proximity of a consistent series of collocates" (2000, p.57), and Schmitt and Carter (2004) state that SP can characterise individual words as well as phrases (p.7).

Hoey (2000), who claims that the idea behind SP goes back to Sinclair (1991), highlights that use of words and phrases show a tendency to occur in a certain semantic environment. He says that SP "occurs when a word associates with a particular set of meanings" and is a "partial generalisation based on the collocates a word has" (Hoey, 2000 p. 232). Furthermore, Baker et al. (2006) define semantic prosody as "the way that words in a corpus can collocate with a related set of words or phrases, often revealing [hidden] attitudes" (p.58). It is a definition that is commonly discussed in the literature as the hidden, subliminal, covert, concealed, and subconscious quality of semantic prosodies (as in

Stewart 2010, p.32). This definition echoes Louw's (2000, p.58-9) view on the primary function of semantic prosodies, which is to express speaker/writer attitude or evaluation.¹⁶⁹ However, a speaker/writer, as Louw also claims, can also rhetorically violate a semantic prosody condition to produce some effect on the listener – for example, irony, insincerity, or humour can be explained by identifying violations of semantic prosodies (Louw 1993, p.137 and p.157). In the current research, just as in Partington's connotative classification of semantic prosody (2004b), a pleasant or favourable affective meaning is judged as positive while an unpleasant or unfavourable affective meaning is judged as negative. When an event is completely neutral, or the context provides no evidence of any other semantic prosody, the instance is labelled 'neutral'. These kinds or classifications of meaning are acquired by virtue of a grouping of words based on their usage, and do not derive from the conventional, dictionary meanings of the words themselves (as in Olohan, 2004, p.200). To conclude, SP has been defined in a variety of ways in the previous literature.¹⁷⁰ Whitsitt (2005), who, like Hunston (2007), has reservations about the way SP has been presented in the literature, claims that SP has been defined from three different points of view: primarily diachronic/synchronic (Louw 1993; Bublitz 1996), primarily pragmatic (Stubbs, 2001; Sinclair 2004b), and connotative (Partington 1998; 2004b; Stubbs, 2001; Hunston, 2002) (as in Stewart 2010, p.16). In this research, the analysis of the SP of natural phenomena in the Qur'an combines the pragmatic point of view (discourse prosody) with the connotative (evaluative prosody). It thus provides a holistic approach to SP as both a corpus-linguistic feature that can be utilised in corpus-based studies (e.g., this research on SP of nature in the Qur'an and translation studies in general).

1- Evaluative and discourse prosodies: two defining features of SP

An examination of the previous literature on SP reveals that it has been defined in terms of either evaluative prosody or discourse prosody or as a conjunction of both (as in Sinclair, 1996; Hunston and Thompson, 2000; Louw, 2000; Stubbs, M., 2001; Hunston, 2002; Partington, 2004b). Both evaluative prosody and discourse prosody are aspects of the

¹⁶⁹ See also Xiao and McEnery (2006). Available from:

[http://www.lancaster.ac.uk/fass/projects/corpus/ZJU/xpapers/Xiao_collocation.pdf], [Accessed 16 July 2018].

¹⁷⁰ To give a few other examples in addition to those already discussed in the main body of the thesis, SP has been defined, for example, as a way by which apparently neutral terms come to carry positive or negative associations through frequently occurring in particular collocations and how they create an aura of meaning capable of affecting words around them (Coffin et al., 2004, xxi) and (Gavioli 2005, p.46); and as a pattern (Berber-Sardinia, 2000, p.94); and as a phenomenon (Lewandowska-Tomaszczyk, 1996, p.153). More definitions of SP can be found in (Stewart, 2010).

definition as mentioned above of SP. *Evaluative prosody* is defined as “a form of meaning which is established through the proximity of a consistent series of collocates, often characterisable as positive or negative [or neutral]” (Louw, 1993, p.158-9, my brackets). It is contingent upon co-text¹⁷¹ and is “inferable employing observation of the word’s habitual co-occurrences” (Louw, 2000, p.50). It is “spread over a unit of language which potentially goes well beyond the single orthographic word and is much less evident to the naked eye” (Partington, 2004b, pp.131-32). Moreover, it is “a positive or negative association a word takes that is expressed covertly” (Hunston and Thompson, 2000, p. 38); and is a result of transferred meaning, “where a word that is typically used in a particular environment takes on connotations from that environment” (Hunston 2002, p. 61). On the other hand, the *discourse prosody*, or the attitudinal meaning or pragmatic function, reflects the attitude of its speaker or writer towards any given pragmatic situation. It is defined as follows:

A discourse prosody is a feature which extends over more than one unit in a linear string. (. . .) Discourse prosodies express speaker attitude. If you say that something is *provided*, then this implies that you approve of it. Since they are evaluative, prosodies often express the speaker’s reason for making the utterance therefore identify functional discourse units (Stubbs, 2001, p.65, italics in the original).

This aspect of SP emphasises the functional side of SP (e.g., it can be a cohesive device) in integrating an item with its surroundings and expressing “something close to the function of an item” (Sinclair, 1996). It is said to have two main functions. Firstly, besides showing the speaker’s or writer’s attitude (Sinclair, 1996, pp. 87-8; Stubbs, 2001, p.65), it has a pragmatic function which is to maintain discourse coherence, as asserted by Stubbs (2001) when he writes:

I will prefer the term ‘discourse prosodies’, both in order to maintain the relation to speakers and hearers, but also to emphasise their function in creating discourse coherence (Stubbs, 2001, p.66).

Secondly, it can be a secondary yet no less attitudinal function to (synchronously) invoke irony through “deliberate injection of a form which clashes with the prosody’s consistent series of collocates” (Louw, 2000, p. 60). In essence, it is produced by deviations from habitual co-occurrence patterns, yielding “a collocative clash, which is perceived albeit subliminally by the reader” (Louw, 1993, p.158). Figure 19

¹⁷¹ *Co-text* is the text occurring around a node, as can be seen in a concordance (Baker, 2006, p.199).

summarises these views and presents them as two pivotal dimensions of SP. SP itself is both a component of the theoretical framework used in this research and a descriptive corpus linguistic feature occurring in the references to natural phenomena in the Qur'an which the research investigates.

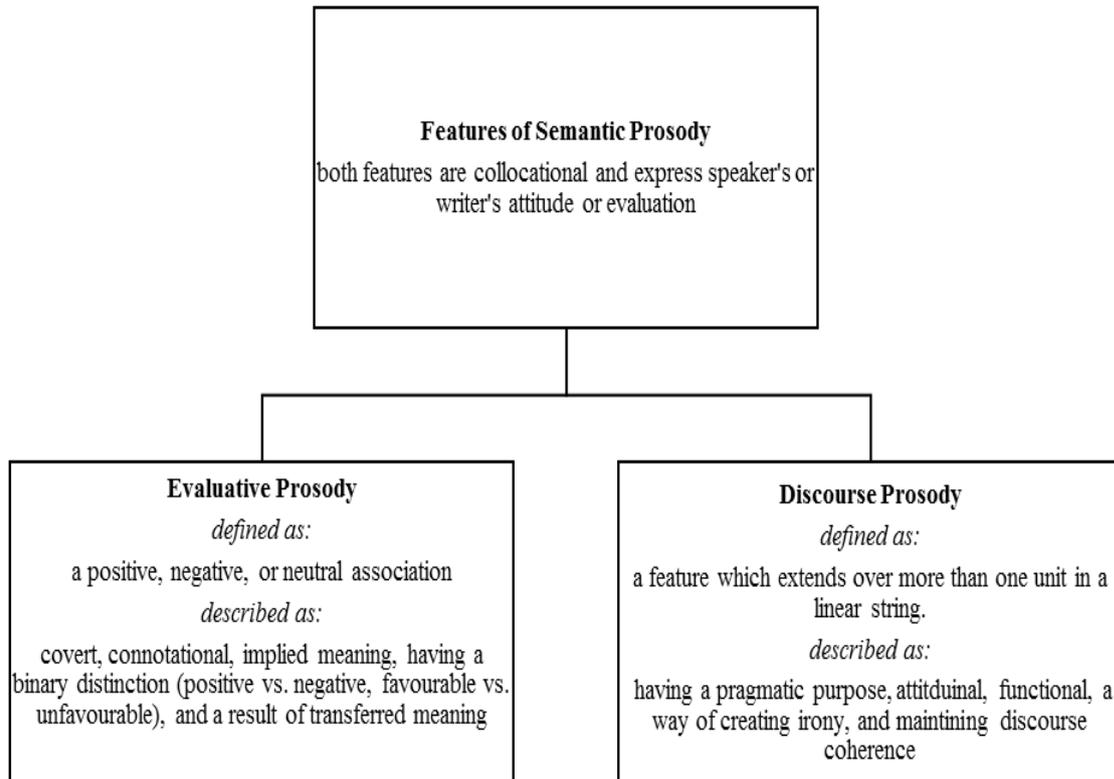


Figure 19: The defining features of SP in this research

2- *Semantic prosody and semantic preference*

Having defined the features of SP, it is necessary to disambiguate it from its sibling corpus linguistic feature in Sinclair's hierarchical model of the Extended Lexical Unit (2004).¹⁷² A distinction between semantic prosody and semantic preference is made by Sinclair (1996; 1998), Stubbs (2001), and Partington (2004b). Their views are summarised as follows:

- Semantic preference and semantic prosody have different operating scopes: the former relates the node item to another item from a particular semantic set, whereas the latter can affect wider stretches of text.

172 See also Figure 15.

- Semantic preference can be viewed as a feature of the collocates, while semantic prosody is a feature of the node word. The two also interact; while semantic prosody “dictates the general environment, which constrains the preferential choices of the node item”, semantic preference “contributes powerfully” to building semantic prosody (Partington, 2004b, p.151).

3- *Semantic prosody and meaning*

Sinclair (2003) states that semantic prosody is sometimes classified under “connotation” (compared to ‘denotation’ for the extensional range of the word) or at times referred to as ‘pragmatic’ meaning. He further claims this kind of meaning is structurally important and essential for the understanding of texts. It is ‘semantic’ because it involves meaning, and it is a ‘prosody’ because it typically ranges over combinations of words in an utterance rather than being attached to just one (Sinclair, 2003, p.117). The key difference from the traditional notion of connotation, however, as McEnery and Hardie claim (2012, pp.135-37),¹⁷³ is that semantic prosodies are not necessarily accessible to intuition, which is often used to make judgments about the connotations of a word. Rather, SP can *only* be discovered by analysis of concordance, as Louw argues (1993, p.159). In support of this view, Hunston (2002, p142) suggests that while SP may in some cases explain connotations which exist intuitively, in other cases the positive or negative associations of a given lexical item may not be accessible to conscious knowledge; this is in fact what Tognini-Bonelli (2001, p.114) refers to when saying that SP is mainly engaged at the subconscious level.¹⁷⁴

The model of analysis for the exploration of collocational behaviour and SP in this research adheres to a definition that is inclusive of semantic prosody as a type of connotative meaning (not necessarily synonymous with it); that is, SP can be said to represent two dimensions or levels of meaning in Dickins’ classification of type of connotative meanings (Dickins, 2019).¹⁷⁵ Based on this classification, one viewpoint might be that SP falls within the collocational meaning as one of the components of connotative meaning. This view is supported by Xiao and McEnery (2006, p.85) who, when writing about the association between connotative meaning and the lexical environment, claim that connotation can be collocational or non-collocational and that SP is purely collocational.

173 Available from: [<https://epdf.pub/corpus-linguistics-method-theory-and-practice.html>], [Accessed 16 July 2018].

174 McEnery and Hardie (2012) strongly emphasise this viewpoint.

175 See also Dickins (2019, pp.135-62).

On the other hand, based on the definition of SP meaning in this research and the preliminary work on collocation extraction in the previous chapter of this thesis, one can claim that SP meaning in some instances resembles the defining features of *associative meaning* and those of *attitudinal meaning* in the classification of connotative meaning in Dickins (2019). Connotative meaning, as noted by Dickins (2014, p.2), “is all kinds of meaning which are not denotative meaning”;¹⁷⁶ that is, “meaning minus denotative meaning”. According to Dickins (2019), there are, perhaps, an endless number of types of connotative meanings, of which fifteen are identified.¹⁷⁷ Hervey and Higgins define associative meaning as “that part of the overall meaning of an expression which consists of expectations that are – rightly or wrongly – associated with the referent¹⁷⁸ of the expression” (Hervey and Higgins, 2002, pp.149–50 as cited in Dickins 2014, p.3).¹⁷⁹ Dickins et al. (2017) give an example of this meaning in the following:

Most people automatically associate ‘nurse’ with the idea of female gender, as if ‘nurse’ were synonymous with ‘female who looks after the sick’ – on the basis that in the real world (at least in Britain and other English-speaking countries at the start of the twenty-first century) nurses are typically female. This unconscious association is so widespread that the term ‘male nurse’ has had to be coined to counteract its effect: ‘he is a nurse’ still sounds semantically odd, even today. (p.97)

Similarly, evaluative prosody is viewed as a particular connotative meaning which a node acquires through habitual association with particular words [not by intuition] (as in Louw 1993, p.158; 2000, p.50). This habitual association becomes “unconscious” in lexical items which co-occur abundantly in a language; that is, it is considered a widespread meaning. For example, a word like ‘death’ in the Qur’an can be wrongly assumed negative by “unconscious association” of its wide spread meaning of ‘grief and despair’; and that would be called its associative meaning. However, the exploration of the SP of the word ‘death’ as it co-occurs with collocates in the Qur’an overshadows this associative meaning and

176 In the distinction between *denotative* and *connotative meanings*, Dickins claims that denotative meaning is also known by other terms, e.g. denotational meaning, denotation, propositional meaning and cognitive meaning (Cruse 1986, pp. 271–77), while connotative meaning, or connotation, is defined negatively as all kinds of meaning which are not denotative meaning.

177 The identified types of *connotative meaning* in Dickins (2019) are the following :1. Associative meaning; 2. Attitudinal meaning; 3. Affective meaning; 4. Allusive meaning; 5. Reflected meaning; 6. Selectional restriction-related meaning; 7. Collocative meaning; 8. Geographical dialect-related meaning; 9. Temporal dialect-related meaning; 10. Sociolect-related meaning; 11. Social register-related meaning; 12. Emphasis (emphatic meaning); 13. Thematic meaning (theme-rheme meaning); 14. Grounding meaning; and 15. Locution-overriding illocutionary meaning.

178 Dickins (2019, p.136) defines a referent as: “what an expression in a particular sense refers to in a particular ‘speech/writing event’”.

179 Dickins (2014, pp.1-38) classifies associative meaning into three main forms as follows: (i) extra linguistic based, (ii) linguistic-based, and (iii) scalar implicature-based.

endows this word with a prosodic meaning which imposes a new associative meaning in the Qur'anic context relating to positive mentions of the command God has over nature; causing life and death being one of His qualities in the creation of nature and the universe.

Furthermore, Dickins et al. (2017) define attitudinal meaning as:

that part of the overall meaning of an expression which consists of some widespread *attitude to the referent*. The expression does not merely denote the referent in a neutral way but also hints at some attitude to it (p.95).

Likewise, the discourse prosody of a lexical unit in SP, or “the unit’s attitudinal and pragmatic function”, reflects the attitude which the speaker/writer has towards the referent of the word (good because desirable and bad because dangerous) (Sinclair 1998, pp. 14-5; 2004b, p.30-5). By way of example, the word ‘river’ in the Qur’an mostly collocates with words that portray the reward in the afterlife, which in turn reveals the Qur’an’s theme of inviting people to worship God, which is shown in praising those who believe and frequently describing their reward in the afterlife.

To conclude, ‘meaning is use’ is the domain of SP as understood in this research. Words do not have fixed meanings which can be recorded, once and for all, in dictionaries. They acquire or change meaning according to the social and linguistic contexts in which they are used. Similarly, “semantic prosodies are not merely connotational” as “the force behind SPs (semantic prosodies) is more strongly collocational than the schematic aspects of connotation” (Louw, 2000, pp.49-50). This again underlines the relationship between collocation and semantic prosody and confirms the theoretical framework of this research. Furthermore, with this definition which links SP to connotative meaning in mind, SP should be investigated via a methodology which incorporates a statistical model (e.g., Evert 2005; 2008) with a concordance to reflect the different types of meanings evolving from collocation (as in Louw, 1993, p.159; 2000, pp.49-50).

4- Two main approaches to the study of SP

Stewart (2010, p. 161) makes a distinction between what he regards as the two most prominent approaches to the study of SP: Sinclair’s, which he calls ‘discourse prosody’ and Louw’s, which he calls ‘semantic prosody’. He summarises their approaches in terms of the elements, which they prioritise as follows:

Sinclair’s SP approach (1996; 2004a; 2004b):

- central to the unit of meaning;
- one of two obligatory elements;
- considered within a synchronic framework;
- a feature of a unit which is larger than the single word/expression (pragmatic function);
- not restricted to semantically 'neutral' lexical items;
- not restricted to descriptions in terms of 'good' and 'bad.'

Louw's SP approach (1993):

- transferred or attached meaning;
- considered within both a diachronic and synchronic framework;
- a feature of the word;
- associated above all with more semantically 'neutral' lexical items;
- generally expressed by means of a binary distinction whose primary terms are 'good' and 'bad' (positive/negative, favourable /unfavourable).

The SP approach adopted in this research incorporates elements from both approaches. It follows Sinclair's Extended Lexical Unit model in placing SP as a central and obligatory element of the lexical item and deals with the uncovered evaluative prosodies from natural phenomena terms. However, like Louw, it employs a binary division (positive and negative), but with the addition of neutral. Also, it includes the element of discourse prosody (Stubbs, 2001), which endows the text with coherence and has the function of uncovering the underlying meanings relevant to the theme of nature in the Qur'an.

5- The importance of SP

From the studies reviewed in the previous section, the following can be said about SP:

- It is an aspect of evaluative meaning (evaluative prosody).
- It is a habitual association (collocational phenomenon).
- It has a pragmatic function (discourse prosody).
- It is a cohesive device.

Accordingly, SP is a significant phenomenon, and its analysis can be a tool for textual analysis. This same tool can further be employed for comparison between texts, and its importance is described in several studies that discuss the significance of collocation and SP both theoretically and practically. Examples of these studies are provided in the following two sub-sections on the theoretical and practical importance of SP.

a. *Theoretical importance*

Although the different definitions of SP adopt different positions regarding its meaning and significance, at least one of the approaches, namely that of lexico-grammatical patterns, can be used to unify the various definitions. The following table shows the six major theoretical strands of research on the interpretation of collocation together with the scholars who discussed them (See Table 15).

Table 15: Theories of lexis and grammar to interpret collocations

<i>Theory</i>	<i>Scholar/s</i>
<i>The Idiom Principle:</i> “Language text is a result of a very large number of complex choices.” (Sinclair 1991, p.109)	Sinclair (1987b; 1991;1996)
<i>Pattern Grammar</i> “An approach to lexis and grammar based on the concept of phraseology and language patterning arising from work on large corpora. Patterns show how words are typically associated with each other and how they form meaningful units.” (Hunston and Francis 2000, p.43-44).	Hunston and Francis (2000)
<i>Lexical Priming</i> “Every word is primed for use in discourse as a result of the cumulative effects of an individual’s encounters with the word.” (Hoey, 2005, p. 13 and Hoey 2004, p. 386).	Hoey (2004; 2005)
<i>Lexical Bundles</i> “Lexical bundles are defined as ‘recurrent expressions, regardless of their idiomaticity, and regardless of their structural status. To be classified as a lexical bundle, a multi-word unit (MWU) has to (a) frequently occur in a register, e.g., ten times per one million words, and (b) occur in multiple texts in the register.” (Biber et al. .1999 , p.990).	Biber, Conard, Johansson, and Leech (1999)
<i>Collo-structional Analysis</i> It is “an extension of collocational analysis specifically geared to investigating the interaction of lexemes and the grammatical structures associated with them. It is specifically geared to investigating pairs of semantically similar grammatical constructions and lexemes that occur in them.” (Stefanowitsch and Gries 2003).	Gries and Stefanowitsch (2003; 2004)
<i>Construction Grammar</i> Constructions are defined as “conventionalised pairings of form and function” and stored as units in the brain, exist on all levels of grammatical analysis and cover “morphemes and words, idioms, partially lexically filled and fully general phrasal patterns.” (Goldberg, 2006, p. 3 and p. 5). “Construction Grammar claims to be usage-based and places some emphasis on frequency as an indicator for the existence of constructions.” (Goldberg, 2006, p. 228).	Goldberg (2006)

Out of the six theoretical lines of study of collocation given in the table above, lexical priming theory is incorporated into larger theories by some scholars. Morley and Partington (2009, p.144) combine the two views on prosody (namely, prosody as a property and a unit of meaning) by adopting Hoey’s lexical priming. They say that they are two ways of viewing the same phenomenon, the *lexical priming* (essentially as a mental phenomenon)

and the discursal (Hoey 2004, p. 136), and both Sinclair and Partington quite clearly adopt both perspectives. A further explanation of SP in the light of lexical priming is provided below:

The theory of priming also enables us to answer one of the questions which are frequently raised about prosody: If the favourable or unfavourable evaluation of an item said to display semantic prosody is not part of its in-built, inherent meaning – as is the case for words like *excessive* or *timely* – then how do language users decide to employ such items in the appropriate environment? The answer is that language users have a set of mental rules derived from the priming process, alongside or integrated with the mental lexicon, of how items should collocate. These rules are not always open to casual introspection (Morley and Partington, 2009, p 148).

Moreover, it has been argued that it is most appropriate to combine the different views on SP (Morley and Partington, 2009, p.144-45). According to Hunston (2007), there are two lines of thought. The first, she claims, represented in the work of Partington, sees semantic prosody as the ‘property’ of an item which expresses itself in patterns of co-occurrence with other items, whereas the second, championed, she says, by Sinclair, and also favoured by herself, considers semantic prosody as the overall discourse function of a “unit of meaning”. The second view of SP is also adopted by Louw (1993), Stubbs (1996), and Xiao and McEnery (2006). Based on this discourse perspective, Sinclair (2004b) claims that there are two types of SP: *textual* and *statistical*. He writes:

Semantic prosody is a textual phenomenon definable as the evaluative intent of the speaker, that is, the attitude s/he has to his/her topic (good because desirable and bad because dangerous). The statistical, or corpus-assisted definition, analyses via a concordance, how an item is actually instantiated many times in many texts (pp. 30-5).

b. Practical importance

On the practical level, the importance of SP is reflected in the findings of studies which focus on insights gained by listeners or readers into the opinions and beliefs of the text producer through discourse analysis with SP. For example, an awareness of SP can be invaluable for the translator and language learner in distinguishing between items considered to be synonyms or translation equivalents. There is steadily a growing body of work in several languages that use this methodology. The fact that it is a significant

linguistic phenomenon which should be given attention and studied is shown in the following statement about corpus linguistic applications in translation studies.

Semantic prosody is linked with notions such as connotations (Stubbs, 2001), attitudinal meaning (Sinclair, 1987b) or evaluative meaning (Hunston, 2007), and deserves further study for its effects on pragmatic translation, and thus on the translation of specific genres and specialised texts (Kübler and Volanschi, 2012, p.103).

In their study of the importance of exploring SP in the evaluation of pragmatic translation,¹⁸⁰ Kübler and Volanschi (2012) conclude, as seen above, that SP is a linguistic phenomenon that could affect the quality of a pragmatic translation.¹⁸¹ Besides, in regard to the importance of SP for translation studies in general, Stewart (2009) writes:

Semantic prosody must be seen as a reality that translators are required to address; otherwise, important source text elements will be left unaccounted for (p.29).

Finally, a list of studies exploring SP in different languages includes the following:

1. Sinclair (1987b; 1991), who investigated the clause *set in*.
2. Louw (1993), who examined the expressions *symptomatic of, utterly, bent on*.
3. Bublitz (1995), who examined the words *cause, happen, commit, somewhat, and prevail*.
4. Sinclair (1996; 1998), who investigated the units of meaning containing the following core items: *naked eye* and *true feelings*.
5. Stubbs (1995; 2001), who examined the verb *cause*.
6. Hunston and Francis (1999), Hunston and Thompson (1999), Hunston (2002), who highlighted the covert nature of SP and the fact that it is a result of transferred meaning; they studied the verb *take on*.
7. Louw (2000), which concentrated on unfavourable prosodies, showing a change of focus from his 1993 work, with no reference to his 'auras' of meaning, and less

180 The term *pragmatic translation* encompasses what is generally called 'specialized translation', which is usually opposed to literary translation. It has been defined by Newmark (1988) as a process taking into account not only the intent of the author and the expectations of the reader, but also the perlocutionary effect it has on the reader. See also (Kübler and Volanschi, 2012).

181 In *pragmatic translation*, the purpose is essentially to translate a message as efficiently and as accurately as possible. The emphasis is on the content of the message as such rather than on its aesthetic form, grammatical form or the cultural context, all of which are subsidiary to the practical, matter-of-fact goal. Instructions, explanations, directions (such as those given in several languages on dress patterns or packaged goods), scientific treatises, government documents and communiques lend themselves quite naturally to pragmatic translation, but myths and tales, literary works or folklore may be similarly treated. See also (Casagrande, 1954, p.335).

emphasis on the notion that meaning is transferred over time from habitual collocates to the node through ‘contagion’.

8. Tognini-Bonelli (2001), who operated within a Sinclairian framework (Sinclair’s units of meaning, 1996), investigated prosodic differences between supposedly ‘true friends’ in Italian and English.
9. Tao (2003), who studied the SP of various Chinese words meaning *occur*.
10. Partington (1998; 2004b), who found differences in both SP and semantic preference in a set of adverbial intensifiers in English.
11. Xiao and McEnery (2006), who discussed collocation, SP, and near-synonyms in English and Chinese.
12. Dam-Jensen and Zethsen (2007), who found systematic prosodic differences between Danish verbs roughly translatable as *cause, lead to*.
13. Al-Sofi et al. (2014), Al-Ubaidi (2013), and Younis (2018) who explored the SP in the Arabic text of the Qur’an (For details, see Section 3.2.2).

To conclude and summarise the first section of this chapter, previous literature inspired the present research to explore nature as a theme in the Qur’an via examining frequently occurring linguistic patterns of concepts related to this theme. This research mainly sheds light on the concept of connotative meaning to explore the semantic prosody of natural phenomena in the Qur’an. These repeated patterns of co-occurrence are the evidence from which both pragmatic and semantic meanings¹⁸² can then be inferred. The theoretical principles for this research can be summarised in the following points:

- 1- Semantic prosody is a result of collocation.
- 2- Lexico-grammatical patterns constitute collocations with a combination of lexis and grammar to represent meaning.
- 3- ‘Meaning is use’ and ‘you know a word by the company it keeps’ are two fundamental perspectives on which finding collocations and the exploration of the semantic prosody of nature in the Qur’an are dependent.

182 In the semantics-pragmatics distinction, *semantics* is viewed as the study of linguistic meaning; the study of the relation between linguistic expressions and their meanings. In contrast, *pragmatics* is the study of context; a study of the way context can influence our understanding of linguistic utterances (Szabo, 2006, p.364).

- 4- Semantic prosodies are the connotative meanings (associative or evaluative meanings and attitudinal meanings) of the node, and semantic preference refers to the semantic sets of collocates of the node.
- 5- Semantic prosody plays an important functional role, which is establishing the pragmatic meanings relevant to the messages and themes in the Qur'an.

3.2 Corpus linguistic studies on the Qur'an

After discussing the studies which helped to determine the theoretical foundations of this research, it is crucial to examine relevant corpus-based studies on the Qur'an as they reflect two areas of the methodology in this research: the theme of nature in the Qur'an and the SP of natural phenomena. This section reviews the corpus-based studies on the language of the Qur'an. It consists of two subsections: one is concerned with corpus-based studies on nature as a Qur'anic theme (e.g. Jain, 1991; Farooqi, 1992; Al-Yahya et al., 2010; Djamil, 2012; Kahrizi et al., 2012; Khan et al., 2013); the other with SP in the Qur'an as a collocational phenomenon and a tool for accuracy (Al-Nasser and Khashan, 2008; Al-Ubaidi, 2013; Al-Sofi et al., 2014; Younis, 2018). These studies vary in their choice of methodology; some rely only on a corpus-based method, while others, just as the present research, add a computational approach to the exploration of both subject matters in the language of the Qur'an.

3.2.1 Corpus linguistic studies on nature in the Qur'an

The review of studies in this section revisits the available ontologies about the content of the Qur'an outlined in Section 2.2. However, unlike the studies discussed in Section 2.2, which mention studies on concepts occurring in the Qur'an in general, this section sheds light on the studies on the semantic representation of nature in particular. For Example, a review of the previous literature has shown that there are studies that focus on the classification of plants in the Qur'an (e.g., Jain, 1991; Farooqi, 1992; Kahrizi et al., 2012), as well as a brief study about the types of rain referenced in the Qur'an (Jaliliyan and Haji-Khani, 2017). Besides, Al-Yahya (2010) focuses on nouns and verbs related to the concept of Time and finds that there are 7 abstract concepts and 11 concrete conceptions of Time in the Qur'an. In a more advanced exploration of nature in the Qur'an using computational methods, Sadi et al. (2016) model nature-related concepts in the Holy Qur'an using OWL

Ontology of Web Language)¹⁸³ / RDF (Resource Description Framework.¹⁸⁴ Their methodology involves identifying nature-related concepts mentioned in the Qur'an and identifying relations among those concepts. These concepts and relations are represented as classes/instances and properties of an OWL ontology. They add all the nature concepts under OWL/Thing concept. They also use their ontology with SPARQL queries¹⁸⁵ to retrieve verses and concepts of interest. However, this study was conducted on one translation of the Qur'an only, namely 'Saheeh International'. Figure 20 shows the class hierarchy according to their ontology.

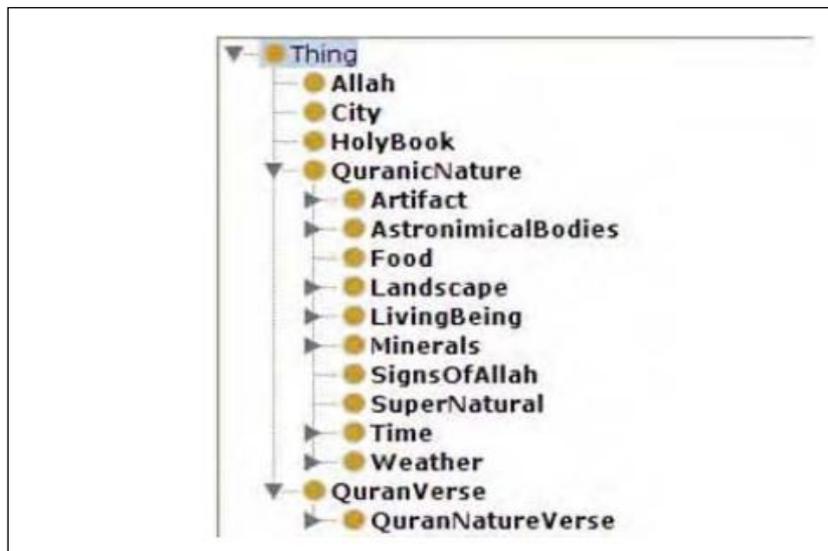


Figure 20: Class hierarchy in the Qur'anic nature ontology (figure taken from Sadi et al., 2016)

To relate Qur'anic nature concepts to Qur'anic verses, they use inverse relationships (Object Properties) called has-part and is-part-of. Their study focuses on only one chapter of the Qur'an, Surah Al-Umran (the second chapter of the Qur'an). An example they mention is verse number 50, which says: "And (recall) when we parted the sea for you and saved you and drowned the people of Pharaoh while you were looking on." In their exploration of nature in the Qur'an here, they divide concepts and relationships, making a triple of concepts, as we see in Figure 21.

183 The *OWL Web Ontology Language* is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics. See also (McGuinness and Van Harmelen, 2004).

184 *RDF* is a directed, labelled graph data format for representing information in the Web See also (McGuinness and Van Harmelen, 2004).

185 *SPARQL* (pronounced 'sparkle', a recursive acronym for SPARQL Protocol and RDF Query Language), recognized as one of the key technologies of the semantic web, is an RDF query language—that is, a semantic query language for databases—able to retrieve and manipulate data stored in Resource Description Framework (RDF) format. Available from: [<https://en.wikipedia.org/wiki/SPARQL>], [Accessed 05 May 2019].

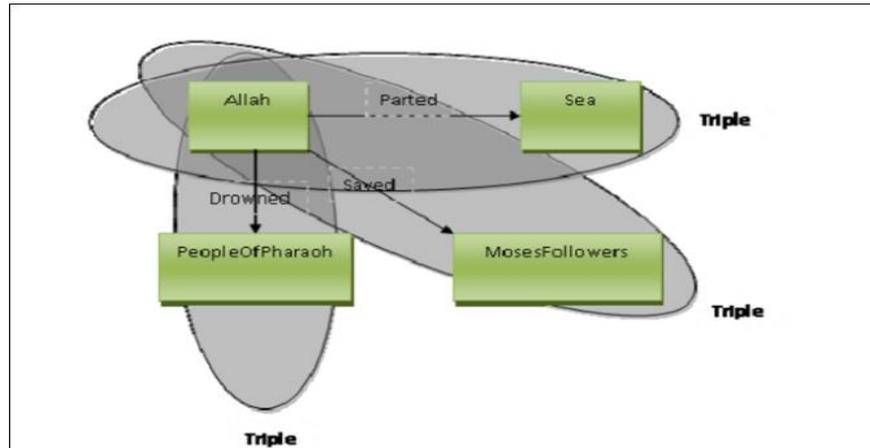


Figure 21: Triple in verse (2:50) (figure taken from Sadi et al., 2016)

Employing a smaller categorisation of natural phenomena and using an English translation of the Qur'an and the Protégé ontology editor tool, Khan et al. (2013) develop a simple ontology based on living creatures including *animals* and *birds* that are mentioned in the Qur'an to provide a Qur'anic semantic search. The authors also propose a framework for a semantic search in the Qur'an using their domain ontology and, similar to Sadi et al. (2016), they evaluate their ontology using SPARQL query language. Their findings show that there are 167 direct/indirect references to *animals* in the Qur'an. In addition, Djamil (2012) explores "Oceanic verses", a term coined by the author, referring to selected ayahs or verses in the Qur'an that contain a word including بحر *baḥr* meaning *ocean* or *sea*, such as بحر *baḥrī* 'the word *sea* in the accusative form', بحرُ *baḥru* 'the word *sea* in the nominative form', بحار *biḥar* 'seas' (plural form), بحرين *baḥrayn* 'two seas in genitive and accusative form', بحران *baḥran* 'two seas in nominative form', and بحرين *baḥrayni* 'two seas in genitive and accusative form with kasra¹⁸⁶'. With a focus on semantics, Djamil develops an ontology, which he calls a categorisation of the aspects of oceanic verses, as revealed in 42 verses in the Qur'an.

3.2.2 Corpus linguistic studies on SP in the Qur'an

To the researcher's knowledge, very little research in the field of corpus-based Qur'anic studies has focused on SP in the Qur'an (Al-Nasser and Khashan, 2008; Al-Ubaidi, 2013; Al-Sofi et al., 2014; and Younis, 2018). It appears that only a few studies were conducted on some lexical items and presented in short articles. For example, Al-Nasser and Khashan

¹⁸⁶ *Kasra* is a diacritic in Arabic, which is a diagonal stroke written below the consonant which precedes it in pronunciation. It represents a short vowel *i* (like the "i" in English "pit"). See also: [https://web.uvic.ca/hrd/hist455/vowels/vowels_pres.htm]

(2008) conduct an analysis of the word مُبِين *mubeen* ‘clear’ in the Qur’an for its SP. In their study of the different corpus linguistic features of this word, they find that it is semantically prosodic with lexical items of negative evaluation. Therefore, they claim that the most frequent of collocations of *mubeen* carry a negative connotation; hence, *mubeen* has a negative SP (See Figure 22).

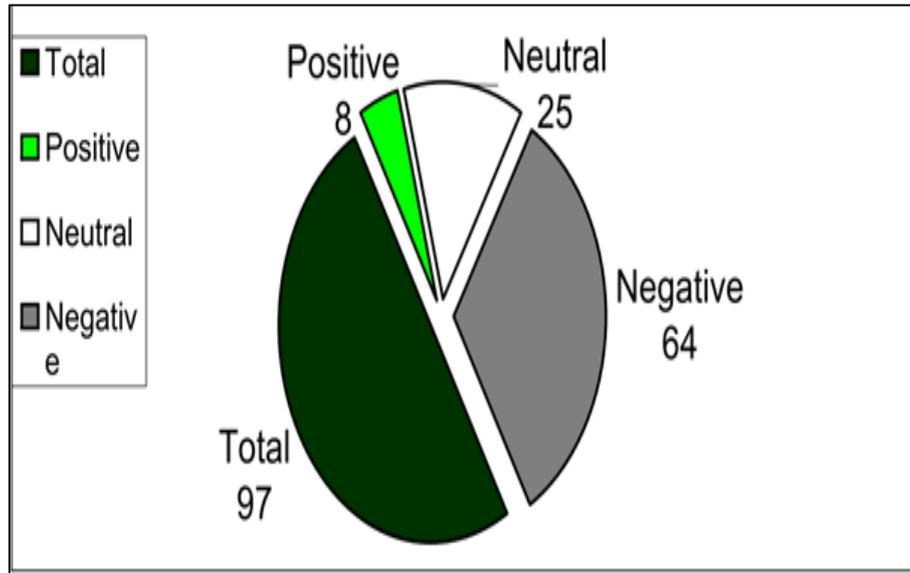


Figure 22: The SP of the word مُبِين *mubeen* ‘clear’ in the Qur’an (figure taken from Al-Nasser and Khashan 2008, p.11)

As seen in the figure above, their findings show that the word *mubeen* occurs 97 times in the Qur’an. Out of these occurrences, it occurs 64 times with connotatively negative items and only eight times with a positive connotation. They report that the negative items are words like عَدُوٌّ *adūw* ‘adversary’, سِحْرٌ *siḥrun* ‘magic’, ضَلَالٌ *ḍalāl* ‘error’, تُغْيَانٌ *tuḡbān* ‘snake’, نَذِيرٌ *naḍīr* ‘warner’, etc. The positive collocates, on the other hand, include words such as رَسُولٌ *rasūl* ‘prophet’, نُورٌ *nūr* ‘light’, نَصْرٌ *naṣrun* ‘victory’. The remaining 25 word-tokens are of neutral evaluation (connotation): شَيْءٌ *ṣai’in* ‘thing’, كِتَابٌ *Kitābun* ‘book’, سُلْطَانٌ *Sulṭānun* ‘authority’, عَرَبِيٌّ *arabīyun* ‘Arabic’, and شِهَابٌ *Ṣihābun* ‘comet’ (pp.11-2).

In an article written in Arabic, Al-Ubaidi (2013) investigates the SP of the collocations of four verbs, which are مَسَّ *massa* ‘touch’, ذَاقَ *dāqa* ‘taste’, كَشَفَ *kaṣafa* ‘reveal’, and أَنْزَلَ *i’anzalanzala* ‘send down’. He finds that the verb ‘sends down’ occurs 62 times and has a positive SP, and collocates with words like *Qur’an, angels, the book, sustenance, water*, etc. He also finds that the word ‘touch’ occurs 53 times and has a negative SP, because it collocates with words with negative evaluative prosody such as *hell, illness, punishment*, etc. Similarly, he finds the verb ‘taste’, which occurs 37 times, to have a negative SP

because it collocates with words such as *punishment* and *illness* too. Finally, he finds that the verb ‘reveals’ occurs 18 times and has a negative prosody because its collocates are also *punishment* and *illness*.

In addition, in the analysis of these verbs, he uses the SP approach in five consecutive steps, which are as follows:

1. Identifying the verbs to be investigated.
2. Making calculations of instances of word use by counting the number of times they appear in the Arabic Qur’anic corpus (frequency).
3. Creating a list of the collocations of each verb and the number of times they occur with it.
4. The movement from the dictionary meaning of the verb to its meaning in the context it occurs in.
5. Establishing a comparison between the dictionary meanings of the verb with its meaning derived from the contexts it appears in (Al-Ubaidi, 2013, p.93).

Al- Sofi et al. (2014) also present a short paper on the SP of four verbs in the Qur’an which are مَسَّ *massa* ‘touch’, ذَاقَ *dāqa* ‘taste’, كَشَفَ *kašafa* ‘reveal’, and جَاءَ *Gā’a* ‘came’. Their conclusion is summarised as follows:

1. The four verbs under study have negative and positive semantic prosodies, not because they have these attitudes, but rather because they are surrounded by a negative or positive semantic environment (collocates).
2. Context plays an essential role in determining the SP of words.
3. SP is one of the most important semantic issues in the study of text meaning because it uncovers the semantic behaviour of words and the semantic set around these words, which in turn creates harmony and cohesion in the text (p.130).

In a more recent corpus-based study incorporating translations of the Qur’an, Younis (2018) investigates the semantic prosody related to the use of certain prepositions (على *alā* ‘on’; إلى *lā* ‘to’; and لِ *li* ‘for’) in verb-preposition constructions by examining the collocational patterns in which they occur. She examines several verbs, including specific forms of the lemmata¹⁸⁷ كَتَبَ *kataba* ‘inscribe or write’, اسْتَمَعَ *istam’a* ‘listened’, نَزَلَ *nzala* ‘send down’, صَبَرَ *šabar* ‘was patient’, أَوْحَى *awḥa* ‘reveal’, and هَدَى *hada* ‘guide’. By

187 In linguistics, a *lemma* is a word considered as its citation form together with all the inflected forms. For example, the lemma *go* consists of *go* together with *goes*, *going*, *went*, and *gone*. Available from: [https://www.collinsdictionary.com/dictionary/english/lemmata].

analysing the concordance lines, the collocates of each verb are identified using the concordance in ArabiCorpus by Parkinson (2012)¹⁸⁸ to discover the various levels of meaning related to each preposition, and the QAC to compare it with parallel corpora of Qur'an translations. Her study uses the corpus-based approach to translation studies by assuming that the concept of semantic prosody can be a tool that raises the level of accuracy when selecting a translation equivalent. She does so by quantifying frequencies and studying the concordance and collocational behaviour of the source word as well as the target word in a corpus of each language (p.139). She does not limit SP to three possibilities, positive/negative/neutral, only, but follows the broader sense of SP as expounded by Sinclair (2004b). For example, she finds that the SP associated with *'alā* is “something hard or difficult or done with effort, something that denotes or implies commitment or obligation” (p.140). On the other hand, the SP that seems to be associated with *'lā* is “the delivering of something that is usually good”. In addition, the collocates of *li-* suggest an SP of “assigning something”; especially in cases where it contrasts with *alā*, *li-* often indicates something good or positive. An important implication of her study is that the “analysis of semantic prosody can help translators achieve the highest possible degree of accuracy” (p.140).

In conclusion, the overview of the studies here showed that not many corpus-based studies have ventured to study extensively the dimensions of nature as a prominent concept in the Qur'an and compare it with its English translations. Moreover, based on the studies above, except for Younis (2018) which employs concordances, none of these studies has applied a corpus-based technique to explore collocations and SP in the Qur'an, and there is no specific mention of the collocation measures and parameters used in any of these studies including Younis (2018). Finally, it appears that the corpus-based study of SP in the Qur'an is still relatively new and only a few short studies were conducted on it; it seems that only one used Natural Language Processing (NLP) or computational methods of any kind. Hence, this research bridges the gap in corpus-based studies on the Quran by covering both SP as a collocational phenomenon and nature as a major theme in the Qur'an.

¹⁸⁸ *ArabiCorpus* (arabicorpus.byu.edu) is a medium-sized, plain text corpus of nearly 200 million words, created by Dilworth Parkinson and hosted by Brigham Young University. Though divided into five genres (Newspapers, Modern Literature, Nonfiction, Egyptian Colloquial, and Premodern), the core of the corpus is the Newspapers category—and thus Modern Standard Arabic with an emphasis on media Arabic. This category consists of full-year datasets of newspapers from different regions of the Arab world—including Egypt, Morocco, and Syria—as well as pan-Arab newspapers such as *Al-Hayat*. The other categories draw from a variety of sources, including modern literature, classical literature, scientific texts, the Qur'an, chats, and plays. (Parkinson, 2012, p.75).

3.2.3 NLP projects and the Qur'an

In recent years, a copious number of NLP studies focused on the language of the Qur'an to develop machine learning projects, information retrieval software and websites on the semantics, syntax, and morphology of the Qur'an. The characteristic feature across the NLP studies on the Qur'an is their focus on building a platform through which information relevant to the Qur'an can be retrieved. The outcomes of these studies are websites and tools that can be used for searching the content of the Qur'an; they are constantly evolving and are widely used by researchers from different backgrounds (e.g., linguists and computational specialists). Of the numerous tools developed for the analysis of the Qur'an, four prominent websites are mentioned below. They are all useful tools, but the most suitable analysis tool for this study is the *Qur'anic Arabic Corpus*.

*Qur'an Analysis (QA)*¹⁸⁹

The QA by Ouda (2015) is a smart search engine for the Qur'an. It is an open-source with which users can search a word in the Qur'an by "meaning" in Arabic and English. They can also analyse the text and get detailed statistics and find hidden lexical and semantic patterns and relations (e.g., collocations). Its major features, as stated on the website, are the following:

1. Basic Statistics: statistics about the total number of Chapters, Verses, Words, Characters and more;
2. Word Frequencies: list of all words in the Qur'an with their frequencies and weights calculated using the TFIDF algorithm;
3. Word clouds: word clouds for each Chapter in the Qur'an in addition to two other clouds for verse endings and beginnings (that is clouds for first and last words in each verse): the bigger the word size, the more it is mentioned;
4. Full Text: listing of all verses in the Qur'an;
5. Charts: a collection of charts such as 'Chapter/Verse distribution';
6. N-Grams: choosing 'N' to produce N-grams of words from the Qur'an;
7. PoS Patterns: getting verses from the Qur'an which match a specific PoS Pattern;
8. PoS Query: listing verses containing any specific PoS Tag from the Qur'an using Qur'anic Arabic Corpus;
9. Repeated Verses: listing of 'full repeated verses' in the Qur'an;

189 Available from: [<http://qurananalysis.com>]

*Qur'any (meaning 'my Qur'an' in Arabic)*¹⁹⁰

In *Qur'any*, Abbas (2009b) incorporates features to search concepts in the Qur'an. The Qur'an corpus is augmented with an ontology or index of key concepts, taken from a recognised expert source: *Mushaf Al Tajweed* by Habash (2001). This is a tool that allows users to search the Qur'an corpus for abstract concepts via an ontology browser. The Qur'an Search Tool in *Qur'any* includes two modules: a search module for keywords, and a search module for concepts. To implement a search for a concept, Abbas (2009b) embeds an index of concepts retrieved from *Mushaf Al Tajweed*¹⁹¹ compiled by Habash (2001). This resource contains a comprehensive hierarchical index or ontology of nearly 1,200 concepts in the Qur'an. Scholars can use the *Qur'any* ontology browser to identify a concept and find the verses that allude to this concept.

*QurSim: A corpus for evaluation of relatedness in short texts*¹⁹²

Sharaf and Atwell (2012) build this tool for semantic observation, where semantically similar or related verses are linked together. The authors have created a dataset called *QurSim*, which consists of 7,600 pairs of related verses for evaluating the relatedness of short texts.

*Qur'anic Arabic Corpus*¹⁹³

The *Qur'anic Arabic Corpus* (QAC), which is the most comprehensive of the previously mentioned websites on Qur'an analysis, was chosen to be used as the linguistic resource in this research. As previously introduced, it is an annotated linguistic resource, which shows the Arabic grammar, i.e., syntax and morphology, for each word in the Holy Qur'an. It is an online Qur'an that includes an ontology that defines 300 concepts in the Qur'an and captures interrelationships using predicate logic. The number of relationships is 350, and the type of relationships between concepts is 'Is-a'. Dukes (2011) built this Qur'anic ontology of concepts based on the knowledge contained in traditional sources of Qur'anic analysis, including the sayings of the Prophet Muhammad (peace be upon him), and Tafsīr books. This research employs the QAC in three parts of the methodology. It is first used when compiling and refining the list of natural phenomena by using the unified root system

190 Available from: [<http://quranytopics.appspot.com/>], [Accessed 16 December 2016].

191 Compiled by Dr Mohamed Habash (2001), Director of the Islamic Studies Centre in Damascus, published by Dar Al-Maarifah in Syria and authenticated by the Al-Azhar Islamic Research Academy in Egypt.

192 See also (Sharaf and Atwell, 2012).

193 Available from: [<http://corpus.quran.com>].

available on this website. Also, it is employed when producing the ontology of nature, a major Qur'anic theme, as claimed by Fazlur Rahman (2009). Most of the terms belonging to sub-concepts related to the theme of nature in the Qur'an are identified with the aid of this tool and for this research. Furthermore, all the examples from the Qur'an in this thesis are taken from this linguistic resource.

3.3 The evaluation of translation

Having chosen the theoretical framework and methodology in the first phase of this research and decided on the translations included in the corpora of this research, a method for evaluation for the second phase, involving evaluating the translations of the Qur'an for their congruency with the representation of SP of nature in the Arabic Qur'an, is chosen. As previously mentioned, this is a corpus-based study to evaluate the representation of nature in the English translations of the Qur'an via SP as a cohesive device and its analysis as a tool for accuracy (as in Ebeling, 2014; Younis, 2018). Accordingly, it requires a method that incorporates corpus linguistics with applications of translation studies. Kennedy (2014) distinguishes between two types of corpora in studies that compare different languages: "parallel corpora", where one corpus consists of a text and its translations, and "comparable corpora", where each corpus consists of the same kind of texts written originally in each language. The corpora in this research belong to the first category because they include the source text (ST), that is, the Qur'an, and five of its English translations, the Target Texts (TT). Besides, this research adopts Olohan's (2004) definition of *descriptive translation studies* as follows:

Descriptive translation studies (DTS), as the name suggests, is interested in describing translations and translation practice as it occurs or has occurred, the role and nature of translation and the impact of translation activity in wider cultural, social, historical, etc., contexts (p.199).

According to Olohan, DTS is an approach to evaluation that can be traced back to 1970s and was developed in opposition to previous, prescriptive approaches to translation that were interested in formulating for translators on how translations should be. DTS is also a corpus linguistic methodology employing various corpus analysis methods (e.g., the frequencies, collocations, and SP in this research) to examine aspects of a text (as in Mason, 2001; Kenny, 2006; Munday, 2014). In the same sense, this study adopts Munday's (2014) method of evaluation, which will be reviewed below. It is a method, which entails

that several linguistic features of the source text are used to compare and evaluate the translations of the Qur'an. On the application level, TT corpora can be compared to the source text via alignment, which means using software to compare the linguistic features of translated texts (Olohan, 2004, p.26 , p.55; Baker et al., 2006) (see also Section 4.2 for more on the evaluation approach). Regarding the application of this method, Mason (2001, p.77) writes that the “analysis of source texts and their translations provides us with information about translator behaviours, or as he calls it later, “how translators use language”. This section reviews several previous studies on evaluating translations generally, and mainly three studies which assess translations of the Qur'an regarding different linguistic aspects (e.g., Abdullah, 2010; Al-Ghamdi, 2015; Younis, 2018).

In a study of the cohesive device of junction and its translation between French (ST) and English (TT) in corpora consisting of *L'Etranger*¹⁹⁴ and its English translation *The Outsider* by Stuart Gilbert, Mason (2001) concludes that for different genres and discourses, certain types of junction are explicit in Camus' French texts and are treated in different ways in the English translation and vice versa. This means that “translators are responding to a lack of correspondence between ST and TT” (p.61). Mason (2001) finds, for example, that the ellipsis (e.g., suppressed connectives) of the causative junction in French is made explicit in English using ‘and’, as in the following lines:

French text: je lui ai demandé si on pouvait éteindre une des lampes. L'éclat de la lumière sur les murs blancs me fatiguait.

English translation: The glare forms the white walls making my eyes smart, and I asked him if he couldn't turn off one of the lamps (2001, p.72).

Kenny's (2006) *German-English Parallel Corpus of Literary Texts* (GEPCOLT) comprises works of mainly German-English literature and their translations into English. Her research focuses on lexis and creativity (i.e., creative words are those which occur only once in a corpus or those which are invented or used by a single writer) in translation using corpus methodology; she makes several contributions with her study. Not only does she compile a sizable parallel corpus with the specific aim of investigating lexical normalisations in translation (i.e., whether translators use more conventional forms in the target text to render lexically creative source-text forms, e.g., unusual word forms or word combinations), but she successfully develops a methodology for carrying out a detailed study of how lexical

194 *L'Etranger* is a 1942 novel by French author Albert Camus.

creativity is dealt with in translation. Based on the assumption that creative word forms will occur very infrequently, a set of creative word forms is extracted from a list of words that occur only once in the corpus (i.e., hapax legmena).¹⁹⁵ Another focus of the investigation is creative collocations; for this, she selects a node that occurred commonly throughout the corpus and examines its collocates. In doing so, she combines the capabilities offered by corpus software tools with qualitative analysis (p.115). Findings of her analysis of creative collocations reveal a combination of linguistic creativity and normalisations on the part of the translator.

Munday (2014) develops a model that attempts to overcome the shortcomings of other approaches for source-target text (ST-TT) analysis by bringing together systemic functional linguistics, corpus linguistic techniques, and consideration of broader socio-cultural and socio-political contexts. For his study, Munday analyses the source text which is a Spanish article written by Gabriel García Márquez, in which the author describes the story of Elián González, the six-year-old Cuban boy who was rescued from the sea while attempting to reach the USA in 1999. The target texts are translations of this article that were published in the *Guardian* and in the *New York Times* and published by the Cuban group *Granma International*. Munday uses various methods of corpus analysis, sentence length, type-token ratio, frequency lists, “to pinpoint aspects of the text that would seem to merit closer examination” (Munday, 2014, p.80). He subsequently links these to the meta-functional profiles of the texts (ideational, interpersonal and textual, based on Halliday, 1994). In this way, Munday identifies significant shifts that have taken place in the translation process and posits explanations for them. For example, he outlines how the anti-USA sentiment of the source text has been handled differently in the target text and indeed mainly been omitted for the readers of the *New York Times* (pp.87-90). Although his corpus is relatively small compared to this research and other corpus linguistic studies, Munday's study is essential to this review because it shows the potential offered by combining a quantitative analysis with a qualitative one, a mixed approach which this research adopts for the corpus-based study of SP of nature in the Qur'an (see also Section 4.2).

In a study similar to the present research in its comparison of lexical items via SP matching, or what she calls “correspondences” (i.e., congruency), Ebeling (2014) conducts a

¹⁹⁵ *Hapax legmena* is a Greek phrase (singular hapax legomenon usually abbreviated to hapax) meaning 'once said' and is used to describe a word that occurs only once in a text or set of texts (Baker et al., 2006, p.81).

contrastive study of SP in English and Norwegian corpora. In her cross-linguistic study, she presents three case studies of English units with established negative prosody containing the core items: *commit* in Partington (1998), *signs of* in Stubbs (2001), and *utterly* in Louw (1993).¹⁹⁶ The Norwegian correspondences of these items are identified via the *English-Norwegian Parallel Corpus*.¹⁹⁷ These correspondences serve as the starting point for an investigation of cross-linguistic prosodies. She finds that while units with *commit* and *signs of* have good Norwegian matches in terms of semantic prosody, units with *utterly* are less stable across the two languages. She concludes by emphasising the importance of carrying out studies of this kind in order to improve the cross-linguistic understanding of extended units of meaning. This in turn, she says, “has implications for how teachers, translators and lexicographers choose to present words in isolation or as part of larger, extended units” (p.176).

In a non-computational but corpus-based evaluative study, Abdullah (2010) evaluates the translations of collocations in the Qur’an.¹⁹⁸ In his research, six translations of the Qur’an are used to analyse the way translators have rendered the Qur’anic verses that contain collocations. The translations are those of Pickthall (1930); Daryabadi (1957); Ahmad Ali (1984); Irving (1985); Mohamed J. Ahmed and Samira Ahmed (1995); and Qaribullah and Darwish (2001). He follows Williams’ paradigm in the classification of types of collocations. Williams suggests the following types of collocations:

1. Verb + Noun e.g., throw a party / accept responsibility
2. Adjective + Noun e.g., square meal / grim determination
3. Verb + Adjective + Noun e.g., take vigorous exercise / make steady progress
4. Adverb + Verb, e.g., strongly suggest / barely see
5. Adverb + Adjective e.g., utterly amazed / completely useless
6. Adverb + Adjective + Noun, e.g., totally unacceptable behaviour
7. Adjective + Preposition e.g., guilty of / blamed for / happy about
8. Noun + Compound Noun, e.g., pay packet/window frame (Williams, 2002 as cited in Abdullah, 2010, pp. 155-156).

¹⁹⁶ Available from: [<https://journals.uio.no/index.php/osla/article/view/695>], [Accessed 04 May 2018].

¹⁹⁷ *The English-Norwegian Parallel Corpus* (1994-1997), Dept. of British and American Studies, University of Oslo. Compiled by Stig Johansson (project leader), Knut Hofland (project leader), Jarle Ebeling (research assistant), Signe Oksefjell (research assistant). <http://www.hf.uio.no/ilos/forskning/forskningsprosjekter/enpc/index.html> The extended version of the ENPC (ENPC+) has been described in Ebeling and Ebeling (2013).

¹⁹⁸ No mention of a collocation statistic being used in his study.

He limits his analysis of Qur'anic verses to those that contain collocations of the type (Noun + Adjective). In this regard, he focuses on six examples of this pattern of collocations. They are the following: عذاب مهين 'shameful torture'; عذاب أليم 'brutal torture'; عذاب شديد 'torture, torment, which is severe'; عذاب مقيم 'torture that lasts'; عذاب غليظ 'torture which is rough, harsh'; and عذاب عظيم 'great torment'.

Abdullah (2010, p.163) asserts that the translators use the nouns doom, torment, punishment, torture, and nemesis as renderings for the ST Noun عَذَاب *adāb* 'torture'. He highlights that the noun 'doom' only collocates with the adjectives 'approaching' and 'impending'. The noun 'torment' collocates with the adjectives 'great', 'inner', 'mental', and 'private'. The noun 'punishment' collocates with the adjectives 'cruel', 'harsh', 'heavy', 'severe', 'unusual', 'appropriate', 'fitting', 'capital', 'corporal', and 'physical'. The noun 'torture' collocates with the adjectives 'brutal', 'systematic', 'mental', and 'physical'. As for the noun 'nemesis', it was used by Ahmad Ali (1984), and it has no collocations. Moreover, Ahmad Ali's renderings of عذاب أليم 'brutal torture' were inconsistent, unlike the other translators. Accordingly, he claims that all translators failed to appropriately render the ST collocations of عذاب in the Qur'an (p.176).

Similar to the present research in the compilation of corpora and evaluation criteria is Al-Ghamdi (2015) who analyses and critically evaluates the English renderings of the Divine Names in five English translations of the Qur'an, namely Pickthall, A. Y. Ali, Arberry, Hilali Khan and Abdel-Haleem. He suggests that all five translations not only fall short in their attempts to distinguish between the near-synonymous Names, particularly the root-sharing ones, but prove unsuccessful in rendering them accurately and consistently. He also finds that the translation of Arberry shows much better quality in terms of accuracy and consistency, in rendering the root-sharing Divine Names, compared to the other translations. In addition to critically revealing shortcomings, inaccuracies, and inconsistencies of the renderings of the Divine Names, his study suggests the use of translation technology solutions (or computer-assisted tools), such as translation memory and bilingual concordances, to improve the quality and consistency of future Qur'an translations in general, and the renderings of the recurring Divine Names in particular.

Also, most closely related to the present research in terms of methodology is Younis' (2018) study who uses a corpus-based methodology to explore the SP of certain prepositions (على *'alā*, 'on'; إلى *'lā*, 'to'; and لِ *li-*, 'for') in verb-preposition constructions and evaluate six translations of the Qur'an (Saheeh International; Pickthall; Yusuf Ali;

Shakir; Mohsin Khan; and Ghali). Using a parallel Qur’anic corpus, the six translations are examined and compared to identify similarities and differences in the rendering of the source verb-preposition construction into English (p.126). In the results of her study, Younis underlines that these corpus-based analyses will, as she claims, “hopefully” help translators render the Arabic text into the target language while keeping the same semantic effect conveyed by an original preposition. She demonstrates that a careful reading of concordance lines of both the source lexical item and the equivalent used by the translator/s in the target language (alignment), alongside, as she says, a comparison of the collocational behaviour of both, can easily be a tool for selecting the translation with the closest meaning. One of the findings of her study is that translators, when rendering the verb-preposition constructions into English in most cases, tend to neglect the meaning of the preposition in favour of the verb. The results of her evaluation of the translation are summarised in Table 16:

Table 16: Summary of preferred translations for verb-preposition constructions under study (Younis, 2018, p. 140)

<i>Verb-preposition combination</i>	<i>Best English equivalent</i>	<i>Translation/s using the best equivalent</i>
<i>kataba+’alā</i>	has decreed upon	Saheeh International
<i>kataba+li-</i>	has decreed for	Saheeh International; Pickthall; Yusuf Ali
<i>yastami’+ilā</i>	give ear to	Pickthall
<i>yastami’+li-</i>	hearken unto	Pickthall
<i>’anzala+’alā</i>	has been sent down on	Ghali
<i>’anzala+ ilā</i>	has been sent down to	Mohsin Khan; Ghali
<i>’iṣbir+’li-</i>	wait with patience/ patiently for	Mohsin Khan; Yusuf Ali; Shakir
<i>’iṣbir+’alā</i>	bear/endure patiently/ with patience	Pickthall; Shakir; Ghali
<i>’awḥa+ ilā</i>	revealed to	Saheeh International
<i>’awḥ a+ li-</i>	has commanded	Saheeh International
<i>yahdi+ ilā</i>	gives guidance towards	Yusuf Ali; Ghali
<i>yahdi+ li-</i>	guides to	Saheeh International; Mohsin Khan; Shakir; Ghali

In short, the present research is similar to Mason (2001), Kenny (2006), Ebeling (2014), and Munday (2014) in following the model of comparing texts based on linguistic features of a ST, and to Al-Ghamdi in two aspects, that is, the choice of the number of translations in the corpora and the criteria of evaluation. Furthermore, although it follows a different approach in applying the evaluation to that of Al-Ghamdi (2015), it is similar to Younis (2018) in employing the analysis of SP as a device for accuracy and lexical cohesion. By comparing the results of SP in the ST and TT, the evaluation is established. Finally, it

resembles aspects of Kenny's approach in its compilation of an extracted list of creative words because this research compiles a list of words describing natural phenomena in the Qur'an.

3.4 Conclusion

The studies described in this chapter were useful in constructing the theoretical framework of this research and in guiding the decisions on the choice of methodology. Also, they showed the gap that this research aims to address, which is in the corpus-based studies on SP whether it is as a linguistic phenomenon, a cohesive device, and a tool for word choice accuracy for translators. Furthermore, the computational methodologies employed in previous studies on the Qur'an provided insights into quantitative and qualitative analyses that can then be employed in the mixed approach applied in this research. To conclude, this chapter links the previous studies to the present research and has paved the way to design a theoretical framework and develop the methodology, which blends a corpus-based study with a computational approach and with an evaluative analysis of translations.

Chapter 4 Methodology

The Corpus-Based Exploration of SP of Nature in the Qur'an and its Translations

Introduction

After presenting an outline of the previous literature that aided the choice of the theoretical framework and methodology of this thesis, it should be stressed that upon choosing a methodology to explore SP of natural phenomena in the Qur'an, some common features of the previous literature of research on SP is adopted both on the theoretical and methodological level. This platform of the adopted theory defines SP as a subtle meaning derived from co-occurrence in language, not intuitively discovered, but detectable through analysis of concordance. In addition, in the analysis of the concordance lines of a particular node, collocations are observed in the first lines of concordance (as in Sinclair, 2003) to make judgments on its sets of meanings or SPs. However, the novelty of the methodology in this thesis lies in the fact that it links the analysis of concordance to find SP with a statistical analysis based on Evert (2008) for collocation discovery via NLP methods, a combination of methods that does not seem to appear even in the latest studies on SP in general (e.g., Zhang, 2013; Ebeling, 2014) and SP in the Qur'an in particular (e.g., Al-Ubaidi, 2013; Al-Sofi et al., 2014; Younis, 2018). From this combination emerges the mixed approach of quantitative and qualitative analyses. It is quantitative in finding and verifying the collocations, and qualitative in interpreting the statistics of collocations to identify the different sets of meanings (SPs) relevant to the theme of nature in the Qur'an. Finally, a further quantitative but not computational analysis is conducted regarding the calculation of SP (evaluative and discourse prosodies) for each of the natural phenomena terms. This methodology is useful in producing datasets with features that can be used to evaluate the translations of the Qur'an for their congruency with the representation of SP of natural phenomena in the Qur'an.

This chapter provides an overview of the methodology of this research which consists of two phases: phase one, which is the central corpus-based part of this research, aims to explore the semantic prosody (SP) of nature as a theme in the Qur'an and its translations; phase two is concerned with the evaluation of the translations for their representation of this theme in comparison with the Arabic source text. As a corpus-based study, this research

examines corpus linguistic features such as frequencies, collocations, and SP (See Figure 24). In this regard, it adopts Sinclair's (1991; 2003; 2004b) concepts and methods for working with corpus data and analysing concordances.

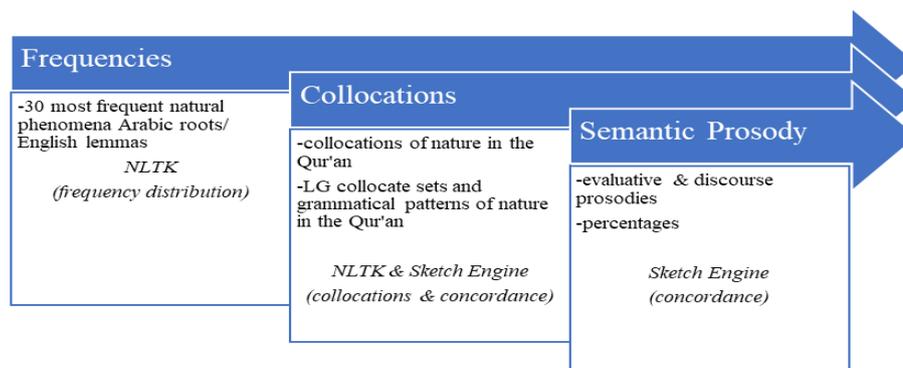


Figure 24: The three corpus linguistic features of the methodology

On the other hand, as an application of a corpus linguistic methodology on an evaluative translation study in the second phase, it follows the comparative model underlying the use of parallel corpora, i.e., corpora consisting of source texts and their translations, and various corpus analysis methods (e.g., the frequencies, collocations, and SP from the previous stage), which are used to examine aspects of the texts (as in Chesterman, 2000; Mason, 2001; Kenny, 2006; Munday, 2014). It is a descriptive¹⁹⁹ study of translation that combines quantitative and qualitative corpus-based analyses, with a focus on (a combination of) lexis, syntax, and discourse features.²⁰⁰ The second part is an application of the findings of the first phase. It presents evidence of the usefulness of corpus-linguistic SP analysis as a means of uncovering aspects of meaning that are not covered copiously by traditional analyses and methods in translation studies to describe and evaluate translations.

This chapter will describe the details of these two phases, including an outline of the methods applied to analyse the Arabic text and English translations of the Qur'an. The findings in the resulting datasets and their discussion are presented in chapter Five. The current chapter is organised as follows: Part 1 describes the corpus-based analysis of SP of natural phenomena in the Qur'an and its five translations. It outlines the steps applied to explore collocations, namely semantic prosody, as a collocational phenomenon in the Qur'an, as seen in the three features that were illustrated previously in Figure 24. Also, Part

¹⁹⁹ For the definition of DTS see Section 3.6.

²⁰⁰ Discourse in translation studies most often is used to refer to attitudinal expression through language, i.e. the "material out of which interaction [between writer and reader] is moulded as well as the themes addressed" (Hatim 1998, p. 98).

2 presents the evaluation of the Qur'an translations via the analysis of SP of nature terms in the Qur'an. It discusses the approaches to the evaluation of data to decide on the most congruent translation of the Arabic representation of nature in the Qur'an. Finally, Section 4.3 gives a conclusion to this chapter.

4.1 Part 1: The corpus-based analysis of SP of natural phenomena

This section provides an overview of the corpus-based methods applied to analyse the Arabic text and the English translations to find their evaluative and discourse prosodies. Figure 24 shows the processes this analysis follows to explore the SP of nature as a Qur'anic theme in the corpora of this research. It illustrates that following the production of an extensive list of natural phenomena in the pre-methodology stage as described in Chapter Two, the general framework of the methodology described in this chapter follows four steps that encompass corpus-linguistic and Natural Language Processing (NLP) computational applications (See the right box in Figure 25). It should be noted that in this part, there will be a reference to the workflow of tasks from the preliminary stage of this research (Chapter Two). These references relate primarily to the use of NLTK in Python in the experiments that were conducted before the actual study; they represent the researcher's attempt to refine a list of natural phenomena in the Qur'an and present a pilot statistical profiling of the Qur'an and these terms via this NLP tool.

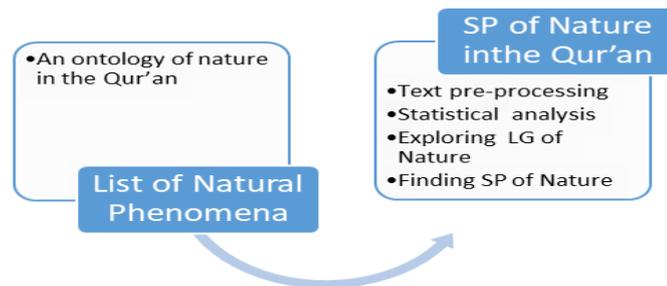


Figure 25: Finding the SP of natural phenomena in the Qur'an

Furthermore, the corpus-based analysis in this section is based on several corpus and computational linguistic methods from previous work on semantic prosody as a collocational phenomenon (Firth, 1957; Stubbs, 1995, Partington, 1998; 2004b; Manning and Schütze, 1999; Sinclair, 2004b; Evert, 2005; 2008; Morley and Partington, 2009; Abdullah, 2010; and Younis, 2018). It is a method that addresses the first two questions of this research:

- 1- Is there SP in the representation of nature in the Qur'an?
- 2- What are the lexico-grammatical patterns, evaluative prosodies, and discourse prosodies of nature in the Qur'an?

To explore the SP of nature in the Qur'an and its translations, it was essential to follow various tasks, the first of which was to clean the texts using computational tools via Python (e.g., removal of stop -words, removal of punctuation, word disambiguation, etc.) in what is referred to in the above mentioned as text pre-processing to speed up the programme and later provide accurate frequencies and collocation extractions of nature terms (See Figure 25). After that, automated statistical analysis using NLTK in Python is undertaken on the corpus by computing the raw frequencies of each of these terms and collocation extraction, namely bigrams (word pairs) for each of the terms, since "collocations are essentially just frequent bigrams" (Bird et al., 2009, p.20). Once the evaluative prosodies and discourse prosodies were manually labelled for each of the occurrences of these terms based on the most frequent collocations, the percentages of their evaluative prosodies (i.e., *positive*, *negative*, and *neutral*) and discourse prosodies (e.g., *reward*, *glorifying of God*, *afterlife punishment*, etc.) were tagged in Excel to label their contexts (e.g., verses in which they occur). These were calculated to find the percentage for each number of positive, negative, or neutral collocations against the number of its frequencies to allocate evaluative prosody and discourse prosody for each of the natural phenomena which repeatedly occur with collocates in the Qur'an. The same list of tasks was applied to produce five other annotated datasets representing the occurrences of nature in the selected translations of the Qur'an. An elaboration on each of the steps to explore the SP of nature in the Qur'an is provided below. However, before elaborating on the steps of this phase of the methodology, the choice of this methodology is justified in the following section.

4.1.1 Finding SP: a corpus-based method and a computational approach

This section presents the rationale behind the choice of the methodology and quantitative analysis in this research. Being corpus-based research that aims to infer SP from concordances, it follows the link between semantic prosody and computational analysis of the corpus data (as claimed by Hunston, 2001; Adolphs and Carter, 2002; Baker et al., 2006; Baker, 2016). This claim can be seen in Hunston (2001, p.142) who says that SP can be "observed only by looking at a large number of instances of a word or phrase, because it relies on the typical use of a word or phrase", and in Adolphs and Carter (2002, p.7) who

state that, “the study of semantic prosody has only become possible with the advent of large corpora and suitable software”. Similarly, Baker et al. (2006) put forward the notion that semantic prosody is about “words in a corpus” and can be observed/revealed by computational methods (as cited in Stewart, 2010, p. 81). Similarly, this research holds the viewpoint that semantic prosody can be inferred by “elicitation and introspection” (Baker et al., 2006 as cited in Stewart, 2010, p.82). According to Stewart, finding SP from corpus data involves two macro stages: (a) choosing a relevant and appropriate search (natural phenomena in the Qur’an), and (b) scanning the corpus data retrieved to identify collocations and semantic preferences and then “translating” these into a prosody to transform data into evidence (pp. 83-5). However, the second stage, as Stewart sees it, can be highly interpretive and involves subjective judgments about the text (p.85). To reduce subjectivity, this research consulted exegetical works, which include that of Tafsīr al-Jalalayn and Asbab Al-Nuzul. These provide the original interpretation of the Qur’an as has been explained by Prophet Muhammad (peace be upon him) and his companions and as reported by prominent Qur’an scholars.

The following sections discuss the corpus-based methods and steps of the computational approach to the analysis of SP in the Qur’an and its translations in the first phase of this research.

4.1.2 Text pre-processing²⁰¹

Text pre-processing via NLTK in Python, which involves “the transformations applied to the data before feeding it to the algorithm” Mayo (2018, www.kdnuggets.com), was the first step of exploring the SP of nature in the Qur’an (See Figure 26).

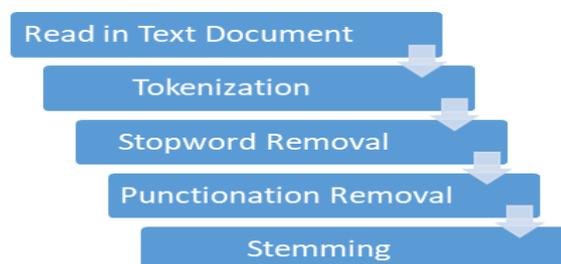


Figure 26: A model for text pre-processing the research corpora²⁰²

201 See also Chapter Two for introductory tasks of experimenting with NLP to explore the list of natural phenomena that were practically similar to this part of the methodology.

202 Adopted from Manning et al. (2008).

This section of the methodology recalls the pre-processing model employed in the pre-methodology stage of this research; Mayo's (2018) model for pre-processing, which is also based on Manning et al. (2008, pp. 47-56), resembles that of Pooja et al. (2016). The elaboration of the pre-processing tasks of the Yusuf Ali translation was therefore repeated in this part of the chapter, but with the addition of the details of pre-processing that are exclusive to the Arabic text. In pre-processing (as in Figure 26), the first step was to read each of the texts via Python as seen in the code below:

```
import nltk
raw= open ('en.pickthall (4).txt', 'r'). read ( )
```

The second step of pre-processing texts was *tokenisation*. It is defined as “the task of cutting a string into identifiable linguistic units that constitute a piece of language data”(Bird et al., 2009, p.109). In this process, the raw texts were split into smaller units (e.g., sentences then words), and the result is a list of words or tokens (i.e. tokens in this research refer to the smallest units that each corpus divides into; typically, each word form and punctuation is a separate token). These tokens provide a platform for counting raw frequencies and ascertaining unusual or significant frequency in the following task of exploring the SP of nature in the Qur'an. Furthermore, the upper-case in the English texts was removed as a form of normalisation or filtering (See the following code).

```
import string
from nltk.tokenize import sent_tokenize, word_tokenize
#tokenise and remove all upper-case in the English texts only
tokens =nltk.word_tokenize (raw1)
raw2= [w.lower( ) for w in tokens]
```

After that, the *normalisation* or *filtering* phase is carried out to eliminate punctuation and stop-words (e.g., *the*, *at*, *which*, and *on*). The removal of stop-words is done to focus on content words, which were shown to reveal the covert meanings of nature in the Qur'an as seen in the exploratory work that is reported in Chapter Two of this thesis. This study is interested in looking at collocations [i.e., lexico-grammatical patterns], which are ranked based on significance. To do this, the significance test of Log-Likelihood was used. This test highlights high frequency words and is bound to generate collocations with an abundant presence of stop-words [i.e., function words]. To go around this and focus on content words, it was plausible to eliminate stop-words.²⁰³ However, it should be stressed that this research does not neglect the role of function words entirely; that is, the allocation of

203 See also: <https://www.lancaster.ac.uk/users/moocs/corpus/people/hardie-wk2/index.htm>

semantic prosody was not only determined via LG patterns with zero-function words. It was elicited through the examination of concordance lines aligned with the raw corpora, where the researcher observes the presence function words (e.g., negation particles, which can occasionally change positive evaluative prosody into a negative one).

In this regard, Arabic, being a language with rich morphology, has stop-words with a variety of forms that need to be removed. Based on a grammatical distribution of the most frequent words, a list of stop-words was compiled and checked against the list recently uploaded and provided by the NLTK update in Python. The 124 words that are not present in the former list were also appended via Python. Firstly, a classification for grammatical features of stop-words in the Qur'an was chosen as follows:

- Demonstrative pronouns
- Interrogative pronouns
- Prepositions
- Relative pronouns
- Expressions of exception
- Personal pronouns
- Conjunctions
- Particles of the accusative case
- Conditional pronouns
- Vocative case of pronouns
- One of the Five Nouns (ذو dw 'possessor of')
- Auxiliaries (verbs 'to be' and its 'sisters')
- Words of time and place
- Other words (as in Abu-Chacra, 2007, p.61)

Secondly, using NLTK in Python (as illustrated in the code shown below), the existing list of Arabic stop-words in the NLTK was obtained.

```
from nltk.corpus import stopwords
```

Another list by Zerrouki and Amara (2009)²⁰⁴ as shown in the table below, was also

204 Available from: [<http://arabicstopwords.sf.net>], [Accessed March 2018]

obtained. The number of distinctive, duplicate, missing, and compiled Qur'an stop-words in each of the lists is also shown.

Table 17: Compiling the Qur'an stop-list

<i>Zerrouki and Amara (2009)</i>	<i>Python.org</i>	<i>Duplicate</i>	<i>Missing</i>	<i>Compiled Qur'an Stop-words</i>
78	71	148	46	343

Stop-words in the Python list were checked against two lists: the list from Zerrouki and Amara, as well as a compiled list of missing words based on the above categorisations (particularly Parekh, 2000).²⁰⁵ The new list (343 words) contains the existing duplicate words already in both lists (148 words); distinctive words specific to Zerrouki and Amara but not in the Python list (78 words); distinctive words specific to the Python list (71 words); and the compiled missing words (46 words). The missing words are mainly the ones that are not present in the existing three lists: inflectional forms of the particles; repeated particles with diacritics that could affect the processing of the text if it has diacritics; or repeated particles with the conjunction *و* 'the Arabic equivalent of the conjunction *and*', which is not separate from the words in the Qur'anic text. The following figure illustrates this task as the Python list was examined and modified. The words shaded in pink are the ones that are missing; the ones shaded in blue are duplicates; the word in the yellow shade is an example of an exclusive item in the list retrieved from python.org.

A	B	C	D
Python.org [https://pypi.python.org/pypi/stop-words/2014.5.26, Mar, 2018]	Zerrouki & Amara	Parekh [www.understantquran.com, Mar, 2018]	Quran Stopwords
إذ	و	عندى	إذ
إذا	سوى	إلى	إذا
إذاما	غير	ولك	إذاما
إذن	لأسباب	ولكم	إذن
أف	من	ولى	أف
أهل	للى	ولا	أهل
أكثر	للى	غله	أكثر
إلا	للكان	عظكم	إلا
إلى	أمن	عظك	إلى
اللى	لكم	للى	اللى
اللى	يا	عظا	اللى
اللى	هنا	عنه	اللى
اللى	من	عظم	اللى
اللى	كم	عظك	اللى
اللى	ك	عظك	اللى
اللى	ك	عظك	اللى

Figure 27: The modification of the available Python Arabic stop-words

205 Parekh (2000) presents a categorisation of the grammar of the Qur'an based on frequency to enhance the understanding of the language of the Qur'an for non-Arab readers of the Qur'an.

After the removal of stop-words from the texts in this research and for a similar purpose, punctuation was removed both from the Arabic Qur'an as well as the five translations to increase the speed of the programme in discovering the frequencies and collocations of natural phenomena in the text-processing step.

```
#remove Punctuation
punct= set (string.punctuation)
raw1= ''.join (w for w in raw if w not in punct)
```

The final step in text pre-processing was *stemming*, whereby morphological tools were used to reduce multiple forms of the word to one form in a process referred to as *stem-based disambiguation*.²⁰⁷ A stemmer in the context of this research is an automatic process in which morphological variants of terms are mapped onto a single representative string called a *stem*. Stemming was especially important to achieve results which reflect the aggregated frequency and extract collocations accordingly; it was applied to the Arabic and English texts (See also Section 2.3 on the justification for using stemming, especially on Arabic). Hence, to explore the presence of SP in the Arabic Qur'anic text and English translations, tokens needed to be reduced to their least minimally inflectional and derivational forms in preparation for their use in processing the text. The lines of code below illustrate the implementation of the Porter Stemmer, a tool developed by Martin Porter at the University of Cambridge in 1980 used to stem the English texts.

```
import nltk
from nltk.stem import PorterStemmer #English texts only
ps= PorterStemmer ()

for w in raw3:
    print (ps.stem (w))
```

The Arabic text, on the other hand, was stemmed using the Information Science Research Institute's (ISRI) Arabic Stemmer by Taghva et al. (2005).²⁰⁸ However, it should be mentioned that again, as when removing the Arabic stop-words (function words),²⁰⁹ the morphologically rich nature of the Arabic language poses challenges for the Arabic stemmer for it misses some of the words belonging to the same concept of nature (See example in Table 18). For instance, the various word forms belonging to the concept *day* in the Qur'an are stemmed, and the ISRI Arabic Stemmer misses several occurrences. To

207 [Chapter Two].

208 ISRI stemmer (Taghva et al., 2005) [found in: <http://nltk.org>].

209 *Function words* are defined as "a set of words sometimes referred to as grammatical words, consisting of pronouns, prepositions, determiners, conjunctions, auxiliary and modal verbs". See also Baker et al. (2006, p.76).

solve this problem, a unifying system is developed whereby one abstract Arabic root²¹⁰ represents a natural term and its different morphological and semantic realisations. The assigning of roots was based on the *Qur'anic Arabic Corpus* and checked against Lane's acclaimed *Classical Dictionary of Arabic* (1863),²¹¹ and variation was reduced by extracting the root for each of the terms. In an experiment to prepare the list of natural phenomena to be computed for the exploration of the semantic prosody of nature in the Qur'an, the root-based disambiguation took place on three levels; they are as follows:

- 1- *Word level*: different forms of the same word with differing stems. An example of this is the root *يوم/ywm/*, which was used to unify the different forms of the word 'day' in Arabic (see also Table 18):

Table 18: The word-level root-disambiguation of the word *day* in Arabic

<i>Arabic Word</i>	<i>English Equivalent</i>	<i>ISRI Arabic Stemmer</i>	<i>Root</i>
يَوْمٌ	a day <i>genitive</i>	يوم	يوم
وَالْيَوْمِ	in the day <i>genitive</i>	ليوم	يوم
يَوْمًا	a day <i>accusative</i>	يوم	يوم
وَالْيَوْمِ	and the day <i>genitive</i>	يوم	يوم
أَيَّامًا	days <i>accusative</i>	ايا	يوم
يَوْمٍ	a day <i>accusative</i>	يوم	يوم
أَيَّامٍ	days <i>genitive</i>	ايوم	يوم
يَوْمَيْنِ	two days	يوم	يوم
يَوْمَيْنِ	<i>accusative/genitive</i>	يوم	يوم
الْيَوْمِ	today/ the day <i>accusative</i>	اليوم	يوم
يَوْمٌ	a day <i>nominative</i>	يوم	يوم
يَوْمٌ	a day <i>genitive</i>	يوم	يوم
لِيَوْمٍ	for a day <i>genitive</i>	ليوم	يوم
الْأَيَّامِ	the days <i>genitive</i>	ايوم	يوم
يَوْمَئِذٍ	on that day <i>accusative</i>	مئذ	يوم

- 2- *Root level*: two or more words have the same root. It is also defined in terms of *homograph ambiguity* in Arabic NLP (Farghaly and Shaalan 2009) as a word belonging to more than one part of speech, such as *قدم qdm* which could be a verb

210 *Root, stem, and base* are all terms used to designate that part of a word that remains when all affixes have been removed. A *root* is a form which is not further analysable, either in terms of derivational or inflectional morphology. It is that part of word-form that remains when all inflectional and derivational affixes have been removed. A root is the basic part always present in a lexeme. In the form 'untouchables' the root is 'touch', to which first the suffix '-able', then the prefix 'un-' and finally the suffix '-s' have been added. In a compound word like 'wheelchair' there are two roots, 'wheel' and 'chair'. A *stem* is of concern only when dealing with inflectional morphology. In the form 'untouchables' the stem is 'untouchable'. A *base* is any form to which affixes of any kind can be added. That is, 'touchable' can act as a base for prefixation to give 'untouchable', but in this process 'touchable' could not be referred to as a root because it is analysable in terms of derivational morphology, nor as a stem since it is not the adding of inflectional affixes which is in question. See also Bauer (1983, pp.20-1).

211 Lane, E. 1863. *Arabic-English Lexicon*, 2 books in 4 vols. London and Edinburgh: Williams and Norgate. Available from: [http://www.tyndalearchive.com/TABS/Lane/], [Accessed 12 June 2018].

meaning ‘to introduce’ or a verb meaning ‘to arrive from’ or a noun meaning ‘foot’. Another example is taken from the context of the natural phenomenon of darkness with the root ظلم *zlm*, which represents two meanings: ‘darkness’ and ‘injustice’. To solve this problem, a number is added next to the root, resulting in 1ظلم to mean *darkness*.

- 3- *Synonym level*: different words that denote the same meaning of the natural phenomena. In this case, the synonyms of each term are included under the same root. Examples of concepts with synonymous terms are shown in Table 19.²¹²

Table 19: Examples of the synonym level root-disambiguation of the words

English Term	Arabic Term	Arabic Synonyms	Root
sky/s	سما	سماوات/سماء/ السماء/ السماوات/ المعارج/سقف/ كسفا	سمو
sun	شمس	الشمس/ سراجا	شمس
ship/ark	فلك	الفلك/ السفينة/ذات ألواح/الجوار	فلك

In addition, it should be mentioned that the English translations of the terms belonging to natural concepts in the Qur’an were also disambiguated at the synonym level, such as the synonymous words ‘skies’ and ‘heavens’ which are represented by the root *sky*. Finally, a list of the Arabic terms and their disambiguated roots is shown in the table below.

Table 20: The root-based disambiguation of natural phenomena terms

Arabic Term	Translation	Root-based disambiguation
ملك	angel/s	1ملك
حيوان	animal	2حي
ذرة	atom	1ذرر
ظهورهم	back/s	1ظهر
نحل	bee	1نحل
عظام	bone/s	1عظم
العجل	calf	1عجل
صدر	chest/s	1صدر
غمام	cloud/s	1غمم
غراب	crow	3غرب
ظلمات	darkness	1ظلم
الفجر	dawn	1فجر
نهار	day2	2نهر
قرين	devil	2قرن
تراب	dust	1ترب
أذن/أذان	ear/s	1أذن
صعيد	earth/soil	1صعد
الشرق	east	2شرق
أعين	eye/s	2عين

212 Alshahrani, H. and Brierley, C. (2018). The Root-Based Disambiguation of Words in the Annotation of Semantic Prosody of Nature as a Qur’anic Theme. A paper presented in *The Workshop on Computational Approaches to Morphologically Rich Languages (#CAMRL2018)*.

<i>Arabic Term</i>	<i>Translation</i>	<i>Root-based disambiguation</i>
وجه/وجوه	face/s	1وجه
أرجل	feet	2رجل
أجنة	fetus	4جنن
بنان	fingertips	1بنن
نار	fire	1نور
جنة/جنات	garden/s	1جنن
قرون	generation/s	1قرن
ذهب	gold	1ذهب
قلب	heart/s	1قلب
الجنة	heaven	2جنن
قدم	heel/s	1قدم
نار/جهنم	hell	2نور
حديد	iron	1حدد
جن/جان	jinn	3جنن
الشمال	left-hand	1شمل
حياة	life	1حيي
دنيا	life2	1دنو
نور	light	3نور
برق	lightning	1برق
رجل/رجال	man/men	1رجل
بعوض	mosquito	1بعض
طور	mount	1طور
قوم	people of	2قوم
أحمر	red	1حمر
القيامة	resurrection	1قوم
نهر/أنهار	river/s	1نهر
سبلا	roads	1سيل
غرابيب	rock	4غرب
ظل	shade	1ظلل
حرير	silk	1حرر
فضة	silver	1فضض
نفس/أنفس	soul/s	1نفس
عيون	spring/s	1عين
أطوار	stages	2طور
بطون	stomach/s	1بطن
حجارة/حجر	stone/s	1حجر
الشروق	sunrise	1شرق
غروب	sunset	2غرب
الأعراف	the heights	1عرف
الصور	the trumpet	1صور
كنز	treasure/s	1كنز
واد	valley/s	دوي
غرب	west	1غرب
رياح/ريج	wind/s	1ريج
جناح	wing/s	1جناح
الارحام	womb	2رحم
العالمين	worlds	1علم
سنة	year/s	1سنو

4.1.3 Statistical analysis

Following the pre-processing of the text and the discovery of the 30 most frequent natural phenomena²¹³ (see Chapter Two), two tasks of statistical analysis were undertaken:

²¹³ They were by default 30 words but have additional terms because some of the terms which were in the original list were by definition strictly positive (e.g. angels) or negative (e.g. devils, death). Hence, they were retained for investigation of their connotative

identifying collocations and ranking them in the Arabic Qur'an and the selected five translations. This involved the identification of four dimensions of collocation discovery (See Figure 28). A discussion of the methods applied to identify these features of collocations is provided in this subsection.

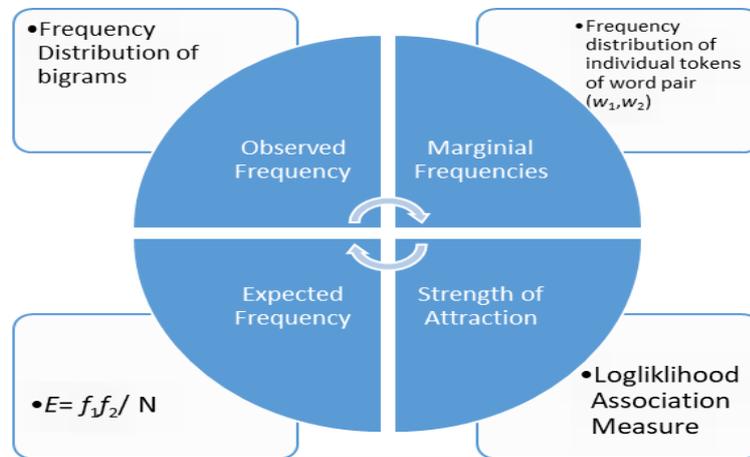


Figure 28: Statistical analysis of collocations of nature in the Qur'an

Moreover, the practical identification of collocations of the natural phenomena in the Qur'an adheres to the most common definitions and methods in corpus and computational linguistics. For example, the extracted collocations were bigrams, following the Manning and Schütze (1999) model. Similarly, the collocation extraction coincides with the definition of collocation by Bartsch and Evert (2014, p.76), which states that collocations are “lexically and pragmatically constrained recurrent co-occurrences of at least two lexical items which are in direct syntactic relation with each other”. Finally, this research highlights the Firthian notion of collocations (1957), which assumes mere “habitual” co-occurrence lexical items that occur a minimum number of times, usually at least five times (as cited in Bartsch and Evert, 2014, p.52). In addition, it follows the established paradigm of Evert (2008, p.41), which claims that collocation in this Firthian sense can be explored in terms of frequencies and association measures. He writes:

Introduced as an intuitively appealing, but fuzzy and pre-theoretical notion by Firth (1957), collocativity can be operationalised in terms of co-occurrence frequencies and quantified by mathematical association measures (p.41).

meanings [evaluative and discourse prosodies], and the following terms in the list were added to enrich the coverage of this theme in the Qur'an and its five translations. Hence, although they are 42 terms, they will be referred to as the 30 most frequent words for the sake of exploring the ones that are intuitively positive or negative by dictionary definition.

This research adopted Evert's five-step model (2008) of collocation extraction to process textual data in the exploration of the SP of nature in the Qur'an in the Arabic text and its five translations. It consists of the following steps:

- 1- Identifying the type of co-occurrence (*surface*, *textual*, or *syntactic*).²¹⁴
- 2- Finding the frequency signature (i.e., co-occurrence frequency and the marginal frequencies f_1 and f_2 in the corpus) for all relevant word pairs (w_1, w_2) as well as sample size N .²¹⁵
- 3- Filtering the occurrence dataset by applying a frequency threshold.
- 4- Calculating the expected frequencies of the word pairs,²¹⁶ using the general equation: $E = f_1 f_2 / N$ for textual and syntactic co-occurrence, and the approximation: $E = k f_1 f_2 / N$ for surface co-occurrence, where k is the total span size [window].
- 5- Applying an association measure [to score and rank the collocations from highest to lowest in the strength of association].

As the first step, the chosen *type of co-occurrence* was the surface co-occurrence, that is, the circumstances in which words are said to co-occur if they appear close to each other in a running text and are measured by the number of intervening words or tokens (Evert 2008, p.5). *Surface co-occurrence* is known as the most common approach in the Firthian tradition in which two words are said to co-occur if they appear within a certain distance or "collocational span" (Sinclair, 1991; Evert, 2008, pp.11-2). This surface co-occurrence is combined with the node-collocate view, that is, looking for collocates within the collocational spans around the instances of a node word (e.g., in this study, a word referring to natural phenomenon). The chosen span size k of tokens, per this model, was ten words (as in Sinclair 1991 and recommended by Baker et al. 2006). These collocational spans are symmetric and are described as five tokens to the left of the node word and five tokens to its right as follows: (L5 node R5).

The second step in this model was *determining the frequency signature* in the surface co-occurrences of the node and collocate patterns of natural phenomena in the Qur'an. The frequency signature of a bigram is comprised of four values: O for the observed co-

214 Evert (2008, p.4) claims that the co-occurrence of words can be defined in many different ways. *Surface co-occurrence*, often combined with a node-collocate view, refers to collocates within the collocational spans around the instances of a given node word. *Textual co-occurrence* considers words to co-occur if they appear in the same textual unit. Typically, such units are sentences or utterances. In *Syntactic co-occurrence*, words are only considered to be near each other if there is a direct syntactic relation between them. Examples are a verb and its object (or subject) noun, pronominal adjectives (in English and German) (as in Evert, 2008, pp.12-4).

215 See also Evert (2004, p.36).

216 Expected frequencies are the values we would get if the proportions of the overall corpus are reflected identically in each section of the corpus. They are found via a process of establishing the random co-occurrence baseline (Brezina, 2018, p. 69).

occurrence frequency in a given corpus (sometimes also denoted by f , especially when specifying frequency thresholds such as $f \geq 5$; f_1 and f_2 for the marginal frequencies of the first and second component of the word pair, respectively); and N for the sample size (Evert, 2005, p. 36). An illustrative example of the application of this step onto surface co-occurrence is found in the text shown in the following figure (Evert, 2008, p.13):

A vast deal of coolness and a peculiar degree of judgement, are **requisite in catching a hat**, A man must not be precipitate, or he runs over it; he must not rush into the opposite extreme, or he loses it altogether. [...] There was a fine gentle **wind, and Mr. Pickwick's hat rolled sportively before it**, The wind puffed, and Mr. **Pickwick puffed, and the hat rolled over and over**, as merrily as a lively porpoise in a strong tide; and on it might have *rolled*, far beyond Mr. Pickwick's reach, had not its course been providentially stopped, just as that gentleman was on the point of resigning it to its fate.

Figure 29: Illustration of surface co-occurrence for the word pair (*hat*, *roll*)²¹⁷

This figure above shows the surface co-occurrences between the words *hat* (in boldface, as the node) and *roll* (in italics, as collocate). The span size is four words, excluding punctuation and limited by sentence boundaries. Brackets below the text indicate collocational spans around instances of the node word *hat*. There are two co-occurrences in this example, in the second and third span, hence $O = 2$. Their overall occurrence gives the marginal frequencies of the two words counts in the text, i.e., $f_1 = 3$ for *hat* and $f_2 = 3$ for *roll*. The sample size N is the total number of tokens in the corpus, counting only tokens that are relevant to the definition of spans. In this example, N is the number of word tokens, excluding punctuation, i.e., $N = 111$ for the text shown in Figure 29. The full *frequency signature* for the pair (*hat*, *roll*) is thus (2, 3, 3, 111). This research applied the same concept, where each of the recurrent co-occurrences of the collocations of nature in the Qur'an is assigned a frequency signature.

On the practical level, the text processing required for this step entails the use of collocation finders in NLTK to provide the bigrams of each of the terms in the list of natural phenomena. For example, the most frequent bigrams in the text were found through the use of NLTK built-ins: `bigrams()` and `FreqDist()`.

²¹⁷ This is a toy corpus used by Evert 2008 consisting of 111-word tokens (excluding punctuation).

```
bigram_fd = nltk.FreqDist(nltk.bigrams(tokens))  
finder = BigramCollocationFinder(word_fd, bigram_fd)
```

To find the marginal frequencies of the word pairs in natural phenomena bigrams, Python was run against the set of individual tokens of the text to calculate the frequency distribution of the terms of the list of natural phenomena as seen in the following code:

```
from nltk.probability import FreqDist  
fdist= FreqDist (word for word in dataset)  
inspect= fdist
```

The third step following Evert's model was to *filter the occurrence dataset by applying a frequency threshold*. In this regard, theoretical considerations suggest a minimum frequency threshold of $f \geq 3$, $f \geq 5$ (Evert, 2008, p. 41), or $f \geq 10$. They are generally based on the size and content of the corpus used after the text pre-processing. Therefore, the chosen threshold for this research was $f \geq 3$ because the size of the corpora is limited to one Arabic text and five English representations of this text individually explored for the collocational phenomenon of SP. Hence, a filter was applied to remove low-frequency candidates (less than three occurrences in the corpus) as in the following code:

```
finder.apply_freq_filter(3)
```

The fourth step applied per Evert's model of text processing was the *calculation of the expected frequencies of the word pairs*, using the approximation $E = k f_1 f_2 / N$ for surface co-occurrence, where k is the total span size (i.e., ten in this research). The expected frequency of co-occurrence by chance E serves as a reference point for the interpretation of O : the pair was only considered collocational if the observed co-occurrence frequency O is substantially higher than the expected frequency E (Bartsch and Evert, 2014, p.49).

The fifth and final step was to apply the association measure to compute the significance of the attractions between the words in word pairs of nature in the Qur'an; that is, to reflect what extent the word co-occurrence is accidental. For this purpose, the association measure employed in this research is the Log-Likelihood Ratio (LLR), an association measure which was briefly described in the previous chapter. Developed by Dunning (1993), LLR is a method used to qualify the association between two words in bigrams by calculating the ratio between two likelihoods: the probability of observing one constituent of a collocation given the other is present, and the probability of observing the same constituent of collocation in the absence of the other (Dunning, 1993, p.61). It was selected as a preferred measure because it is said to perform well on all corpus sizes, and is claimed by Dunning (1993) to be more reliable with low frequencies (Daille et al., 1994, p. 174). It is defined as shown in the formula below, where k , i , and j are the respective counts. LLR measures

how much the observed joint distribution of words x and y differ from their distribution under the null hypothesis of independence, i.e., how sign the association between them is in the given corpus (Toivonen et al., 2013, p.3).

$$LLR(x,y) = -2 \sum_{i=1}^2 \sum_{j=1}^2 k_{ij} \log(p_{ij}^{null}/p_{ij}),$$

The LLR statistic gives a measure of the likelihood that two samples are not independent (i.e., generated by the same probability distribution) (See the following code). To compute the LLR score for collocates of natural phenomena in the Quran, the frequency signatures in the second step of Evert's model (2008) were employed to provide a scored list of collocation for each nature terms as seen in the following lines of code:

```
from nltk.collocations import *
bigram_measures = nltk.collocations.BigramAssocMeasures()
print('%0.2f' % bigram_measures.likelihood_ratio(n_ii, (n_ix, n_xi), n_xx))
```

An example of the application of this code on the collocates of term 'earth' in Ali's translation is shown following code and its results:

Input:

```
print('%0.2f' % bigram_measures.likelihood_ratio(190 , (200 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(227 , (2895 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(52 , (153 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(54 , (379 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(25 , (33 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(21 , (34 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(21 , (35 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(51 , (419 , 775) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(21 , (49 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(14 , (14 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(19 , (419 , 39) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(20 , (51 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(35 , (419 , 337) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(31 , (287 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(18 , (47 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(29 , (419 , 239) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(20 , (80 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(26 , (213 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(17 , (54 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(29 , (419 , 310) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(17 , (419 , 73) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(24 , (419 , 191) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(14 , (41 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(14 , (42 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(12 , (24 , 419) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(15 , (419 , 63) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(16 , (419 , 82) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(19 , (419 , 144) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(8 , (419 , 8) , 63688))
print('%0.2f' % bigram_measures.likelihood_ratio(11 , (419 , 22) , 63688))
```

Output:

```
1932.47
859.38
334.22
243.21
216.25
167.00
165.13
151.20
145.49
141.14
137.99
134.00
133.56
120.45
119.45
119.35
112.70
107.14
104.70
104.27
92.98
100.21
88.84
88.03
87.80
82.71
81.27
81.01
80.53
80.46
```

The researcher also experimented with another form of automated bigram exploration via LLR in NLTK with Python. In this type of code as seen in the sample below, she only specified the nature term, and the program ranked the collocations automatically.

```
from nltk.collocations import BigramCollocationFinder
from nltk.metrics import BigramAssocMeasures
from nltk.collocations import *
from nltk import FreqDist
#Pickthall Bigrams and their LLR Scores
night_filter = lambda *w: 'night' not in w
finder = BigramCollocationFinder.from_words(text, window_size = 10)
finder.apply_freq_filter(3)
finder.apply_ngram_filter(night_filter)
m= finder.score_ngrams(bigram_measures.likelihood_ratio)
for item in m:
    print (item)
```

Output:

```
((('night', 'day'), 17.311175543238136)
 (('night', 'moon'), 7.716540258444388)
 (('causes', 'night'), 7.209854396688973)
 (('night', 'sun'), 7.029688723460942)
```

To conclude, when the LLR score of a pair is low, the words are said to be independent, and it is high otherwise (i.e., words are highly associated) (Sharoff et al., 2006, p.3). More on the parameters of the collocation identification (e.g., the collocation notation parameters)

in this corpus-based research and its visual representation is discussed in Section 4.1.5 of this chapter.

The following subsection describes how the researcher interpreted the quantitative results of this subsection (e.g., the produced and ranked lists of the ranked bigrams of nature terms in the Qur'an and its translations) to find their evaluative and discourse prosodies.

4.1.4 Finding the SP of nature in the Qur'an

In keeping with the theoretical framework built for this research, it is of importance to give a clear picture of how the present study was conducted according to this methodology to develop a systematic way of finding the SP of natural phenomena in the Qur'an. This system was followed to find SP of nature in the Arabic Qur'an in this research and later used to compare the Arabic Qur'an to its translations. To elaborate, as previously mentioned, the resulting 30 most frequent natural phenomena were examined for their collocations by computing their bigrams via LLR. The bigrams for each term were the ones to consider in this research for two main purposes. The first purpose of finding the bigrams of nature is to extract the lexico-grammatical patterns of nature in the Qur'an, which are the founding blocks of unveiling the subliminal meanings connected with the theme of nature in the Qur'an. The second purpose of employing bigrams of nature in this research was to annotate the evaluative prosodies (EPs) and discourse prosodies (DPs) of natural phenomena in the Qur'an. The percentage of the highest EPs and PDs for each of the nature terms was assigned as its EP and DP. Hence, the exploration of SP was premised on the analysis of both the evaluative and discourse forms of SP via discovering patterns of usage and meanings (as in Sinclair, 2004a and Stubbs, 2007, p.155).

To obtain these meanings after processing the text,²¹⁸ Sinclair's (2003, xvi-xvii) seven-step procedure of reading concordances was followed. The steps are as follows:

- 1- *Initiate*. This involves examining the lexico-grammatical patterns for each node after collocation extraction in the previous section by looking at the words that occur immediately to the right and left of the node and identifying the highest ranking pattern. Table 21 shows the categorisation of the collocates of nature terms into syntactic sets.

218 Examples provided in this subsection are from both the Arabic text as well as the English translations.

Table 21: A sample of the syntactic categorisation of collocates of natural phenomena

<i>Nouns (n.)</i>	<i>Verbs (v.)</i>	<i>Adjectives (adj.)</i>	<i>Adverbs (adv.)</i>
sign	said	many	indeed
signs	sent	grateful	verily
truth	made	subservient	certainly
dominion	praise	six	
Qur'an	say	merciful	
[Prophet] Lut	disbelieve	lost	
Noah	establish	spacious	
Thamud	know	evil	
Moses	judge	good	
woe	created	righteous	
bounty	merge	sealed	
Pharaoh	belong	right	
throne	emerge	black	
thing	assemble	white	
mischief	earth swallow	the warned	
sustenance	appointed	alike	
creation	send		
All-Mighty	descend		
All-Wise	wronged		
effort	earned		
prophet	paid		
injustices	burden		
warning	witness		
brother	punished		
worshiper	doeth		
triumph	guide		
ignorance	worship		
promise	waste		
disbelief	abode		
doom	pray		
punishment	taste		
glitter	invite		
glory	enter		
reward	reject		
Eden	glorify		
righteousness	give		
works	gave		
believe	remind		
bliss	flow		
disease	understand		
book	spread		
Adam	bow		
enemy	said		
fire	follow		
creator	subjected		

Nouns (n.)	Verbs (v.)	Adjectives (adj.)	Adverbs (adv.)
blessings	come		
darkness	submit		
light	raise		
bounties	warn		
blast	conceal		
flood	drink		
grapevines	drown		
light	split		
	remember		
	cause		

Interpret. After finding the lexico-grammatical (LG) patterns and syntactically categorising collocates surrounding the nodes, these LG patterns were used to navigate concordance lines and to interpret their meanings when aligned with their verses. For example, it was useful to use concordances via Sketch Engine to navigate the context of the collocation (e.g., bigram) found in the textual analysis of the collocation extractions. An example of the use of copied concordances is shown below in the figure of the Excel sheet, where the word نور²¹⁹ *nwr* ‘light’ is analysed. This nature phenomenon Arabic root includes both the literal meanings of ‘light’, for example, in words such as ضياء ‘translated as light of the sun’ and مصابيح ‘translated as lanterns when describing the light of the stars in the sky’ as well as an attribute of God.²²⁰ In addition, verified collocates of this root as a token from the previous subsection were marked in red in the concordance lines, and the corresponding verses are aligned to the right. With the microanalysis of the word via consultation of the Tafsīr and Asbab Al-Nuzul (e.g., the contexts and occasions of the Revelation of the Qur’an), the SP of natural phenomena in the Qur’an was explored.

H	
Concordance lines	Verse
الله وليا نصيرا ناس جاء لك م برهان ريكم وانزلنا > نور3 < مبينا امنوا ب الله واعتصموا ف س يدخل هم رحمة و فضل و	4
يهدى الله اتبع رضوانه سبيل السلام و يخرج هم ظلم1 > نور3 < باذنه و يهدي هم صراط مستقيم ك فر قالوا الله المسيح	5 16
الله يتولون اولئك ب المؤمنين انزلنا التوراة هدى > نور3 < يحكم النبيون اسلموا هادوا و الريانيون و الاحبار	
هم ل يجادلوكم اطعمتمو هم ل مشركون مبناحي1 و جعلنا > نور3 < يمضي ناس مثل ه ظلم1 ب خارج زين ل الكافرين يعملون	
يضع اصم هم و الاغلال امنوا و عزروه و نصره و اتبعوا > نور3 < انزل اولئك المفلحون قل ناس رسول الله جميعا ملك	
الها واحدا اله سبحانه يشركون يريدون يظفوننا > نور3 < الله ف و هر و ياني الله يتم نور3 كره الكافرون اربل	9 32
يريدون يظفوننا نور3 الله ف و هر و ياني الله يتم > نور3 < كره الكافرون اربل رسول هر ب الهدى و دين الحق ل يظهر هر	9 32
نفس1 نفعنا ضرر قل يستوي الاعمى عين2 تستوي ظلم1 > نور3 < جعلوا ل ل هر شركاء خلقوا ك خلق هر ف تشاب هر الخلق قل الله	
الله الرحمن الرحيم كتاب انزلناه ل تخرج ناس ظلم1 > نور3 < ب اذن رب هم صراط العزيز الحميد الله سمو ارض و ويل ل	
الحكيم ارسلنا موسى باية نا اخرج قوم2 الظلمات > نور3 < و ذكر هم يوم الله ل آيات صبار شكور قال موسى ل قوم هر	4 15

Figure 30: A concordance of the nature term *light* in the Qur’an

219 The number ‘3’ next to the Arabic word in the figure represents the root-disambiguation of the word (See also Section 4.1.2).

220 See also the Quranic Arabic Corpus, Available from : <http://corpus.quran.com/search.jsp?q=light>

However, since many concordance searches of nature terms in this research can produce hundreds of lines in the Qur'an datasets (source text and target texts), this research follows Sinclair (1999) in selecting 30 random lines in which the statistically significant bigrams appear and noting patterns in them, then selecting a different 30, noting the new patterns, and so on until a further selection of 30 lines reveals nothing new (Baker et al., 2006, p.44). This is done to perform a content analysis to interpret the meanings of the co-occurrences of nature in the Qur'an. Thus, the analysis of the first 30 concordance lines for the nature terms helps in forming a hypothesis that may link repeated words (e.g., they are from the same word class or have similar meanings). Their link was generally established by categorising the meanings into categories.

- 2- *Consolidate*. This means looking for other evidence (other collocations) to support the hypothesis formed from the first 30 lines of the concordance in Step 2, extending beyond the word positions first studied and revising the hypothesis by looking at the next 30 concordance lines of patterns of natural phenomena. This was done by conducting a form of content analysis of nature in the Qur'an to elicit the pragmatic functions²²¹ of the lexico-grammatical patterns of natural phenomena (See figure below and for details and examples see qualitative analysis in sub-section 4.1.5).

Concordance lines	Verse	Tafseer	Translation of tafseer	Pragmatic Meaning	
3 نور	إِنَّا أَنْزَلْنَا الْقُرْآنَ فِيهَا هُنَّىٰ وَنُورٌ يَجْعَلُ الْبَارِئِينَ السَّمَوَاتِ وَالْأَرْضِ فِيهَا نَبَأٌ بَشِيرٌ لِّمَنْ هُوَ مِنَ الْإِنْسِ كَذِبٌ	إِنَّا أَنْزَلْنَا الْقُرْآنَ فِيهَا هُنَّىٰ وَنُورٌ يَجْعَلُ الْبَارِئِينَ السَّمَوَاتِ وَالْأَرْضِ فِيهَا نَبَأٌ بَشِيرٌ لِّمَنْ هُوَ مِنَ الْإِنْسِ كَذِبٌ	إِنَّا أَنْزَلْنَا الْقُرْآنَ فِيهَا هُدًى مِّنَ الضَّلَالَةِ	[5:44] Surely We revealed the Qur'an in it as a guidance and light for the believers	emphasis of truth of believers
1 و جعلنا	وَمَنْ كَانَ مِنَّا فَاتِنًا فَجَعَلْنَا لَمَنَؤُنَّ مِنَّا فِي النَّاسِ حَرَجًا	وَمَنْ كَانَ مِنَّا فَاتِنًا فَجَعَلْنَا لَمَنَؤُنَّ مِنَّا فِي النَّاسِ حَرَجًا	«أَوْ مِنْ كَانِ مَنَّا فَاتِنًا فَجَعَلْنَا لَمَنَؤُنَّ مِنَّا فِي النَّاسِ حَرَجًا»	[6:122] The following was revealed in it as a guidance and light for the believers	emphasis of truth of believers
نور	يَضَعُ أَصْرَهُمُ وَالْإِغْلَالَ إِنْتُوا وَعِزُّهُمْ وَتَضْرِبُ نَاصِيَهُمْ	يَضَعُ أَصْرَهُمُ وَالْإِغْلَالَ إِنْتُوا وَعِزُّهُمْ وَتَضْرِبُ نَاصِيَهُمْ	«الَّذِينَ يَتَّبِعُونَ الرَّسُولَ النَّبِيَّ الَّذِي جَاءَهُمْ بِبَيِّنَاتٍ مِّنَ رَبِّهِمْ يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[7:157] those who follow the Messenger of Allah, they desire to extinguish the light of Allah	emphasis of truth of believers
3 نور	إِنَّمَا وَجَدَهَا بِرَبِّهَا وَأَنَّهَا مِن بَيْنِ أَيْدِيهِمْ وَأَنَّهَا فِي بَيْنِ يَدَيْهِمْ	إِنَّمَا وَجَدَهَا بِرَبِّهَا وَأَنَّهَا مِن بَيْنِ أَيْدِيهِمْ وَأَنَّهَا فِي بَيْنِ يَدَيْهِمْ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[9:32] They desire to extinguish the light of Allah	emphasis of truth of believers
3 نور	نَفْسًا نَّفَعًا خَيْرًا لِّلَّذِينَ آمَنُوا وَلَئِن لَّمْ يَكُن لَّهُمْ نَفْسٌ مِّنْ نَّفْسٍ لَّوْ كُنُوا نَارًا	نَفْسًا نَّفَعًا خَيْرًا لِّلَّذِينَ آمَنُوا وَلَئِن لَّمْ يَكُن لَّهُمْ نَفْسٌ مِّنْ نَّفْسٍ لَّوْ كُنُوا نَارًا	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[9:32] They desire to extinguish the light of Allah	emphasis of truth of believers
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[13:16] Say, O Muhammad (ﷺ), we only worship Allah	emphasis of truth of believers
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[14:1] Alif lām rā': God knows what you do not know	emphasis of truth of believers
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[14:5] And verily We sent Messengers before thee with clear signs and revelation	emphasis of truth of believers
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[21:48] And verily We gave Noah the Scripture and revelation	emphasis of truth of believers
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:35] God is the Light of the world	glorifying of God
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:40] Or, [it is that] those who disbelieve	disbelievers
3 نور	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	إِنَّمَا نَحْنُ نَعْبُدُ اللَّهَ وَنَحْنُ عَلَىٰ بَدَنِ اللَّهِ كَمَا نَحْنُ نَعْبُدُ اللَّهَ	«يُرِيدُونَ أَن يُقْلِقُوا نُورَ اللَّهِ»	[24:40] Or, [it is that] those who disbelieve	disbelievers

Figure 31: The discourse prosodies of the word *light*

- 3- *Report*. It means the labelling of the evaluative prosodies for each of the patterns of natural phenomenon as being *positive*, *negative*, or *neutral* based on the acquired connotative meanings from the co-text²²² in each of the co-occurrences. Figure 32

221 Pragmatic functions are also referred to as discourse prosodies in this thesis.

222 *Co-text* is the text occurring around a node, as can be seen in a concordance (Baker et al., 2006, p.199).

shows three different SP meanings of the word *water*.²²³ It can have positive and negative EPs when its neighbouring collocates, such as ‘tidings’, ‘send’, and ‘scalding’ respectively refer to the emphasis on the message of the Qur’an, glorifying God and his creation, and punishment in the afterlife.²²⁴

verse	Tafseer	Water	.T	Bigram	Dis. SP
And it is He who sends the winds as good tidings before His mercy, and We send down from the sky pure water	[25:48] And He it is Who sends forth the winds (al-riyāḥa; a variant has al-rīḥa, 'the wind') dispersing before His mercy (nushuran, 'dispersing'), scattered before the [coming of the] rain (a variant reading has nushran, the singular of which is nashūr, similar [in pattern] to rasūl, 'messenger' [plural rusul]; another reading has nashran, as a verbal noun; and a third variant has bushran, in other words, [the winds function as] mubashshirāt, 'bearers of good tidings', the singular of which is bashīr); and We send down from the				
Then indeed, they will have after it a mixture of scalding water.	[37:67] Then, lo!, on top of it they will have a brew of boiling water, which they drink and which mixes with what they have eaten and becomes a brew thereof.	pos		(sends,sky),(sky,water),(sends,w	glorifying of God /emphasis o
		neg		(scalding,water)	punishment 2

Figure 32: The discourse prosodies of the word *water*

The following step was to manually annotate the verses in which these patterns occur with the appropriate evaluative and discourse prosodies for each of the nature terms in the Qur’an. It should be noted that when calculating the evaluative and discourse prosodies of nature terms, it was necessary to go back to the contextual meanings of the terms and consult the Tafsīr of the verses for the further assertion of the meanings and to avoid subjectivity.

- 4- *Recycle*. This step means starting with the next most important pattern nearest to the node and looking for anything unusual in any remaining data.
- 5- *Results*. The manually calculated SP of nature in the Qur’an (the percentage of the evaluative and discourse prosodies for each of the natural phenomena) was reported (See also quantitative analysis in section 4.1.5). The highest percentage of EPs and DPs for each natural phenomenon word is said to be its evaluative prosody and discourse prosody. For example, Figure 33 illustrates the different pragmatic functions of the words *people*, *day*, *sky/heavens*, *earth*, and *water* in the Arabic dataset. Their discourse prosodies meanings or pragmatic functions were found by

223 Translation of 1997. *Ṣaḥīḥ International The Qur'an: Arabic Text with Corresponding English Meanings*. Riyadh: Abul-Qasim Publishing House.

224 Referred to as ‘punishment 2’ in this research.

calculating the percentage of the DPs over the total frequencies of each of the words in the text.

2	people		day			
3	Row Labels	Count of Dis. SP	Row Labels	Count of Dis. SP	sky/heavens	
4	believers	3	believers	4	Row Labels	Count of Dis.
5	disbelievers	42	believers	4	emphasis on message	
6	disbelievers/emphasis on message	1	disbelievers	4	emphasis on message	
7	disbelievers/glorifying of God	1	emphasis on message	3	glorifying of God	
8	emphasis on message	43	emphasis on message /reward	1	glorifying of God/emphasis on message	
9	emphasis on message	2	glorifying of God	2	glorifying of God/emphasis on message	
10	emphasis on message/disbelievers	1	glorifying of God/emphasis on message	3	punishment 2/glorifying of God	
11	glorifying /emphasis on message	1	Islamic teaching	3	reward	
12	glorifying of God	3	punishment 2	13	Grand Total	
13	glorifying of God/disbelievers	1	punishment 2/glorifying of God	1		
14	glorifying of God/emphasis on message	5	reward	1		
15	glorifying of God/emphasis on message	1	wrong doers	1		
16	glorifying of God/emphasis on truth	1	Grand Total	40		
17	hypocrites	1				
18	Islamic teaching	2				
19	miracle story	7	earth		water	
20	punishment 1	11	Row Labels	Count of Dis. SP	Row Labels	Count of Dis.
21	punishment 2	1	disbelievers	2	glorifying of God	
22	transgressors	1	emphasis on message	1	glorifying of God /emphasis on message	
23	wrong doers	5	glorifying of God	17	glorifying of God/emphasis on message	
24	Grand Total	133	glorifying of God/emphasis on message	3	glorifying of God/emphasis on message	

Figure 33: A dataset of the pragmatic functions to find discourse SP in the Arabic text

- 6- *Repeat*. In this step, a new selection of data (30 concordance lines of another natural phenomenon term) from the corpus is taken, and the same steps of reporting the analysis are followed.

The following section will discuss the types of analyses employed to explore the SP of nature as a theme in the Qur'an with examples.

4.1.5 Data analysis in this research

Given that the datasets resulting from this research were both statistical and linguistic, it was important to employ a mixed-method approach, i.e., a quantitative and qualitative analysis, in the data analysis. The quantitative side involved a statistical analysis of the frequencies and bigrams identified using NLP applications with Python to rank collocations of nature in the Qur'an. Following that, the qualitative analysis was performed to elicit SPs (i.e., EPs and DPs) of natural phenomena in the Qur'an. Finally, to obtain EPs and DPs for each of the nature terms, the percentage of their evaluative and discourse prosodies in relation to their total frequencies in the text were calculated. Figure 34 summarises the approach to data analyses used in finding the SP of natural phenomena in the Qur'an. It can be broken down into three steps, and its results are used in the comparison in the following stage.

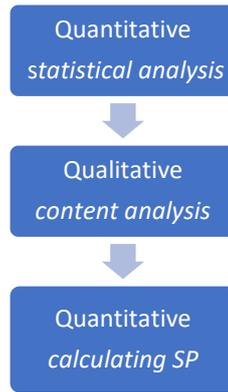


Figure 34: The data analysis of SP of natural phenomena in the Qur'an

1- *The quantitative analysis:*

When determining the SP of the 30 most frequent words describing natural phenomena (cf. Sinclair, 2004b), this research examined the units of meaning of the collocational framework containing a word that represents a concept related to nature, as well as its surroundings which ascribe the word with positive, negative or neutral associations only through the statistical and qualitative analyses of their collocations (See also Coffin et al., 2014; and see Table 27 for the most frequently mentioned natural phenomena in the Qur'an, Section 5.1). To do so, it was necessary to include some words that are neutral by default in the list of the 30 most frequent terms. However, there are words in the list that have by default and dictionary definition either positive or negative associations²²⁵ already, such as the afterlife creations of *Heaven* or the *Garden* (positive association with promised reward) and *Hell* (negative association with severe punishment). Other words include living creatures belonging to the spiritual beings of *angels* (positive association with benevolence), *devils* (negative association with evil spirits), and the weather phenomenon *flood* as a noun (negative association with destruction). Interestingly, the word *death*, which, according to its meaning as per the dictionary, has a negative association with grief, is seen in the Qur'an as part of nature's cycle of life and an inevitable beginning of the afterlife, leading to the Hereafter. See the following examples:

كُلُّ نَفْسٍ ذَائِقَةُ الْمَوْتِ ثُمَّ إِلَيْنَا تُرْجَعُونَ ﴿٥٧﴾

225 <https://en.oxforddictionaries.com>

Saheeh International: Every soul will taste death. Then to Us will you be returned. Verse (29:57)

وَهُوَ الَّذِي خَلَقَ السَّمَوَاتِ وَالْأَرْضَ فِي سِتَّةِ أَيَّامٍ وَكَرَّاتٍ عَشْرَةً عَلَى الْمَاءِ
لِيَسْأَلَكُمْ بِكُمْ أَحْسَنُ عَمَلًا وَلِيُنذِرَ لَكُمْ مَبْعُوثَاتٍ مِنْ بَعْدِ الْمَوْتِ
لِيَقُولَ الَّذِينَ كَفَرُوا إِنْ هَذَا إِلَّا سِحْرٌ مُبِينٌ ﴿٥٧﴾

Saheeh International: And it is He who created the heavens and the earth in six days - and His Throne had been upon water - that He might test you as to which of you is best indeed. But if you say, 'Indeed, you are resurrected after death,' those who disbelieve will surely say, 'This is not but obvious magic.' Verse (11:7)

وَاللَّهُ أَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَحْيَا بِهِ الْأَرْضَ بَعْدَ مَوْتِهَا إِنَّ فِي ذَلِكَ لَآيَةً لِقَوْمٍ يَسْمَعُونَ ﴿١٦﴾

Saheeh International: And Allah has sent down rain from the sky and given life, thereby to the earth after its lifelessness. Indeed in that is a sign for a people who listen. Verse (16:65)

The aforementioned words are interesting for their salience by frequency in the Qur'an, rendering them integral parts of the co-occurrence of natural phenomena. Although it seemed unlikely that they will display the tendency to acquire other connotative colourings, they were not disregarded. Therefore, although they are not excluded, the list was merely expanded to include additional terms, the last of which is the word 'star', which occurs 13 times in the Qur'an.

A further step in the implementation of the distributional approach (adopted from Evert, 2005; 2008) was taken through statistical analysis to reveal both the structures of collocations of natural phenomena in the Qur'an. An example of the study of the nodes belonging to natural phenomena in the Qur'an, both in Arabic and English, are in the following figure, which shows their frequencies, morphological and semantic variations, and translations.

Arabic Quran	F	G	H	K	L
	Nature Term	Arabic Frequency	Pickthall	All Frequency	Arberry
ناس-أهل-قوم-أقوام	people	534	people	511	people
يوم-أيام-يومان	day	481	day	520	day
أرض-خسف	earth(s)	462	earth(s)	419	earth(s)
سما-سماوات	sky(s)/heavens	324	sky(s)/heavens	247	sky(s)/heavens
نفس-أنفس-نفوس	soul(s)	295	soul(s)	176	soul(s)
القيامة-البعث	resurrection	276	resurrection	20	resurrection
النار-جهنم	hell	255	hell/fire	81	hell/fire
حياة-الدنيا	life	255	life	213	life
رجل-رجال-إنسان	man/human/men	151	man/human/men	574	man/human/men
جنة-جنت	heaven/the garden	144	heaven/the garden	129	heaven/the garden
قلب-قلوب	heart(s)	142	heart(s)	161	heart(s)
العالمين	worlds	61	beings	49	all beings
قرية-قرى	village/town	55	village/town	17	village/town
جبل-جبال-رواسي-أعلا	mountain	53	mountain	45	hills
أنهار	river(s)	53	river(s)	55	river(s)
نساء	women	49	women	98	women
نور-ضياء-مصباح	light	47	light	78	light
صدر-صدور	chest(s)/breast(s)	44	chest(s)/breast(s)	14	chest(s)/breast(s)
بحر-بحرين	sea	39	sea (s)	42	sea (s)
سنة-سنتين	year(s)	35	year(s)	36	year(s)
عين-عينات-سلسيل	springs	20	watersprings	32	fountain(s)

Figure 35: The frequencies of natural phenomena nodes

Furthermore, following Evert (2008), this research filtered the bigrams of nature in the Qur'an by calculating their the LLR association scores. The ranked lists of collocates of the word 'garden' as seen in the figure are the ones that were later used to annotate the verses for SP.

Sahih	140	N=57428	k=10				
Index	Position	Collocate	Observed Freq	f2	Freq Sign	Expected freq	LLR
1	R	flow	35	41	35_41_140_57428	0.999512433	22.8437196
2	R	rivers	36	53	36_53_140_57428	1.292052657	22.7638358
3	L	admit	15	28	15_28_140_57428	0.682593857	11.1264208
4	L	righteous	26	167	26_167_140_57428	4.071184788	10.5338736
5	R	abide	19	55	19_55_140_57428	1.340809361	9.31628277
6	R	perpetual	11	17	11_17_140_57428	0.414431984	9.18173307
7	R	residence	11	26	11_26_140_57428	0.633837153	7.93435523
8	L	deeds	16	135	16_135_140_57428	3.291077523	6.9902751

Figure 36: A sample of the statistical representation of the collocations of nature in the Qur'an in accordance with Evert's model (2008)

In addition, it should be noted that it is at this point that the visual representation of some of the collocations of nature was implemented using *GraphColl* in *LancsBox* and via the Log-likelihood statistic. To check that the LLR scores from both Python and *LancsBox* were similar, the researcher ran the same dataset in both. To visually represent a word via *LancsBox*, the researcher employed the following lines of code in Python to score each of the found collocations in Steps 2 and 3 of Evert's model (2008):

```
from nltk.collocations import *
bigram_measures = nltk.collocations.BigramAssocMeasures()
print('%0.2f' % bigram_measures.likelihood_ratio(n_ii, (n_ix, n_xi), n_xx))
```

An example of the resulting scores of the word 'earth' in Ali's translation is shown in the table below:

Table 22: The Python and *LancsBox* LLR scores of bigrams of the word 'earth' in Ali's translation

Index	Position	Collocate	Observed Freq	f2	Freq Sign	Python	LancsBox
1	L	heavens	190	200	190_200_419_63688	1932.47	1931.17761
2	L	Allah	227	2895	227_2895_419_63688	859.38	857.975075
3	L	created	52	153	52_153_419_63688	334.22	333.803201
4	L	things	54	379	54_379_419_63688	243.21	242.784882
5	L	belongs	25	33	25_33_419_63688	216.25	216.04542
6	L	belong	21	34	21_34_419_63688	167	166.822789
7	L	dominion	21	35	21_35_419_63688	165.13	164.958461
8	R	say	51	775	51_419_775_63688	151.2	150.824677
9	L	heaven	21	49	21_49_419_63688	145.49	145.320649
11	L	belongeth	14	14	14_14_419_63688	141.14	141.022313

<i>Index</i>	<i>Position</i>	<i>Collocate</i>	<i>Observed Freq</i>	<i>f2</i>	<i>Freq Sign</i>	<i>Python</i>	<i>LancsBox</i>
12	R	mountains	19	39	19_419_39_63688	137.99	137.832691
13	M	creation	20	51	20_51_419_63688	134	133.835404
14	R	signs	35	337	35_419_337_63688	133.56	133.28924
15	L	made	31	287	31_287_419_63688	120.45	120.207916
16	L	sky	18	47	18_47_419_63688	119.45	119.300733
17	R	see	29	239	29_419_239_63688	119.35	119.123114
18	L	knows	20	80	20_80_419_63688	112.7	112.537886
19	L	life	26	213	26_213_419_63688	107.14	106.938765
20	L	mischief	17	54	17_54_419_63688	104.7	104.563352
21	R	verily	29	310	29_419_310_63688	104.27	104.047599
22	R	glory	18	73	17_419_73_63688	92.98	100.64114
23	R	power	24	191	24_419_191_63688	100.21	100.01898
24	L	gives	14	41	14_41_419_63688	88.84	88.7242555
25	M	rain	14	42	14_42_419_63688	88.03	87.9139996
26	L	travel	12	24	12_24_419_63688	87.8	87.7014941
27	L	praise	15	63	15_419_63_63688	82.71	82.587466
28	R	end	16	82	16_419_82_63688	81.27	81.1436209
29	R	forth	19	144	19_419_144_63688	81.01	80.8638602
30	R	spacious	8	8	8_419_8_63688	80.53	80.4680481
31	R	spread	11	22	11_419_22_63688	80.46	80.366752

Finally, the statistical analysis in this methodology can be represented via Brezina et al.’s (2015) and Brezina’s (2018) collocation notation parameters as seen in the following table:

Table 23: The collocation parameters notation (CPN) of natural phenomena in the Quran

<i>Statistic ID</i> ²²⁶	<i>Statistic name</i> ²²⁷	<i>Statistic cut-off value</i>	<i>L and R span</i>	<i>Minimum collocate freq. (C)</i>	<i>Minimum collocation freq. (NC)</i>	<i>Filter</i>
6a	LL	6.63 ²²⁸	5L-5R	3	3	function words removed; ranked collocates considered

It is also summarised as follows:

Collocates of natural phenomena in the Quran: 6a-LL (6.63), 5L-5R, C3-NC3; function words removed; ranked collocates considered

²²⁶ In their Appendix 1, which lists the default association measures implemented by *GraphColl*, Brezina et al. (2015, pp168-72) give each association measure a unique identifier called a *Statistic ID*. They are also available in Evert (2004, Section 3; 2010).

²²⁷ The name of the employed *association measure*.

²²⁸ It is the default cut-off in the LancsBox toolkit, and it is claimed that the statistical significance at a confidence level of 99% is generally reflected in LL scores of 6.63 and over (Rayson 2009).

An example of a visual representation in LanksBox is in Figure 37, which shows the bigrams of the word **العالمين** *al'ālamīn*, translated as 'beings' by Arberry.

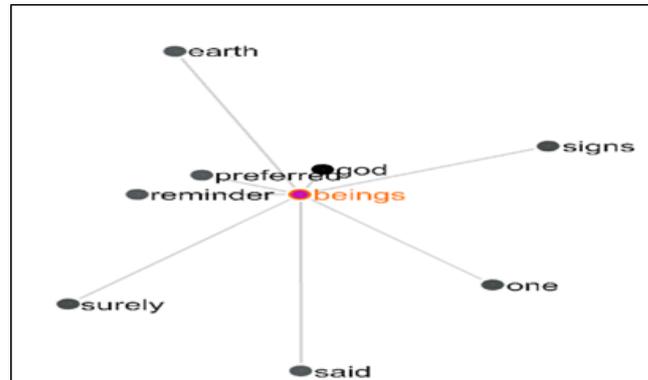


Figure 37: Collocates of **العالمين** *al'ālamīn*, translated as 'beings' by Arberry in the Quran : 6a-LL (6.63), 5L-5R, C3-NC3; function words removed; ranked collocates considered

The following task in the quantitative analysis was to examine the bigrams of nature terms as lexico-grammatical patterns and categorise them into groups of pragmatic functions of nature in the Qur'an. Their examination, following Brezina's approach to the analysis of the lexico-grammatical features (2018, p.104), included a quantitative analysis of their frequencies both in the Arabic ST and those in their translations in the five English TTs as sub-corpora of the whole corpus of natural phenomena occurrences in the Qur'an and its translations. A sample of an overview of the statistical profiling of the lexico-grammatical patterns of nature in the corpora of this research is shown in the figure below, and results of their comparison will be discussed in the following chapter.

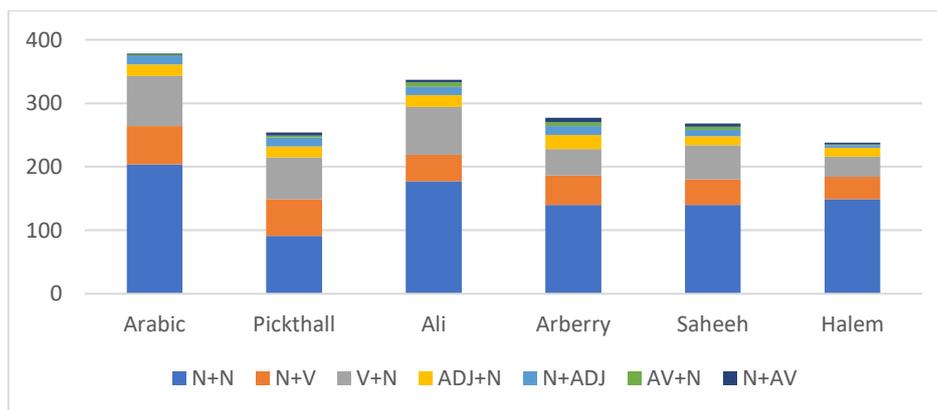


Figure 38: The LG patterns of natural phenomena in the Qur'an and its translations

However, the categorisation of the LG patterns of natural phenomena terms in this research was qualitative and will be discussed in the following sub-section. Subsequently, to

quantitatively obtain evaluative and discourse prosodies, the percentage of the EPs and SPs over the total frequencies in the text were calculated. Once the datasets of the Arabic Qur'an and its five translations were annotated with these two attributes, the qualitative analysis was conducted to compare the translations of the Qur'an as will be seen in Part 2 of this chapter.

2- *The qualitative analysis*

After ranking the list of bigrams, or the lexico-grammatical patterns, for each of the nature terms as seen above in Figure 38, each of them was checked against their original textual environments and allocated a SP for each contextual environment (as in Sinclair, 1996; 1998; Stubbs, 2001; Partington, 2004b). The six datasets of pre-processed text were each run in Sketch Engine to generate the concordance lines of the 30 most frequent natural phenomena in the Qur'an. However, it should be mentioned that in exploring the lexico-grammatical patterns of nature in the Qur'an, the use of Sketch Engine was only in applying the concordance function, and the Wordsketch function was not used for two reasons. Firstly, the data that was run to generate concordance lines based on the Log-Likelihood association measure was the pre-processed texts, which have the nature roots in Arabic, the English stems representing natural phenomena, and no function words or punctuation. This meant that a Wordsketch, although very sophisticated and visually useful in discussing collocates, was not necessary for the fact that the uncovering of SP meanings of nature in this research requires a look back at the raw corpora for more evidence via alignment. Secondly, this research employs *GraphColl* in LancsBox, which can perform a similar task to that of Wordsketch but enhances the view of nodes and collocates to be a network of collocations that is connected through common denominators, namely collocate sets.

Once they were marked as nodes with their collocates, the LG patterns were examined qualitatively in line with Sinclair's (2003) seven-step procedure for reading concordances to uncover the evaluative and discourse prosody of each of the terms by aligning each concordance line with its original verse (See Figures 30 and 31 in the previous section). Exegeses (Tafsīr) were consulted to ensure that the researcher's perception (judgment) of the interpretation of the contextual environment in which we find the co-occurrence is accurate.

erse	Tafseer	People	Day	Earth	Sky	Rain	Worlds	Village	Mounta	River	Women	Light	Chest
167] لئن لم يكن من ربنا رزقنا لكاننا من الخاسرين	[7:67] He said, "O my people, there is no folly in me, but I am a messenger from the Lord of the Worlds. [7:68] Or do you marvel that a reminder from your Lord should come to you through the tongue of a man from among you, that he may warn you? And remember when He made you vicegerents, on earth, after the people of Noah, and increased your stature in extension, in strength and in height — (the tallest of them measured 100 feet, the shortest, 60). Remember then God's bounties, His graces, so that you might prosper', triumph.						pos						
169] لئن لم يكن من ربنا رزقنا لكاننا من الخاسرين	[7:69] And, we sent, to Midian their brother Shu'ayb. He said, "O my people, worship God! You have no god other than Him. Verily there has come to you a clear proof, a miracle, from your Lord, to [prove] my truthfulness. So give full measure and weight and do not defraud, diminish [the value of], people's goods, and do not work corruption in the earth, by way of unbelief and acts of disobedience, after it has been set right, through the	pos											
7164] وَالْمَرْءُ ظَنَنَ أَنَّهْ يُلْقِيْهِمْ كِسْفًا مِّنْ سَائِلَاتٍ	[7:84] And We rained upon them a rain, the stones of baked clay (hijrat al-sajil), and it destroyed them. So behold what was the end of the sinners! [7:85] And, we sent, to Midian their brother Shu'ayb. He said, "O my people, worship God! You have no god other than Him. Verily there has come to you a clear proof, a miracle, from your Lord, to [prove] my truthfulness. So give full measure and weight and do not defraud, diminish [the value of], people's goods, and do not work corruption in the earth, by way of unbelief and acts of disobedience, after it has been set right, through the					neg							
	[7:85] And, we sent, to Midian their brother Shu'ayb. He said, "O my people, worship God! You have no god other than Him. Verily there has come to you a clear proof, a miracle, from your Lord, to [prove] my truthfulness. So give full measure and weight and do not defraud, diminish [the value of], people's goods, and do not work corruption in the earth, by way of unbelief and acts of disobedience, after it has been set right, through the												

Figure 39: The annotation of the evaluative prosodies of the natural phenomena

Using the resulting categories of statistically proven collocations and their lexicogrammatical patterns, each of the natural phenomena words was assigned various evaluative and discourse prosodies. By observing the lines of the concordance of each of the grammatical pattern (i.e., bigram), the nature terms were assigned the different SPs they hold as the text unfolds. As has been highlighted previously, the unit of meaning adopted in this research is Sinclair’s Lexical Extended Unit (2004a); that is, SP is assigned within the parameters of the lexical unit. This means that the annotation of verses or translated verses, including more than one natural phenomenon can be dual such as the annotated verse shown in the following figure:

Verse	.T Tafseer	People	Earth	Sky	Bigram	Dis. SP
40] 57] لَخَلْقِ السَّمٰوٰتِ وَالْاَرْضِ اَكْبَرُ مِنْ اَنْ يَّخْلُقَ الْاِنْسَانَ وَلَكِنَّ اَكْثَرَ النَّاسِ لَا يَعْلَمُوْنَ	[40:57] The following was revealed regarding the deniers of resurrection: Assuredly the creation of the heavens and the earth, as an [unprecedented] first act, is greater than the creation of mankind, a second time — which is the restoration [of them after death]; but most people, namely, the disbelievers of Mecca, do not know, this, and so they are like the blind, whereas those who know it are like those who have sight.	neg	pos	pos	(many,people),(creation,ec	disbelievers/emphasis on message

Figure 40: An example of dual annotation of discourse prosodies

In addition, a sample of the SP annotated datasets is provided in the figure below.

D	E	AV	AW
Verse	Tafseer	Bigram	Dis. SP
50 43	بَلْ نَحْنُ نُحْيِي وَنُمِيتُ وَآلَيْنَا الْحَيَاتِ [50:43] Indeed it is We Who give life and brin	(life,death)	glorifying of God
50 44	وَالْيَوْمَ نَخْلِفُ الْاَرْضَ لَكُمْ وَنَحْنُ الْوَارِثُونَ [50:44] On the day when (yawma substitutes for the previous yawma, with the intervening [statement] being a parenthetical) the earth is split asunder (read tashaqqaqu, or tashshaqqaqu, where the original second tā' [of tatashaqqaqu] has been assimilated with it [the shin]) from them, [they will come] hastening forth (sirā'an: [sirā'] the plural of sarī', a circumstantial qualifier referring to	(earth,day)	horror of judgment day
51 22	وَقِيلَ السَّعْيُ وَنَحْنُ نَزَّلْنَا السَّمَانَ [51:22] And in the heaven is your provision, that is, the rain from which results the vegetation that is [your] provision, and	(sustenance,heavens)	glorifying of God/emphasis on message
51 23	وَقِيلَ السَّعْيُ وَنَحْنُ نَزَّلْنَا السَّمَانَ [51:23] So by the Lord of the heaven and the earth, it, that which you are promised, is as assuredly true as [the fact] that you have [power of] speech (read mithlu, in the	(heavens,earth)	glorifying of God/emphasis on message
51 46	وَقَوْمَ نُوحٍ مِنْ قَبْلُ إِنَّهُمْ كَانُوا قَوْمًا فَاسِقِينَ [51:46] And the people of Noah (read wa qawmi Nūhin, as a supplement to Thamūda, in other words: 'in the destruction of these [two peoples] by what [destructive power]	(noah,people)	punishment 1

Figure 41: The annotation of the discourse prosodies of nature in the Qur'an in Arabic

The qualitative analysis developed for the categorisation system contains thirteen content categories derived from contextual environments in which natural phenomena occur in the Qur'an (Hsieh and Shannon, 2005, p.1277). An example of a concept related to the theme of nature is the word *signs*, which refers to various phenomena, ranging from the universe, its creation, the alternation between day and night, and rainfall, the life and growth of plants. We find other references to miracles, the rewards of belief and the fate of unbelievers. The following are Arabic examples with some of their translations which illustrate a diverse number of occurrences of the word *signs* with natural phenomena in the Qur'an²²⁹:

وَمِنْ آيَاتِهِ خَلْقَ السَّمَوَاتِ وَالْأَرْضِ وَمَا بَيْنَهُمَا مِنْ دَابَّةٍ وَهُوَ عَلَىٰ جَمْعِهِمْ إِذَا يَشَاءُ قَدِيرٌ ﴿٢٩﴾

Arberry: And of His signs is the creation of the heavens and earth and the crawling things He has scattered abroad in them, and He is able to gather them whenever He will. Verse (42:29)

وَآيَةٌ لَهُمُ الْأَرْضُ الْمَيْتَةُ أَحْيَيْنَاهَا وَأَخْرَجْنَا مِنْهَا حَبًّا فَمِنْهَا يَأْكُلُونَ ﴿٣٣﴾

Yusuf Ali: A Sign for them is the earth that is dead: We do give it life, and produce grain therefrom, of which ye do eat.

Verse (36:33)

229 <http://corpus.quran.com/>

تَجْرِي بِأَعْيُنِنَا جَزَاءَ لِمَنْ كَانَ كُفِرًا ﴿١٤﴾ وَلَقَدْ تَرَكْنَاهَا آيَةً فَهَلْ مِنْ مُدْرِكٍ ﴿١٥﴾

Saheeh International: Sailing under Our observation as a reward for he who had been denied. And We left it as a sign, so is there any who will remember?

Verses (54:14-15)

إِذْ قَالَ اللَّهُ يُعِيسَى ابْنَ مَرْيَمَ اذْكُرْ نِعْمَتِي عَلَيْكَ وَعَلَىٰ وَالِدَتِكَ إِذْ أَيَّدْنَاكَ بِرُوحِ
الْقُدُسِ تُكَلِّمُ النَّاسَ فِي الْمَهْدِ وَكَهْلًا وَإِذْ عَلَّمْنَاكَ مَا تَشَاءُ مِنَ الْحِكْمَةِ
وَالتَّوْرَةِ وَالْإِنْجِيلِ وَإِذْ تَخْلُقُ مِنَ الطِّينِ كَهَيْئَةِ الطَّيْرِ بِإِذْنِي فَتَنفُخُ فِيهَا فَتَكُونُ
طَيْرًا بِإِذْنِي وَتُتْرَقُ إِلَىٰ آلِكَمَّةٍ وَالْأَبْرَصِ بِإِذْنِي وَإِذْ تُخْرِجُ الْمَوْتَىٰ بِإِذْنِي وَإِذْ
كَفَفْتُ بِحِجِّ إِسْرَائِيلَ عَنْكَ إِذْ جِئْتَهُمْ بِالْبَيِّنَاتِ فَقَالَ الَّذِينَ كَفَرُوا مِنْهُمْ إِنْ
هَذَا إِلَّا سِحْرٌ مُؤْتَمِرٌ ﴿١٥﴾

Saheeh International: (The Day) when Allah will say, ‘O Jesus, Son of Mary, remember My favour upon you and upon your mother when I supported you with the Pure Spirit and you spoke to the people in the cradle and maturity; and (remember) when I taught you writing and wisdom and the Torah and the Gospel; and when you designed from clay (what was) like the form of a bird with My permission, then you breathed into it, and it became a bird with My permission; and you healed the blind and the leper with My permission; and when you brought forth the dead with My permission; and when I restrained the Children of Israel from (killing) you when you came to them with clear proofs and those who disbelieved among them said, ‘This is not but obvious magic.’ Verse (5:110)

To perform a qualitative analysis of nature in the Qur’an, two steps were followed: the first was to organise all the recurrent collocates of the thirty most frequent natural phenomena per their lexico-grammatical patterns by coding, using the annotations on the Excel sheets. In the second step, the collocations of nature terms were ascribed to different categories, as follows:

- Description of believers
- Shunning of disbelievers
- An emphasis on the message
- Description of evil doings of Satan
- Praising and glorifying of God
- Description of the horror of judgment day
- Description of hypocrites
- Islamic teaching
- Part of a miracle story
- Punishment in the present life (present-life punishment)
- The punishment in the afterlife (afterlife punishment)
- Description of transgressors
- Description of wrongdoers

The resulting categories were the discourse prosodies of natural phenomena in the Qur'an (See Section 5.1. for results).

In summary, the first stage of this research produces the following: statistical information about the natural phenomena in the Qur'an and its five translations and datasets of annotated SP of nature terms in the Qur'an and its five translations. These results, which will be reported and discussed in further detail in the following chapter, are the basis of the comparison of the translation corpus-based methodology in the next section of this chapter.

4.2 Part 2: The evaluation of the Qur'an translations via SP of nature terms

This section describes the second stage of this research in which the datasets of the previous stage were used to evaluate the translations of the Qur'an via evaluative and discourse prosodies as tools to contribute to *accuracy* and *consistency* in the translation of the Qur'an. It is the part of the methodology that involved the qualitative analysis in evaluating the translations in regard to being the most congruent in their representation of the theme of nature in the Qur'an. In addition, it is the section that describes the translated datasets following Stewart's (2000) approach to analysing translated texts, aimed to identify the difference in the translations but not to judge them. In other words, it only describes the differences in the representation of SP of natural phenomena and the results show the representation closest to the Arabic, not the most accurate in translation. In essence, it is not within the scope of this study to provide a comprehensive linguistic evaluation of the translations of the Qur'an or prescribe the best translation/s to be used by the reader of the Qur'an. One contribution of this study, however, is that it demonstrates the importance of an awareness of the analysis of SP as a tool for achieving better accuracy and consistency in translation. Therefore, this part of the research addresses the following questions:

- 1- Which of the five translations is most congruent, and which is the most divergent from the representation of nature in the Qur'an?
- 2- How can variances of the representation of SP of nature in the English renderings of the Qur'an be justified in terms of consistency and accuracy?

In sum, this section provides an overview of the steps taken to analyse the datasets of the English renderings of the Qur'an; references will be made to three components of this

corpus-based study of the parallel corpora of the Qur'an (frequencies, collocation and LG, and SP).

4.2.1 Why semantic prosody?

This subsection presents the rationale behind the selection of the analysis of SP, which is a tool to achieve both accuracy and consistency in comparing and evaluating the translations in this research. SP was chosen as the linguistic feature employed to compare the translations of the Qur'an, for it is the starting point for a writer/translator's choice of a lexical item. Before a writer or speaker chooses his/her lexical items to express meaning, he/she should first choose the SP of their words.

The selection of the item is controlled by the prosody, because the whole point of expressing oneself in this way is to pre-evaluate the actions, which would otherwise be evaluated positively by the reader/listener (Sinclair, 2004a, p. 175).

From the point of view of the speaker/writer, the textual process of constructing a lexical item is described as follows:

1. First, the speaker/writer selects semantic prosody of x applied to a semantic preference y .
2. The semantic preference, in turn, controls the collocational and colligational patterns.
3. The final component of the lexical item is the (invariable) core (Sinclair, 2004b, p. 34).

Furthermore, in the same context, Sinclair (2004b) asserts that “the initial choice of semantic prosody is the functional choice which links meaning to purpose”, and that “all subsequent choices within the lexical item relate back to the prosody” (p.34). He gives an example of the pattern “the naked eye”, which is explored as a lexical item and described as follows:

The speaker/writer selects the prosody of difficulty applied to a semantic preference of visibility. The semantic preference controls the collocational and colligational patterns and is divided into verbs, typically *see*, and adjectives, typically *visible*. With *see*, etc., there is a strong colligation with modals — particularly *can*, *could* in the expression of difficulty— and with the preposition *with* to link with the final segment. With *visible*, etc., the pattern of collocation is principally with degree adverbs and the negative morpheme *in-*; the following preposition is *to*. The final component of the item is the ‘core’, the almost invariable phrase *the naked eye* (Sinclair, 2004b, p. 34, italics in original).

It can be suggested that it is of considerable significance that the translator is familiar with the overall semantic prosody of the lexical items in the ST in order to achieve more accuracy. In addition to accuracy, consistency is another criterion for the evaluation of a translation. In the exploration of SP, the concordance gives an impression of group cohesion and textual unity and may lead to “perceived, and generalised, semantic consistency” (Stewart, 2010, pp.108–13). We might infer accordingly that the use of SP as a device for achieving accuracy and consistency can also be implemented in the evaluation of translated texts. Therefore, this research first introduces the SP of nature as a significant theme in the Qur’an and then presents a comparative/evaluative study of the English renderings of the Qur’an.

4.2.2 The evaluation of the five translations

The datasets were explored further to compare them to the Arabic source. An overview of the general statistics of the tagged datasets via LanksBox is shown in the following table:

Table 24: Statistical overview of datasets in this research

<i>Corpus</i>	<i>Tokens</i>	<i>Types</i>
Arabic	78253	15519
Pickthall	58988	5856
Ali	63688	6268
Arberry	54706	5195
Saheeh	57428	5084
Haleem	60280	5544

Moreover, just as in Mason’s (2001), Kenny’s (2006), and Munday’s (2014) methods of source-target text (ST-TT) analysis and evaluation, several linguistic features of the source text were used to compare and evaluate the translations of the Qur’an. Accordingly, the 30 most frequent natural phenomena terms in the Arabic text were explored in the five translations, and their collocations and SPs were analysed. The following table shows the translations of the natural phenomena as they appear in the five translations of the Qur’an.

Table 25: The translations of nature terms in the five selected translations

<i>Arabic Words</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
ناس-أهل -	people/mankind/folk	people	people	people	people
قوم-أقوام					
يوم-أيام-	Day	day	day	day	day
يومين					
أرض	earth/s	earth/s	earth	earth/s	earth/s

Arabic Words	Pickthall	Ali	Arberry	Saheeh	Haleem
سمااء- سماوات	sky/s/heavens	sky/s/heavens	sky/s	sky/s/heavens	sky/s/heavens
نفس-أنفس- نفوس	soul/s	soul/s	soul/s	soul/s	soul/s
القيامة-البعث	resurrection	resurrection	resurrection	resurrection	resurrection
النار-جهنم	hell/fire	hell/fire	hell/fire	hell/fire	hell
حياة-الدنيا	life	life	life of the world/life	life	life
رجل-رجال- إنسان	man/human/men	man/human/men	angel/s	man/human/men	man/human/men
جنة-جنات	heaven/the garden	heaven/the garden	heaven/the garden	heaven/the garden	heaven/the garden
قلب-قلوب	heart/s	heart/s	heart/s	heart/s	heart/s
يد-يدين- شمال	hand/s	hand/s	hand/s	hand/s	hand/s
إنس-بشر	mankind	mankind	face/s	mankind	mankind
شيطان- شياطين	devil/s	devil/s	satins/s	devil/s	devil/s
ملك-ملائكة- الصفات- الزجرات- التاليات- المقسمات- النازعات- الناشطات- السابحات- السابقات- المدبرات	angel/s	angel/s	angel/s	angel/s	angel/s
ليلة-ليل-ليال	night/s	night/s	night/s	night/s	night/s
بيوت-بيت- مساكن	homes/houses	homes/houses	homes/houses	homes/houses	homes/houses
موت	death	death	death	death	death
وجه-وجوه	face/s	face/s	face/s	face/s	face/s
ماء	water/rain	water/rain	water/rain	water/rain	water/rain
العالمين	worlds	worlds	all beings	worlds	worlds
قرية-قرى	township/s	town	city/s	village/town	village/town
جبل-جبال- رواسي-أعلام	hill/s/mountain/s	mountain/s	mountain/hills	mountain/hills	mountain
أنهار	river/s	river/s	rivers	river/s/streams	river/s
نساء	women	women	women	women	women
نور-ضياء- مصابيح	light	light	light	light	light
صدر-صدور	chest/s	chest/s/breast/s	chest/s	chest/s/breast/s	chest/s/breast/s
بحر-بحرين	sea /s	sea /s/ocean /s	sea /s	sea /s	sea /s
ثمره-ثمرات- فاكهة-فواكه	fruit /s	fruit /s	fruit /s	fruit /s	fruit /s
شمس-سراجا	sun	sun	sun	sun	sun
أنعام	cattle	cattle	cattle	cattle	cattle
جوار-سفينة- فلك	ship /s	ark	ship /s/ark	ship /s	ship /s/ark
شجر-شجرة	tree /s	tree /s	tree /s	tree /s	tree /s
جن-جان	jinn	jinn	jinn	jinn	jinn
قمر	moon	moon	moon	moon	moon
سنة-سنيين	year/s	year/s	year/s	year/s	year/s
عيون-عينا- سلسبيل	water springs	spring /s/fountain/s	spring/s/fountain /s	spring/s/fountain /s	springs /fountain

<i>Arabic Words</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
ريح-رياح-عاصفات-ناشرات-فارقا-ملقيات	wind/s	wind/s	wind/s	wind/s	wind /s
أغرق-غرق	drowning /flood	flood	drowning /flood	drowning	drowning
ظلمات-غاسق	darkness	darkness	darkness /shadows	darkness	darkness
نخيل-نخل	date-palm/s	date-fruit /s	palm tree/s	date-palm /s	date-palm/s
نجم-نجوم	star/s	star/s	star/s	star/s	star/s

In the comparison of frequencies of nature phenomena terms in the Qur'an and its five translations as seen in Table 26 below, this research followed McEnery and Wilson's approach (2001) in quantifying data and comparing frequency counts in different corpora.

Table 26: The frequency counts of nature terms in the Qur'an and its translations

<i>Nature Term</i>	<i>Arabic</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
People	534	238	511	521	732	785
Day	481	519	520	462	497	440
earth(s)	462	393	418	388	403	367
sky(s)/heavens	324	247	247	191	261	276
soul(s)	295	116	176	122	105	114
resurrection	276	77	20	70	84	74
Hell	255	94	81	24	82	91
Life	255	140	213	122	161	203
man/human/men	151	354	574	497	206	244
heaven/the garden	144	139	129	123	140	136
heart(s)	142	130	161	132	124	131
hand(s)	113	135	138	126	107	97
Mankind	107	209	72	37	62	23
devil(s)	103	54	2	0	27	7
angel(s)	95	98	100	92	126	96
night(s)	93	106	109	96	105	108
homes/houses	77	112	130	71	120	144
Death	71	76	80	43	88	72
face(s)	69	58	79	79	65	83
water/rain	61	76	69	76	54	83
Worlds	61	44	49	0	71	31
village/town	55	6	17	3	3	47
mountain	53	24	45	47	55	59
river(s)	53	58	55	56	59	18
Women	49	109	98	101	103	91
Light	47	65	78	57	52	52
chest(s)/breast(s)	44	24	14	33	44	5
Sea	39	44	42	46	46	43
fruit	35	37	60	38	52	34
sun	34	35	37	36	35	34
Cattle	32	33	34	32	7	9
Ship	32	33	35	35	36	32
tree (s)	29	28	30	25	49	40
jinn(s)	27	31	32	27	33	37

<i>Nature Term</i>	<i>Arabic</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
Moon	27	27	28	27	28	26
year(s)	35	36	36	26	32	26
springs	20	33	32	25	26	26
wind (s)	23	29	30	27	31	31
drowning	23	26	30	26	25	22
darkness	23	25	30	25	32	28
datepalm(s)	20	11	8	11	21	14
star(s)	13	15	13	14	19	16

In this regard, McEnery and Wilson argue that although simple frequency counts have often been used in corpus-based research as an approach to quantifying data, they have certain disadvantages. The main one arises when one wishes to compare one dataset with another. They state that “arithmetical frequency counts simply count occurrences”; they do *not* show the prevalence of a type in terms of the proportion of the total number of tokens within the text (2001, pp.82-3). Therefore, since the sample sizes (i.e., number of tokens) of the corpora in this research are different as seen previously in Table 24, it is essential to “normalise the data using some indicator of proportion” as follows:

Ratio²³⁰ = number of occurrences of the type/number of tokens in entire sample*(base of normalisation) (As in McEnery and Wilson, 2001, p. 83 and McEnery and Hardie, 2012, p.49)

Accordingly, to compare the frequencies in this research, the counts of a natural phenomenon term in Table 26 are divided by the number of tokens for each corpus as in the figures in Table 24 and multiplied by 1,000 in this research (as in Gries, 2010b, p.7). Once these ratios of natural phenomena terms in the English translations are calculated, they are compared to the ratios of the ST²³¹ to indicate which of the translations is closest in its representation of the proportion of the most frequent nature terms in the Qur’an.

Moreover, alignment is used to perceive collocations and SP as linguistic features of the analysis in the source text together with its translations. Simply put, *alignment* means linking a unit of text in one language with a unit of text in another language.

The applications of parallel corpora include comparing the lexis or grammar of different languages, looking at the linguistic features of translated texts. For many of these purposes an important first step in processing the parallel corpus is alignment (Baker et al., 2006, p. 127).

²³⁰ Also called *normalised frequency* (McEnery and Hardie, 2012,49) and *relative frequency* (RF) (Brezina, 2018, p.43).

²³¹ This refers to the source text, which is the Arabic Qur’an.

In addition, in this corpus-linguistic study of parallel corpora, alignment was used to derive the translation equivalents of nature terms. To do so, the compiled dataset of verses of the Qur'an with the 30 most frequent natural phenomena terms were aligned with their translations. Then, each word referring to nature in the Qur'an is identified with its five renderings in the chosen Qur'an translations in this research (See also Table 26). In this regard, there are software tools that assist in this process, and extensive research has been conducted in the fields of computational linguistics and corpus linguistics to improve alignment and methods (Olohan, 2004, p.26 and p. 55). This research employed Excel as well as LancsBox for a visual representation of the alignment of the verses and concordances of natural phenomena in the source text and target texts.

The image shows two screenshots of the LancsBox v 4.0 interface. The top screenshot shows a search for the word 'moon' with 27 occurrences (5.68). The bottom screenshot shows a search for the word 'moon' with 28 occurrences (4.40). Both screenshots display a table with columns for Index, File, Left, Node, and Right. The 'Node' column contains the word 'moon' in various contexts, and the 'Right' column shows the corresponding Arabic text from the Qur'an.

Index	File	Left	Node	Right
1	Arabic appended and without stop:	رَأَى الْوَيْلَ لِمَنْ يَلْمِ	هجر	رَأَى الْوَيْلَ لِمَنْ يَلْمِ
2	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
3	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
4	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
5	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
6	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
7	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
8	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
9	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
10	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
11	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
12	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
13	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ
14	Arabic appended and without stop:	نَحْمِيكَ لِي وَنَحْمِيكَ	هجر	نَحْمِيكَ لِي وَنَحْمِيكَ

Figure 42: Alignment of concordances of the word *moon* via LancsBox v 4.0

In addition, this research most importantly aligned the evaluative and discourse prosodies of the 30 most frequent natural phenomena in Arabic with their counterparts in the translations to compare them for their correspondence. Both congruence and divergence with the Arabic SP representation of nature terms were marked (as in Ebeling, 2014), and the translations were ranked in accordance with their congruency scores. The figure below shows the discrepancies of the representation of SP meanings of nature in the Qur'an (orange cells). The number of terms included in this evaluation was 42, and the divergence was deducted from the total number of terms to show the congruency score of SP in the five translations of the Qur'an.

A	B	C	D	E	F	G	H	I	J	K
	Arabic		Pickthall		Ali		Arberry		Sahih	
Nature Term	Arabic Dis SP	Arabic Ev.SP	Dis. SP	Ev. SP	Ali Dis SP	Ali Ev SP	Arberry Dis SP	Arberry Ev SP	Sahih Dis SP	Sahih Ev S
people	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos/neg	disbelievers	neg
day	punishment 2	neg	emphasis on message	pos	punishment 2	neg	punishment 2	neg	punishment 2	neg
earth(s)	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos
sky(s)/heavens	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos
soul(s)	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos
resurrection	punishment 2	neg	emphasis on message	pos	punishment 2	neg	punishment 2	neg	punishment 2	neg
hell	punishment 2	neg	punishment 2	neg	punishment 2	neg	punishment 2	neg	punishment 2	neg
life	emphasis on message	pos	disbelievers	neg	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos
man/human/men	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	emphasis on message	pos	glorifying of God	pos
heaven/the garden	reward	pos	reward	pos	reward	pos	reward	pos	reward	pos
hear(s)	disbelievers	neg	disbelievers	neg	disbelievers	neg	disbelievers	neg	disbelievers	neg
hand(s)	miracle story	pos	miracle story	pos	reward	pos	miracle story	pos	miracle story	pos
marking	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	glorifying of God	pos
devil(s)	evil doings of Satan	neg	evil doings of Satan	neg	evil doings of Satan	neg	evil doings of Satan	neg	evil doings of Satan	neg
angel(s)	miracle story	pos	glorifying of God	pos	emphasis on message	pos	miracle story	pos	emphasis on message	pos
night(s)	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos
homes/houses	Islamic teaching	neu	punishment 2	neg	reward	pos	glorifying of God	pos	reward	pos
death	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	emphasis on message	pos	glorifying of God	pos
face(s)	punishment 2	neg	punishment 2	neg	punishment 2	neg	punishment 2	neg	emphasis on message	pos
water/rain	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	punishment 2	neg
worlds	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos
village/town	punishment 1	neg	none	neu	none	neu	none	neu	none	neu
mountain	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos
river(s)	reward	pos	reward	pos	reward	pos	reward	pos	reward	pos
women	reward	pos	believers	pos	Islamic teaching	neu	believers	pos	reward	pos
light	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	glorifying of God	pos	emphasis on message	pos
chest(s)/breast(s)	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos	glorifying of God	pos

Figure 43: The comparison of the SP of nature terms between the Qur’an and its translations

4.3 Conclusion

This chapter described the quantitative and qualitative analyses to explore SP of natural phenomena in the Qur’an and to use the results of the latter to evaluate five translations of the Qur’an. By using SP as a medium for the study of collocational behaviour of nature in the Qur’an, this research mapped the evaluative and discourse prosodies of nature in the translations of the Qur’an.

To conclude, this chapter produced the following datasets for the Arabic Qur’an and five translations:

- 1- The 30 most frequent natural phenomena in Arabic and their equivalents in each of the five translations;
- 2- Bigrams of nature in the Qur’an produced from the corpora used in this research;
- 3- The evaluative and discourse prosodies of SP of natural phenomena in all of the corpora;
- 4- The evaluation of the most congruent translation in its representation of nature as a theme in the Qur’an.

Chapter 5 Results and Discussion

Introduction

This chapter presents the findings of the exploration of nature in the Qur'an reached through the analysis of the collocational phenomenon of SP in the Arabic text and five of its translations. Furthermore, it reports the findings of the evaluation of the five translations regarding their extent of congruency with the Arabic representation of SP of natural phenomena in the Qur'an. Using statistical and linguistic features revealed in the representation of nature as a theme, the four sections of this chapter first present the SP of the theme of nature in the Qur'an and then compare it to its counterparts in the English translations. The first section presents the results of the corpus-based analysis of the SP of natural phenomena in the Qur'an. It reports examples of frequencies of nature terms, examples of collocation (i.e., bigrams) and their functional lexico-grammatical patterns, and the evaluative and discourse prosodies of nature terms. The second section reports the results of the evaluation of the five translations of the Qur'an regarding their representation of the SP of nature in the Qur'an. It will describe the results of comparing the translated texts to the Arabic text. The third section discusses the results of the interpretation of both the quantitative and qualitative data analyses in light of the questions posed in this research and previous literature. The fourth section provides a summary of this chapter. In reporting and discussing the results of this research, the first three sections will focus on the flow of tasks designed for this corpus-based research exploring SP of nature in the Qur'an and its translation (see below Figure 22, which is the repeated Figure 24 from Chapter Four).

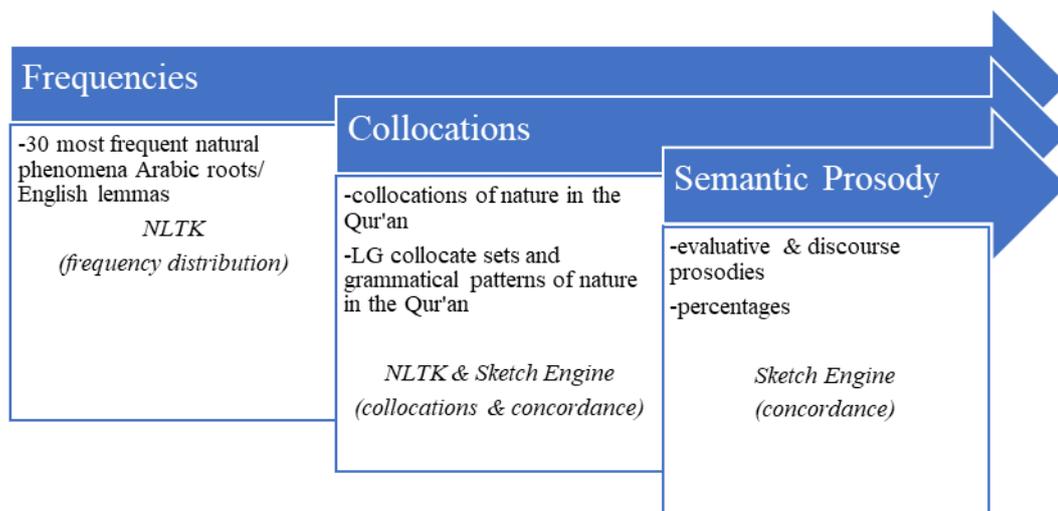


Figure 44: Flow of reporting results (repeated Figure 24)

5.1 Results of exploring the SP of natural phenomena in the Qur'an

This section presents the *frequencies*, *collocations*, and the *SP of nature terms* in the Arabic Qur'an in three sub-sections. It shows several tables which generally display the results of each task followed by two examples of each to illustrate these findings since not all of the results can be included in this chapter of the thesis.²³²

5.1.1 Frequencies of natural phenomena in the Qur'an

The first set of results is in Table 27, which provides a summary of the results related to the list of natural phenomena used in this research. It shows the Arabic words and disambiguated roots; transliterations of roots and English equivalents; and, most importantly, their observed and relative frequencies.

Table 27: The 30 most frequent natural phenomena in the Qur'an

Arabic Words	Disambiguated Root	Transliteration	Nature Term	Observed Freq	Relative Freq
ناس-أهل-قوم-أقوام	قوم2	qwm2	people	534	6.84
يوم-أيام-يومين-يوم منذ-نهار	يوم	ywm	day	481	6.15
أرض-خسف	أرض	rd	earth/s	462	5.90
سما-سماوات	سمو	smw	sky/s/heaven/s	324	4.14
نفس-أنفس-نفوس	نفس1	nfs	soul/s	295	3.77
القيامة-البعث	قوم1	qwm 1	resurrection	276	3.53
النار-جهنم	نور2	nwr 2	hell	255	3.26
حياة-الدنيا	حيي1	hyy	life	255	3.26
			man/human/me		
رجل-رجال-إنسان	رجل1	rġl 1	n	151	1.93
جنة-جنات	جنن2	ġnn 2	garden/s	144	1.84
قلب-قلوب	قلب1	qlb 1	heart/s	142	1.81
يد-يمين-شمال	يدي	ydy	hand/s	113	1.44
إنس-بشر	بشر	bšr	mankind	107	1.37
شيطان-شياطين	شطن	štn	devil/s	103	1.32
ملك-ملائكة-الصفاء- الزجرات-التاليات- المقسمات-النازعات- الناشطات-الساحات- السائقات-المديرات	ملك1	mlk 1	angel/s	95	1.21
ليلة-ليل-ليل	ليل	lyl	night/s	93	1.19
بيوت-بيت-مساكن	سكن	skn	homes/houses	77	0.98
موت	موت	mwt	death	71	0.91
وجه-وجوه	وجه1	wġh 1	face/s	69	0.88
ماء	موه	mwh	water/rain	61	0.78
العالمين	علم1	'lm 1	worlds	61	0.78
قرية-قرى	قري	qry	village/town	55	0.70
جبل-جبال-رواسي-أعلام	جبل	ġbl	mountain	53	0.68
أنهار	نهر1	nhr	river/s	53	0.68
نساء	نسو	nsw	women	49	0.63
نور-ضياء-مصابيح	نور3	nwr 3	light	47	0.60
صدر-صدور	صدر1	šdr 1	chest/s/breast/s	44	0.56
بحر-بحرين-بحار	بحر	bħr	sea/s	39	0.50
ثمره-ثمرات-فاكهة-فواكه	ثمر	tmr	fruit	35	0.45
شمس-سراجا	شمس	šms	sun	34	0.43
أنعام	نعم	n 'm	cattle	32	0.41
جوار-سفينة-فلك	فلك	flk	ship	32	0.41
شجر-شجرة	شجر	šġr	tree/s	29	0.37

232 A detailed version of the results is available in the electronic appendices of this thesis.

Arabic Words	Disambiguated Root	Transliteration	Nature Term	Observed Freq	Relative Freq
جن-جان	جن3	ġnn 3	jinn	27	0.35
قمر	قمر	qmr	moon	27	0.35
سنة-سنين-عام--	سنو1	snw 1	year/s	35	0.45
عيون-عينا-سلسبيل	عين1	'yn 1	springs	20	0.26
ريح-رياح-عاصفات-	ريح	ryh	wind /s	23	0.29
ناشرات-فارقات-ملقيات	غرق	ġrq	drowning	23	0.29
أغرق-غرق	ظلم1	zlm 1	darkness	23	0.29
ظلمات-غاسق	نخل	nhl	date palm/s	20	0.26
نخيل-نخل	نجم	njm	star/s	13	0.17
نجم-نجوم					

As seen above, the table illustrates that the natural phenomenon with the highest frequency, 534 occurrences, in the Qur'an belongs to the term *people* (as in 'people of'). It takes several morphological, semantic, and syntactic forms (e.g., ناس-أهل-قوم-أقوام) and has been root-disambiguated on three levels (word-root-synonym)²³³ to be قوم2 *qwm*2. The second most frequent word is *day*, with 481 occurrences in the Qur'an. It takes several morphological and syntactic forms (e.g., يوم-أيام-يومين-يومئذ) and has been root-disambiguated on two levels (i.e., word-synonym) to be يوم *ywm*. In addition, the diagram below shows that most of the frequencies of natural phenomena in the Qur'an belong to living creatures, such as *people, man, trees, cattle*, etc., while a notably large number of occurrences belong to the category of astronomical bodies, such as *earth, sun, moon, star, sky*. Finally, the third-ranking category in terms of frequencies of nature terms in the Qur'an is temporal events, signified by words such as *day, night, year*, etc.

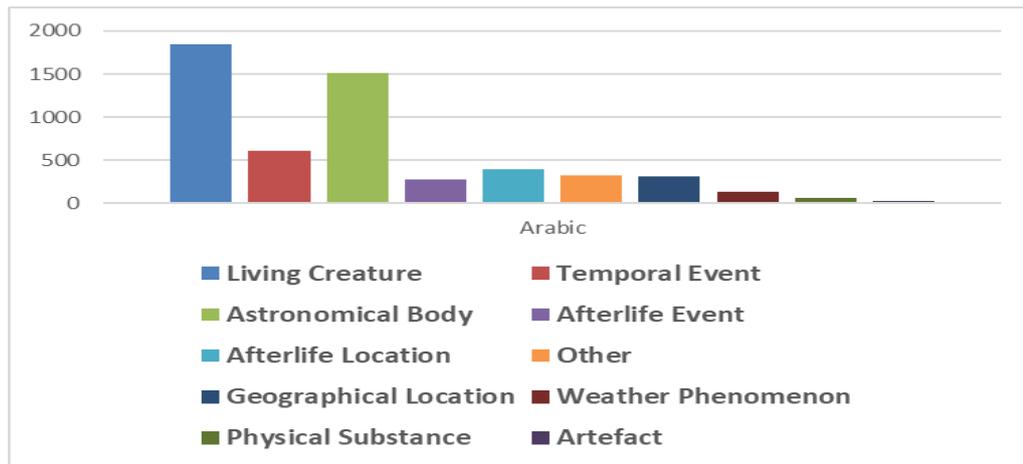


Figure 45: The semantic categories of natural phenomena in the Qur'an

233 As in Section 4.1.2.

5.1.2 Collocations of natural phenomena in the Qur'an

Collocations of nature in the Qur'an were computed by way of statistical analysis to obtain the most significant collocates; Table 28 and Figure 46 illustrate an example of results of the collocations of the word شمس *šams* 'sun', which occurs 34 times in the Qur'an.

Table 28: Example 1- the collocations of the node word شمس *šams* 'sun' in the Qur'an

Position	Collocate	Translation	Observed Frequency	Frequency Signature	Python LLR
R	قمر	moon	23	23_27_34_78253	354.62
L	نهر2	day	14	14_54_34_78253	161.74
L	ليل	night	15	15_93_34_78253	157.99
L	وَسَخَّرَ	subjected	9	9_11_34_78253	131.54
R	يُجْرِي	run	4	4_4_34_78253	62.42
R	نجم	star	5	5_13_34_78253	60.87
L	سَبَبًا	reason	3	3_4_34_78253	42.22
R	لِأَجَلٍ	time	3	3_4_34_78253	42.22
R	مُسَمًّى	appointed	4	4_21_34_78253	41.98
R	غرب2	sunset	3	3_5_34_78253	39.99
R	غرب1	west	3	3_11_34_78253	33.84
R	بِأَمْرِهِ	His command	3	3_12_34_78253	33.23
L	وَجَعَلَ	made	3	3_29_34_78253	27.45
R	نور3	light	3	3_47_34_78253	24.44
L	خَلَقَ	created	3	3_58_34_78253	23.15
L	سمو	sky	3	3_324_34_78253	12.91
R	قَالَ	he said	3	3_416_34_78253	11.48
L	أَرْضَ	earth	3	3_462_34_78253	10.88
L	الله	Allah	3	3_592_34_78253	9.5

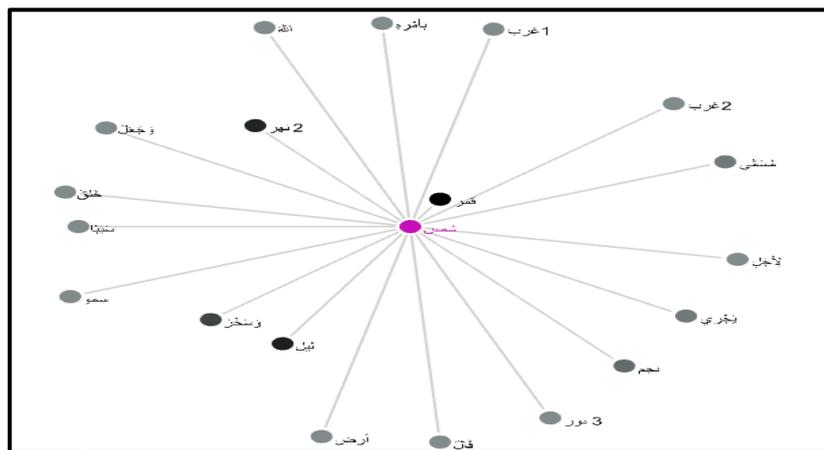


Figure 46: Collocates of شمس *šams* 'sun' in the Qur'an: 6a-LL (6.63), 5L-5R, C3-NC3; function words removed; ranked collocates considered

Figure 46 shows that the term 'sun' in the Quran is closest in its association to the collocate 'moon'. It is also attracted in similar proximity to the collocates 'made subservient', 'night', and 'day'. This type of affinity, when examined in the verses of the Quran reveals the

pragmatic meaning of glorifying God’s creation. In addition, a sample of the resulting dataset which shows three collocational co-occurrences of the word شمس *šams* ‘sun’ in the verses of the Qur’an and their corresponding English interpretation (Tafsīr) is shown in Table 29 (blue words are nodes, red words are the statistical collocates).

Table 29: Example 1- Verses with collocational co-occurrences of the node word شمس *šams* ‘sun’ in the Qur’an and their English interpretation

Verse	Tafsīr Aljalalyin
29 31 أَلَمْ تَرَ أَنَّ اللَّهَ يُولِجُ اللَّيْلَ فِي النَّهَارِ وَيُولِجُ النَّهَارَ فِي اللَّيْلِ وَسَخَّرَ الشَّمْسَ وَالْقَمَرَ كُلًّا يَجْرِي إِلَىٰ أَجَلٍ مُّسَمًّى وَأَنَّ اللَّهَ بِمَا تَعْمَلُونَ خَبِيرٌ	(31:29) Seest thou not that Allah merges Night into Day and he merges Day into Night; that He has subjected the sun, and the moon (to his Law), each running its course for a term appointed; and that Allah is well-acquainted with all that ye do? (35:13) He, God, makes the night pass, enter, into the day, so that it becomes longer, and He makes the day pass into the night, so that it becomes longer, and He has disposed the sun and the moon, each, of them, moving, in its course, to an appointed term — (to) the Day of Resurrection. That is God, your Lord; to Him belongs (all) sovereignty. As for those on whom you call, (whom) you worship, besides Him, in other words, other than Him — and they are the idols — they do not possess (even) so much as the husk of a date-stone.
13 35 يُُولِجُ اللَّيْلَ فِي النَّهَارِ وَيُولِجُ النَّهَارَ فِي اللَّيْلِ وَسَخَّرَ الشَّمْسَ وَالْقَمَرَ كُلًّا يَجْرِي لِأَجَلٍ مُّسَمًّى ذَلِكُمُ اللَّهُ رَبُّكُمْ لَهُ الْمُلْكُ وَالَّذِينَ تَدْعُونَ مِنْ دُونِهِ مَا يَمْلِكُونَ مِنْ قِطْمِيرٍ	(81:1) When the sun is folded away, enfolded and stripped of its light; (81:2) and when the stars scatter, (when) they are extinguished and hurtle down towards the earth.
81 1 إِذَا الشَّمْسُ كُوِّرَتْ 2 81 وَإِذَا النُّجُومُ انكَرَتْ	

Example 2 in Table 30 shows the nature term ماء *mā* ‘water’ as a more widely distributed term occurring 61 times. The table shows collocates with which it co-occurs in the Qur’an, followed by a *GraphColl* and another table illustrating its collocational co-occurrences in the verses and their English interpretations.

Table 30: Example 2- the collocations of the node word ماء *mā* ‘water’ in the Qur’an

Position	Collocate	Translation	Observed Frequency	Frequency Signature	Python LLR
L	سمو	sky	28	28_324_61_78253	225.85
R	أرض	earth	22	22_462_61_78254	147.55
L	أَنْزَلَ	send down	8	8_47_61_78255	72.77
R	فَأَخْرَجْنَا	raise	5	5_6_61_78256	66.58
R	فَأَنْبَتْنَا	grow	4	4_4_61_78257	57.52
R	حَيِي 1	life	8	8_139_61_78258	54.57
L	وَأَنْزَلَ	He sent down	4	4_11_61_78259	43.11
L	وَأَنْزَلْنَا	we sent down	4	4_12_61_78260	42.26
R	موت	death	5	5_71_61_78261	35.91
R	ثمر	fruit	4	4_35_61_78262	32.69
L	وَجَعَلْنَا	we made	4	4_37_61_78263	32.22
R	شرب	drink	4	4_38_61_78264	32
M	لِقَوْمٍ	people	4	4_50_61_78265	29.71
R	زرع	plant	3	3_12_61_78266	29.61
M	جبل	mountain	4	4_53_61_78267	29.23
L	خَلَقَ	created	4	4_58_61_78268	28.49
L	دبب	beast	3	3_15_61_78269	28.1

Table 31: Example 2- the collocational co-occurrences of the node word ماء mā' 'water' in the Qur'an and their English Interpretation

Verse	Tafsīr Aljalalyin
22 2 الَّذِي جَعَلَ لَكُمُ الْأَرْضَ فِرَاشًا وَالسَّمَاءَ بِنَاءً وَأَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجَ بِهِ مِنَ الثَّمَرَاتِ رِزْقًا لَكُمْ فَلَا تَجْعَلُوا لِلَّهِ أَنْدَادًا وَأَنْتُمْ تَعْلَمُونَ	(2:22) He Who assigned to you, created (for you), the earth for a couch, like a carpet that is laid out, neither extremely hard nor extremely soft to make it impossible to stand firm upon it; and heaven for an edifice, like a roof; and sent down from the heaven water, wherewith He brought forth, all types of, fruits for your provision; so set not up compeers to God, that is partners in worship, while you know that He is the Creator, that you create not and that only One that creates can be God.
164 2 إِنَّ فِي خَلْقِ السَّمَاوَاتِ وَالْأَرْضِ وَالْخِلَافِ اللَّيْلِ وَالنَّهَارِ وَالْفُلْكِ الَّتِي تَجْرِي فِي الْبَحْرِ بِمَا يَنْفَعُ النَّاسَ وَمَا أَنْزَلَ اللَّهُ مِنَ السَّمَاءِ مِنْ مَاءٍ فَأَخْرَجَ بِهِ الْأَرْضَ بَعْدَ مَوْتِهَا وَبَثَّ فِيهَا مِنْ كُلِّ دَابَّةٍ وَتَصْرِيفِ الرِّيَّاحِ وَالسَّحَابِ الْمُسَخَّرِ بَيْنَ السَّمَاءِ وَالْأَرْضِ لَآيَاتٍ لِقَوْمٍ يَعْقِلُونَ	(2:164) They then asked for a sign to prove this, and the following was revealed: Surely in the creation of the heavens and the earth, and the marvels contained in them, and the alternation of the night and day, passing and returning, increasing and diminishing, and the ships that run in the sea, and do not become cracked and sink, with what profits men, of trade and merchandise, and the water, the rain, God sends down from the heaven with which He revives the earth, with vegetation, after it is dead, after it has dried out, and He scatters abroad in it all manner of crawling thing, by dividing them and spreading them throughout on account of the vegetation, for they thrive on the fertile pastures it produces; and the disposition of the winds, changing it from south to north, from cold to warm, and the clouds compelled, subjugated by God's command, moving to wherever God wishes, between heaven and the earth, without being attached (to either of the two) — surely there are signs, indicating His Oneness, exalted be He, for a people who comprehend, (a people) who contemplate.
44 11 أَوْقِيلْ يَا أَرْضُ ابْلَعِي مَاءَكَ وَيَا سَمَاءَ أَطْلَعِي وَغِيضِ الْمَاءِ وَقُضِي الْأَمْرُ وَاسْتَوَتْ عَلَى الْجُودِيِّ وَقِيلَ بُعْدًا لِلْقَوْمِ الظَّالِمِينَ	(11:44) And it was said, 'O earth, swallow your waters, that have sprung forth from you — and it reabsorbed it (all) except for what came down from the sky and formed rivers and seas — and O heaven, abate!', withhold the rain, and it did. And the waters subsided, decreased. And the affair was accomplished, the matter of the destruction of Noah's people was complete, and it settled, (and) the ship came to rest, upon al-Jūdī, a mountain in Mesopotamia (al-jazīra, near Mosul; and it was said: 'Away with — perish — the evildoing, the disbelieving, folk!'
32 14 اللَّهُ الَّذِي خَلَقَ السَّمَاوَاتِ وَالْأَرْضِ وَأَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجَ بِهِ مِنَ الثَّمَرَاتِ رِزْقًا لَكُمْ وَسَخَّرَ لَكُمُ الْفُلْكَ لِتَجْرِيَ فِي الْبَحْرِ بِأَمْرِهِ وَسَخَّرَ لَكُمُ الْأَنْهَارَ	(14:32) God it is Who created the heavens and the earth, and He sends down water from the heaven, and with it He brings forth fruits as sustenance for you. And He has disposed for you the ships, that they may run upon the sea, (laden) with passengers and cargo, at His commandment, by His leave, and He has disposed for you the rivers.
16 10 هُوَ الَّذِي أَنْزَلَ مِنَ السَّمَاءِ مَاءً لَكُمْ مِنْهُ شَرَابٌ وَمِنْهُ شَجَرٌ فِيهِ تُسِيمُونَ	(16:10) He it is Who sends down water from the heaven, whence you have drink, for you to drink, and whence are trees, which grow because of this (water), whereat you let your animals graze.
65 16 أَوَلَا اللَّهُ أَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجَ بِهِ الْأَرْضَ بَعْدَ مَوْتِهَا إِنَّ فِي ذَلِكَ لَآيَةً لِقَوْمٍ يَسْمَعُونَ	(16:65) And God sends down water from the heaven and therewith revives the earth, with plants, after its death, (after) its drying up. Surely in that, which is mentioned, there is a sign, indicating the (truth of the) Resurrection, for a people who listen, listening and then reflecting.
54 11 فَفَتَحْنَا أَبْوَابَ السَّمَاءِ بِمَاءٍ مُثَمَرٍ	(54:11) Then We opened (read fa-fatahnā or fa-fattahnā) the gates of the heaven with torrential waters, (waters) pouring down intensely,
54 12 وَجَعَلْنَا الْأَرْضَ بُرُوجًا مَازِيَةً وَالسَّمَاءَ مِزَابًا مَازِيَةً	(54:12) and We made the earth burst forth with springs, that flowed forth, and the waters, the waters of the heaven and the earth, met for a purpose, a circumstance, that was preordained, (a matter) which had been decreed since pre-eternity, namely, their destruction by drowning.
78 14 وَأَنْزَلْنَا مِنَ الْمُعْصِرَاتِ مَاءً تَنْجَاتٍ	(78:14) and sent down from the rain-clouds (mu'sirāt), the clouds due to give rain (similar to (the term) mu'sir, which denotes a girl nearing menstruation) cascading water? pouring forth,
79 31 أَخْرَجَ مِنْهَا مَاءَهَا وَمَرْعَاهَا	(79:31) from it He has brought forth (akhraja: a circumstantial qualifier with a suppressed (preceding) qad, that is to say, mukhrijan, 'bringing forth (from it)') its waters, by making its springs gush forth, and its pastures, what cattle graze, of trees and herbage, and what humans consume of foods and fruits (the use of mar'ā to express this (of the earth) is figurative),
25 80 إِنَّا صَبَبْنَا الْمَاءَ صَبًّا	(80:25) that We pour down water, from the clouds, plenteously;

It should be mentioned, however, that the researcher found that the word *مطر* *maṭar* ‘rain’²³⁵ in the Arabic Quran is associated with the meaning of punishment in the present life for the irrevocable sinners; it is a historical sign or portent with a negative connotational colouring. For example, there are various mentions in the Quran of the sky raining stones on the punished or warned people (See figure below with a key for translations of collocates on the right).

rain	مطر
the warned	المنذرين
became worse	فساء
the punished	الغابرين
He will	يشاء
stones	حجر
sky	سمو

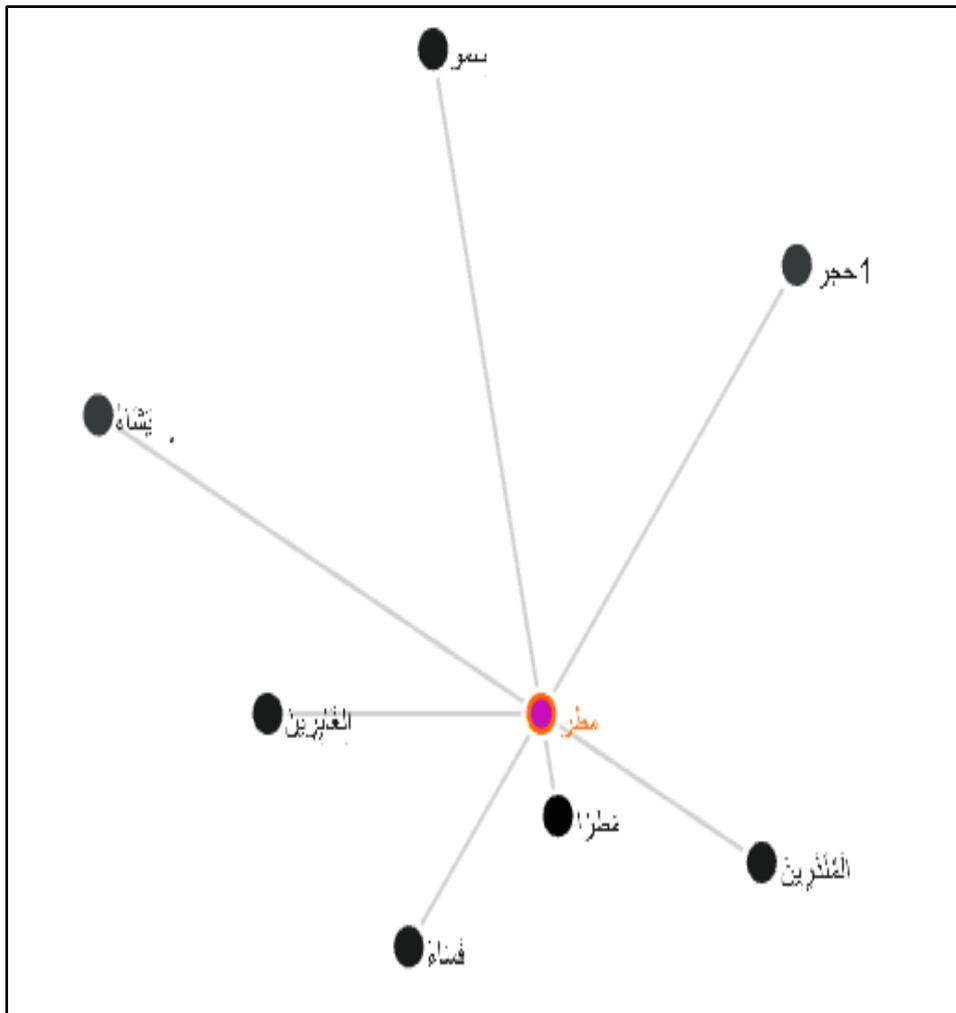


Figure 48: Collocates of *مطر* *maṭar* ‘rain’ in the Qur’an: 6a-LL (6.63), 5L-5R, C3-NC3; function words removed; ranked collocates considered

²³⁵ This word was not included in the list of the 30 most frequent natural phenomena, and only mentioned here for an illustrative purpose.

5.1.3 Lexico-Grammatical (LG) patterns of natural phenomena in the Qur'an

Another set of results that led to the portrayal of the hidden meanings of nature in the Qur'an is what Halliday and Matthiessen (2004) call the pragmatic functions²³⁶ of lexico-grammatical (LG) patterns. Firstly, the list of LG patterns has been identified from the computed bigrams (of natural phenomena in the Qur'an. By observing the patterns of each of the co-occurrences of nature terms with collocate, eight grammatical patterns have been extracted. They are as follows:

Noun (N)+Noun (N);

Noun (N) + Verb (V);

Verb(V)+ Noun (N);

Adjective (ADJ)+Noun (N);

Noun (N)+Adjective (ADJ);

Noun (N)+ Adverb (AV);

Adverb (AV)+ Noun (N)

To employ these patterns in eliciting the pragmatic functions (See Table 32) per Lexico-Grammatical functional theory (Halliday and Matthiessen, 2004), the researcher analysed them through the use of concordance in Sketch Engine (as in Louw, 1993). For each line where the nodes appear, a syntagmatic analysis has been conducted to elicit the collocate sets which have been categorised in accordance with the meanings that they 'transfer'²³⁷ (as in Hunston and Thompson, 2000) to the nodes as they repeatedly co-occur in the Qur'an. Table 32 below illustrates the pragmatic functions emerging in the analysis of the LG patterns and the collocates that indicate the transferred meanings, and Table 33 provides an example of the LG patterns and their pragmatic function (see also Appendix C for all LG patterns).

Table 32: LG pragmatic functions of natural phenomena and their relevant collocates

<i>LG Pragmatic Functions</i>	<i>Collocates</i>
believers	believe (v.) righteous (adj.) righteousness (n.) good (adj.)
disbelievers	rejection (n.) rejecter (n.) disbelief (n.) disbelieve (n.)

²³⁶ Referred to as 'discourse prosodies' in this thesis (in accordance with the SP definition of: Sinclair 1996, Louw, 2000; Stubbs 2001; 2004).

²³⁷ For this a *Tafsīr* was nonetheless consulted whenever was necessary.

<i>LG Pragmatic Functions</i>	<i>Collocates</i>
	(not) grateful (adj.)
emphasis on message	truth (n.) signs/s (n.) understand (v.) submit (v.) indeed (adv.) certainly (adv.)- verily (adv.) remind (v.) see (v.) witness (v.) remember (v.) say (addressed to Prophet Mohammed) (v.) invite (people to believe) (v.) light (to mean the message) (n.) darkness (to mean disbelief) worship (v.) promised (reward in the afterlife) guide -guidance warning faith
evil doings of Satan glorifying of God	evil (adj) praise (v.) glory (n.)-glorify (v.) dominion (n.) All-Mighty (name of God) (n.)- All-Wise (name of God) (n.) created (v.)-creation (n.)-Creator (n.) (He) made (v.) (God) knows (v.) subjected and subservient (natural phenomena are made subservient to man) (adj.) bow to God (v.) give/gave (v.) bounty/bounties (n.) grateful (adj.) spacious (earth) (adj.) establish (His throne) (v.) sustenance (from God) (n.) spread (earth) (v.)
the horror of day of judgment	split (heavens) (v.) raise (the resurrection of people on the day of judgment) (v.) come (to God on the day of judgment) (v.) doom (e.g., day of doom) (n.)
hypocrites	sealed (hearts) (adj.)- concealed (adj.) disease (hearts) (adj.)
Islamic teaching miracle story	pray (v.) white (hand) (adj) signs (miracles granted to prophets) (n.)
Present-life punishment	the warned (people) (adj.) earth swallow (v.) blast (n.) drown (v.) flood (n.)
afterlife punishment	fire (n.) abode (v.) enter (hell) (v.) penalty-punishment- punished (v.) black (face) (adj.) drink (scalding water as punishment in Hell) (v.) woe (n.)

<i>LG Pragmatic Functions</i>	<i>Collocates</i>
	companions of the Fire (n.)
reward	abode (v.) flow (v.) enter (garden) (v.) triumph (n.) bliss (n.) Eden (n.) grapevines (n.) reward (n.) drink (water from the springs in the Garden) (v.) companions of the Garden (n.) righteous -righteousness
transgressors	lost (adj.)
wrong doers	wrong (adj.) -wronged (v.)-mischief (n.)

Table 33: Samples of the LG patterns of nature in the Qur'an and their functions

<i>N+N</i>	<i>Pragmatic Function</i>	<i>N+V</i>	<i>Pragmatic Function</i>	<i>V+N</i>	<i>Pragmatic Function</i>
people people	disbeliever	people praise	glorifying of God	sent people	emphasis on message
signs people	emphasis on message	people say	disbelievers	made people	glorifying of God
truth people	emphasis on message	people disbelieve	disbelievers	gather day	emphasis on message
day rejecters	afterlife punishment	day know	glorifying of God	created day	glorifying of God
woe day	the horror of judgment day	day established	glorifying of God	knows day	glorifying of God
day faces	afterlife punishment	day judgement	emphasis on message	merge day	glorifying of God
day faces	reward	earth see	emphasis on message	made day	glorifying of God
earth day	glorifying of God	say earth	emphasis on message	created earth	glorifying of God
heavens earth	glorifying of God	earth assemble	the horror of judgment day	know earth	glorifying of God
mischief earth	wrongdoers	earth appointed	glorifying of God	made earth	glorifying of God
dominion earth	glorifying of God	sky made	glorifying of God	belong earth	glorifying of God
earth mountains	glorifying of God	heavens raise	glorifying of God	emerge earth	glorifying of God
earth mountains	the horror of judgment day	soul earned	emphasis on message	praise earth	glorifying of God
water earth	Present-life punishment	soul know	glorifying of God	swallow earth	Present-life punishment
<i>ADJ+N</i>	<i>Pragmatic Function</i>	<i>N+ADJ</i>	<i>Pragmatic Function</i>	<i>AV+N</i>	<i>Pragmatic Function</i>
many people	disbelievers	earth six	glorifying of God	verily heart	emphasis on message
grateful people	emphasis on message	day subservient	glorifying of God		
six days	glorifying of God	heavens six	glorifying of God		
merciful heavens	glorifying of God	heavens all	glorifying of God		
seven heavens	glorifying of God	soul good	believers		
wrong soul	wrongdoers	soul astray	transgressors		
lost soul	wrongdoers	soul one	glorifying of God		
punished soul	afterlife punishment	resurrection wrong	wrong doers		
merciful man	glorifying of God	life good	believers		

righteous garden	believers	heart sealed	hypocrites
active angel	glorifying of God	faces black	afterlife punishment
merciful angel	glorifying of God	rivers eternal	Reward
greatest angel	glorifying of God		
merciful night	glorifying of God		
<hr/>			
<i>N+AV</i>	<i>Pragmatic Function</i>		
<hr/>			
earth verily	emphasis on the message		
life verily	emphasis on message		
night verily	emphasis on message		
<hr/>			

5.1.4 The SP of natural phenomena in the Qur'an

After determining the bigrams of natural phenomena in the Qur'an and identifying the patterns and collocate sets of the LG patterns and their functions, the SP of natural phenomena in the Qur'an has been examined. Looking back at the pragmatic functions illustrated in the previous two tables, it is evident that the use of the concordance in Sketch Engine was effective in classifying the collocates of nature via content analysis (results of Section 4.1.3 of the methodology). The same principle was applied to find SP; that is, the pragmatic functions of the bigrams were used to refer to the discourse prosodies (following the definition of SP provided by Sinclair, 1996, 2004b; Louw, 2000; and Stubbs, 2001). This overlap between finding LG pragmatic functions via concordance and SP makes it safe to say that SP is uncovered by way of using concordance lines (findings of Section 4.1.4 of the methodology). In essence, the nature terms were statistically explored; their collocations were extracted and verified, and the concordance lines where the collocations occur were analysed to make a judgment of SP based on co-occurrence. The following step was to label each of the nature terms with the SP that best suits it in the context of the Qur'an and based on its occurrence (frequency) and co-occurrences (collocations) in the Qur'an. This was conducted via the calculation of the percentage of the evaluative and discourse prosodies in relation to their total frequency in the Qur'an using Microsoft Excel. The table below displays the results of the SP of the 30 most frequent nature terms in the Qur'an.

Table 34: The evaluative prosodies (EP) and discourse prosodies (DP) of nature in the Qur'an

<i>Arabic Words</i>	<i>Translated</i>	<i>Frequency</i>	<i>Pos</i>	<i>Neg</i>	<i>Neu</i>	<i>Pos%</i>	<i>Neg%</i>	<i>Neu%</i>	<i>EP</i>	<i>DP</i>
ناس-أهل-قوم-أقوام	people	534	67	66	2	12.5	12.4	0.4	pos	Emphasis on message
يوم-أيام-يومين	day	481	72	100	8	15	20.8	1.7	neg	Afterlife punishment
أرض-خسف	earth	462	220	35	2	47.6	7.6	0.4	pos	Glorifying of God
سماء-سماوات	sky/skies/heavens	324	200	9	0	61.7	2.8	0	pos	Glorifying of God

Arabic Words	Translated	Frequency	Pos	Neg	Neu	Pos%	Neg%	Neu%	EP	DP
نفس-أنفس-نفوس	soul/s	295	30	14	2	10.2	4.7	0.7	pos	Emphasis on message
القيامة-البعث	resurrection	276	33	48	0	12	17.4	0	neg	Afterlife punishment
النار-جهنم	hell	255	0	62	0	0	24.3	0	neg	Afterlife punishment
حياة-الدنيا	life	255	49	33	5	19.2	12.9	2	pos	Emphasis on message
رجل-رجال-إنسان	man/human/men	151	26	18	0	17.2	11.9	0	pos	Glorifying of God
جنة-جنتات	heaven/the garden	144	63	0	0	43.8	0	0	pos	Reward
قلب-قلوب	heart/s	142	24	24	5	16.9	16.9	3.5	pos/neg	emphasis on message/Disbelievers
يديمين-شمال	hand/s	113	9	1	1	8	0.9	0.9	pos	Miracle story
إنس-بشر	mankind	107	3	2	0	2.8	1.9	0	pos	Emphasis on message
شيطان-شياطين	devil/s	103	0	13	0	0	12.6	0	neg	Evil doings of Satan
ملك-ملائكة-الصفات-	angel/s	95	15	4	0	15.8	4.2	0	pos	Miracle story
الزاجرات-التاليات-										
المقسمات-النازعات-										
النشاطات-المباحات-										
السابقات-المدبرات										
ليلة-ليل-ليل	night/s	93	32	2	4	34.4	2.2	4.3	pos	Glorifying of God
بيوت-بيت-مساكن	homes/houses	77	1	0	2	1.3	0	2.6	neu	Islamic teaching
موت	death	71	20	1	0	28.2	1.4	0	pos	Glorifying of God
وجه-وجوه	face/s	69	0	9	0	0	13	0	neg	Afterlife punishment
ماء	water/rain	61	31	8	2	50.8	13.1	3.3	pos	Glorifying of God
العالمين	worlds/universe	61	45	0	0	73.8	0	0	pos	Glorifying of God
قرية-قرى	village/town	55	0	3	1	0	5.5	1.8	neg	Present-life punishment
جبل-جبال-رواسي-أعلام	mountain	53	12	4	0	22.6	7.5	0	pos	Glorifying of God
أنهار	river/s	53	43	0	0	81.1	0	0	pos	Reward
نساء	woman	49	5	4	3	10.2	8.2	6.1	pos	Reward
نور-ضياء-مصباح	light	47	25	4	0	53.2	8.5	0	pos	Emphasis on message
صدر-صدور	chest/s/breast/s	44	6	2	0	13.6	4.5	0	pos	Glorifying of God
بحر-بحرين	sea	39	16	2	1	41	5.1	2.6	pos	Glorifying of God
ثمر-ثمرات-فاكهة-فواكه	fruit	35	15	0	0	42.9	0	0	pos	Glorifying of God
شمس-سراجا	sun	34	16	1	0	47.1	2.9	0	pos	Glorifying of God
أنعام	cattle	32	3	0	0	9.4	0	0	pos	Glorifying of God
جوار-سفينة-فلك	ship	32	12	7	0	37.5	21.9	0	pos	Glorifying of God
شجر-شجرة	tree /s	29	6	2	0	20.7	6.9	0	pos	Glorifying of God
جن-جان	jinn	27	5	11	0	18.5	40.7	0	neg	Afterlife punishment
قمر	moon	27	17	1	0	63	3.7	0	pos	Glorifying of God
سنة-سنيين	year/s	35	1	0	0	2.9	0	0	pos	Emphasis on message
عيون-عينا-سلسبيل	springs	20	3	0	0	15	0	0	pos	reward
ريح-رياح-عاصفات-	wind /s	23	9	1	0	39.1	4.3	0	pos	Glorifying of God
ناشرات-فراق-ملاقات										
أغرق-غرق	drowning	23	0	4	0	0	17.4	0	neg	Present-life punishment
ظلمات-غاسق	darkness	23	0	9	0	0	39.1	0	neg	Disbelievers
نخيل-نخل	date palm/s	20	13	0	0	65	0	0	pos	Glorifying of God
نجم-نجوم	star/s	13	3	1	0	23.1	7.7	0	pos	Glorifying of God

The above table is evidence that words have a negative, positive, or a neutral prosody if they typically co-occur with units that have negative, positive, or neutral meanings (McEnery and Hardie, 2012). For instance, the word *day* occurs more frequently than would be expected by chance alone with contexts that can be evaluated as negative, such as afterlife punishment. Moreover, the word *death*, which is intuitively related to ‘loss and grief’, occurs more frequently than expected in contexts of glorifying of God, where the

Qur'an describes God's power in giving life and death (e.g., the cycle of life in nature, and the contrast between the present life and the afterlife), His reviving of the dead land with water and life, and His power of bringing the dead back to life. The lists below illustrate the positive, negative, and neutral natural phenomena terms in the Qur'an.

1- The positive nature terms

The revealed meanings (evaluative and discourse prosodies) of the positive natural phenomena in the Qur'an fall into one of four categories: glorifying of God in the Qur'an, the emphasis on the message of the Qur'an, the reward in the afterlife, and miracle stories. In terms of their conceptual categories in accordance with the ontology of nature terms established in this thesis, they are living creatures (people, soul/s, mankind, angel/s, hand, tree/s, date palm/s, cattle, chest, woman), astronomical bodies (earth, sky/heavens, sun, moon, star, universe), temporal events (night/s, year/s), geographical locations (mountain, spring, sea/s, rivers/s), afterlife locations (heaven, garden), artefact (ship), some weather phenomena (water, wind), and others (life, death).

2- The negative nature terms

On the other hand, the negative natural phenomena in the Qur'an are associated with the following discourse prosodies: afterlife punishment, present life punishment, the shunning disbelievers, and the evil doings of Satan. Their conceptual categories are as follows: temporal event (day, darkness), after-life event (resurrection), some living creatures (face/s, heart/s, jinn, devil/s), geographical location (village-town), afterlife creation (hell), and weather phenomena (drowning, flood, rain).

3- The neutral nature terms

Finally, the neutral natural phenomena in the Qur'an are associated with one discourse prosody, which is the Islamic teachings. The term home-house/s, a geographical location, was found to be the only neutral natural phenomena term out of the 30 most frequent nature terms in the Qur'an.

In the following lines, two examples of the findings relevant to the uncovering of the SP of natural phenomena in the Qur'an will be presented.²³⁸ They are the terms ريح *ryḥ* 'wind/s' and قلب/قلوب *qlb* 'heart/s', each of which will be represented by its most significant

²³⁸ This representation of SP results follows Sinclair (2003, p.19 and p.117) and the SP section in (Hoey, 2000, p. 232)

collocates, a sample of its concordance lines, and the discourse prosodies it acquired through collocation.

a. The term ريح/رياح *ryḥ* ‘wind/s’ in the Qur’an

The term ريح/رياح ‘wind/s’ has been found to occur in both in the singular ريح *ryḥ*²³⁹ and plural رياح *ryaḥ*. Once the disambiguation was conducted on this term and its SP was explored in the Arabic text, it had positive evaluative prosody with discourse prosody of glorifying God. Its root-based disambiguation included the synonyms level (synonyms of the word *wind* were included) and the word level (singular and plural were included). The co-occurrences in the Qur’an of both forms of this term will be discussed in the following.

The analysis of the term ريح/رياح ‘wind/s’ in the Qur’an revealed that in its singular form, the singular *wind* occurs 9 times and has 6 collocates. The most frequent collocates of this word can be categorised as follows: 4 nouns; 2 verbs; and zero adjectives. They are the following words: مطر ‘rain’, عاد ‘Aad’²⁴⁰, جاء ‘came’, شيء ‘thing’, ارسلنا ‘we sent’, and يوم ‘day’. On the other hand, in its plural form, the word *winds* was found to occur 14 times in the Qur’an and has seven collocates. The most frequent collocates of this word can be categorised as follows: 6 nouns; 1 verb; and zero adjectives. They are the following words: يرسل ‘he sends’, ماء ‘water’, بشرا ‘tidings’, ذكرا ‘reminder’, غمم ‘clouds’, رحمة ‘mercy’, and ارض ‘earth’. As a singular word, the term is associated with the discourse prosody of present-life punishment (see examples in the collocate sets below), whereas it is mainly related to the glorifying of God’s creation and his tidings and mercy in its plural form.

Table 35: The evaluative prosody of the words ريح *ryḥ* ‘wind’ and رياح *ryaḥ* ‘winds’ in the Qur’an

Word	Frequency	Pos	Neg	Neu	Pos%	Neg%	Neu %	EP
wind	9	4	7	0	44%	77.7%	0%	Neg
winds	14	9	0	1	64%	0%	7%	Pos

The discourse prosodies of the term ريح/رياح *ryḥ* ‘wind/s’ that were extracted based on this exploration were the following: glorifying of God; present-life punishment; miracle story; and emphasis on the message of the Qur’an. Moreover, a sample of the concordance

²³⁹ ريح is for the singular form, but when disambiguated both singular and plural are represented with ريح.

²⁴⁰ Aad (عاد) are a historic people mentioned in the Qur’an who were sent the Prophet Hud. Available from: [http://corpus.quran.com/concept.jsp?id=aad], [Accessed 04 July 2017].

lines²⁴¹ followed by a sample of the annotated dataset of the term *wind* (in both singular and plural forms) are shown below:

يسيركم ب ر ب حر ف ل ك و جرين ريح1 طيبة و فرحوا جاءة ها <ريح1>مطر و جاء هم موج مكان و ظنوا احيط دعوا الله
ه عذاب غليظ مثل كفروا برب هم اعمال هم كرماد اشتدت <ريح1>يوممطر يقدرون كسبوا شيء الضلال البعيد تر الله
علمناه صنعة ل يوس ل تحصنكم باسكم شاكرون و ل سليمان <ريح1>مطر تجري ب امر ه ارض باركنا و كنا شيء عالمين
يروا الله خلق هم اشد قوة باياة نا يجحدون فارسلنا <ريح1> صرصرا يوم نحسات ل نذيق هم عذاب الخزي حيي1 دنو1 و
عارضنا مستقبل اودية هم قالوا عارض مطر استعجلتم <ريح1> عذاب اليم تدمر شيء ب امر ربها ف اصبحوا يرى سكن
ف اخذناه و جنود ه ف نبذنا هم اليم مليم عاد ارسلنا <ريح1> العقيم نذر شيء انت جعلته ك الريم ثمود قيل تمتعوا
القران ل الذكر مذكر كذبت عاد عذابي و نذر ارسلنا <ريح1> صرصرا يوم نحس مستمر تنزع ناس اعجاز نخل منقعر
عاد ب القارة ثمود ف اهلكوا ب الطاغية عاد ف اهلكوا <ريح1> صرصرا يوم عاتية سخرها سبع ليل و ثمانية حسوما فترى
ناس انزل الله سمو موه حيي1 ارض موت و ب ث د ب و ب و تصريف <ريح> غمم1 المسخر سمو ارض ل آيات ل قوم يعقلون ناس
ها و ادعو ه خوفا و طمعا رحمت قريب المحسنين يرسل <ريح> ب شرا يدي رحمة ه اقلت غمم1 تقالا سقناه ل بلد ميت ف
س تم برازين شيء خزان ه نزل ه ب قدر معلوم و ارسلنا <ريح> ل واقع ف انزلنا سمو موه ف اسقيناكمو ه ب خازنين حيي1
موه انزلناه سمو ف اختلط زرع ارض ف اصبح هشيم تذروه <ريح> شيء مقتدرا المال و البنون زينة حيي1 دنو1 و
جعل ليل لباسا و النوم سباتا و جعل نهر2 نشورا ارسل <ريح> ب شرا يدي رحمة ه و انزلنا سمو موه طهورا ل نحبي بلدة
ارض الاله قليلا تنكرون يهدي ك م ظلم1 بر بحر يرسل <ريح> بشرا يدي رحمة ه الاله تعالى الله بشركون يبدا الخلق
و عملوا الصالحات فضل ه يحب الكافرين اية ه يرسل <ريح> بشرايات و ل يذيقكم رحمة ه و ل تجري ف ل ك ب امره و ل
ف انتقمنا اجرهموا حقا نصر المؤمنين الله يرسل <ريح> فنتير غمم1 ف يبسطه سمو يشاء و يجعله ه سمو فترى مطر

81 21 أولسئلمان الریح عاصفة تجري بأمره إلى الأرض التي باركنا فيها وكنا بكل شيء عالمين	(21:81) And, We disposed, for Solomon the wind to blow strongly — in another verse it is (described as being) rukhā'an, 'to blow softly' (Q. 38:36); in other words, (it is either) blowing violently or gently (respectively), according to what he (Solomon) wanted — making its way, at his command, to the land which We have blessed, namely, Syria; and We have knowledge of all things, among them the fact that God, exalted be He, knew that what He gave to Solomon would prompt him to be subservient to his Lord, and so God did this in accordance with His knowledge.	miracle story
41 16 فأرسلنا عليهم ريحا صرصرا في أيام نحسات لنديقهم عذاب الخزي في الحياة الدنيا ولعذاب الآخرة أشد وأهم لا يؤصرون	(41:16) So We unleashed upon them a raging wind, cold and violent, but without rain, during (some) ill-fated days (read nahisātin, or nahsātin), (days that were) calamitous for them, that We might make them taste the chastisement of disgrace, humiliation, in the life of this world; yet the chastisement of the Hereafter is indeed more disgraceful, more severe, and they will not be helped, to have it warded off from them.	Present-life punishment
46 24 فلمأ رأوه عارضا مستقبل أوديتهم قالوا هذا عارض ممطرنا بل هو ما استعجلتم به ريح فيها عذاب اليم	(46:24) Then, when they saw it, that is, (when they saw) what chastisement (really) was, as a sudden cloud, a cloud that appeared (out of nowhere) on the horizon, heading towards their valleys, they said, 'This is a cloud that will bring us rain!'. God, exalted be He, says: Nay, but it is what you sought to hasten, of the chastisement — a hurricane (rīhun substitutes for mā, 'what') containing a painful chastisement,	Present-life punishment
51 41 وفي عاد إذ أرسلنا عليهم الريح العقيم	(51:41) And (also) in, the destruction of, 'Ād, was a sign, when We unleashed against them a barren wind, (a wind) which brings nothing of good, for it does not bear any rain and does not pollinate any trees; this (wind) was the west wind (al-dabūr).	Present-life punishment
54 19 إننا أرسلنا عليهم ريحا صرصرا في يوم نحس مستمر	(54:19) Indeed We unleashed upon them a clamorous wind, intensely noisy, on a day of prolonged ill fortune, (nahsin mustamirr means) either one of continuous ill fortune or one of severe ill fortune — and this was the last Wednesday of the month —	Present-life punishment

<p>69 6 وَأَمَّا عَادُ فَأَهْلِكُوا بِرِيحِ صَرْصَرٍ عَاتِيَةٍ</p>	<p>[69:5] As for Thamūd, they were destroyed by the [overwhelming] Roar, an excessively severe cry.</p>	<p>Present-life punishment</p>
<p>2 164 إِنَّ فِي خَلْقِ السَّمَاوَاتِ وَالْأَرْضِ وَإِخْتِلَافِ اللَّيْلِ وَالنَّهَارِ وَالْفُلْكِ الَّتِي تَجْرِي فِي الْبَحْرِ بِمَا يَنْفَعُ النَّاسَ وَمَا أَنْزَلَ اللَّهُ مِنَ السَّمَاءِ مِنْ مَاءٍ فَأَخْبَا بِهِ الْأَرْضَ بَعْدَ مَوْتِهَا وَبَثَّ فِيهَا مِنْ كُلِّ دَابَّةٍ وَتَصْرِيفِ الرِّيَّاحِ وَالسَّحَابِ الْمُسَخَّرِ بَيْنَ السَّمَاءِ وَالْأَرْضِ لَآيَاتٍ لِقَوْمٍ يَعْلَمُونَ</p>	<p>(2:164) They then asked for a sign to prove this, and the following was revealed: Surely in the creation of the heavens and the earth, and the marvels contained in them, and the alternation of the night and day, passing and returning, increasing and diminishing, and the ships that run in the sea, and do not become cracked and sink, with what profits men, of trade and merchandise, and the water, the rain, God sends down from the heaven with which He revives the earth, with vegetation, after it is dead, after it has dried out, and He scatters abroad in it all manner of crawling thing, by dividing them and spreading them throughout on account of the vegetation, for they thrive on the fertile pastures it produces; and the disposition of the winds, changing it from south to north, from cold to warm, and the clouds compelled, subjugated by God's command, moving to wherever God wishes, between heaven and the earth, without being attached (to either of the two) — surely there are signs, indicating His Oneness, exalted be He, for a people who comprehend, (a people) who contemplate.</p>	<p>glorifying of God</p>
<p>7 57 وَهُوَ الَّذِي يُرْسِلُ الرِّيَّاحَ بُشْرًا بَيْنَ يَدَيْ رَحْمَتِهِ حَتَّىٰ إِذَا أَقْلَبْتَ سَحَابًا فَقَالَ سُقْتَاهُ لِيُبْدِ مَيِّتٍ فَأَنْزَلْنَا بِهِ الْمَاءَ فَأَخْرَجْنَا بِهِ مِنْ كُلِّ الثَّمَرَاتِ كَذَلِكَ نُخْرِجُ الْمَوْتَى لَعَلَّكُمْ تَذَكَّرُونَ</p>	<p>(7:57) He it is Who sends the winds, unfolding with His mercy, that is, dispersing before the rains (a variant reading (for nushuran, 'unfolding') has nushran; another reading has nashran as the verbal noun; and a third variant has bushran, meaning mubashshiran, '(with which) He is bearing good tidings (of His mercy)'; the singular of the first reading is nashūr, similar (in pattern) to rasūl, 'messenger'; the singular of the last is bashīr) until, when they, the winds, bear heavy clouds, (clouds heavy) with rain, We lead it, that is, (We lead) the clouds (herein is a shift from the third (to the first) person), to a dead land, one devoid of any vegetation, in other words, (We lead it there) in order to revive it, and then We send down thereon, on that land, and bring forth thereby, by this water, fruits of every kind. Like that, bringing forth, We shall bring forth the dead, from their graves, through revivification, so that you might remember, and believe.</p>	<p>glorifying of God</p>
<p>15 22 وَأَرْسَلْنَا الرِّيَّاحَ لَوَاقِحَ فَأَنْزَلْنَا مِنْ السَّمَاءِ مَاءً فَاسْتَفْقَيْنَا كَمْوَهُ وَمَا أَنْتُمْ لَهُ بِخَازِنِينَ</p>	<p>(15:22) And We send the winds as fertilisers, to fertilise the clouds, whereupon they fill up with water, and send down out of the heaven, (out of) the clouds, water, rain, so that We give it to you to drink, for you are not the storers thereof, that is, the stores thereof are not in your control.</p>	<p>glorifying of God</p>

b. The term قلب/قلوب *qlb* 'heart/s' in the Qur'an

The term قلب *qlb* 'heart/s' was found to occur 142 times in both the singular قلب and plural قلوب. Once the disambiguation was conducted on this term and its SP was explored in the Arabic text, it had negative evaluative prosody with discourse prosody in the description of disbelievers, and a positive one in the description of believers. Its root-based disambiguation included the synonyms level (synonyms of the word *heart* were included such as فؤاد/أفئدة), the word level (its two forms were included), and root level as قلب

because the same root قلب can mean *bring down; overthrow; topple*.²⁴² In terms of its exploration, the co-occurrences of this term were with 53 collocates. The most frequent ones are 'they believe', 'disease', 'hearing', 'vision', 'has concealed', 'hypocrites', 'reminders', 'covers', 'be assured', 'reminder', 'they understand', 'Allah'. It is evident from its nearest surroundings that this word tends to collocate equally with several positive words like *faith* and *assured*, and with negative words such as *hypocrites*, *hardened hearts*, *covers on their hearts*, *concealed hearts*, and *disease*. The discourse prosodies of these words are then equally distributed between the description of disbelievers and, the emphasis on the truth of the message; the reassured hearts of the believers as opposed to the concealed hearts of the disbelievers and hypocrites.

A sample of concordance lines of the word *heart/s* followed by a sample of the SP annotated dataset can be found below²⁴³.

كفروا سواء انذرتهم تنذر هم يؤمنون ختم الله <قلب>1 <اذن>1 عين2 غشاوة عذاب عظيم ناس يقول امنا ب الله
ب مؤمنين يخادعون الله امنوا يخدعون نفس1 يشعرون <قلب>1 <مرض> ف زاد هم الله مرضا عذاب اليم يكذبون
ب بعض ها حيي1 الله الموتى و يريكم اياة ه تعقلون قست <قلب>1 <حجر>1 اشد قسوة حجر1 يتفجر نهر1 يشقق ف
خذوا اتيناكم ب قوة اذن1 قالوا اذن1 و عصينا شرب <قلب>1 <عجل>1 ب كفر هم قل ب نسما بامرکم ايمانكم
و اقام الصلاة و ايتاء الزكاة يخافون يوم تتقلب <قلب>1 <عين>2 ل يجزي هم الله احسن عملوا و يزيد هم فضل ه
الله و رسول ه ل يحكم فريق معرضون الحق ياتوا مدعين <قلب>1 <مرض> ارتابوا يخافون يحيف الله و رسول ه اولئك
اسرائيل نزلناه الاعجمين ف قراه مؤمنين سلكتاه <قلب>1 <المجرمين يؤمنون> يروا العذاب الاليم ف ياتي هم
ه و ل دا يشعرون و اصبحت قلب1 موسى فارغا ل تبدي ربطنا <قلب>1 <ل تكون المؤمنين و قالت ل اخته قصيه عين2
مثل جنة هم باية ل يقولن كفروا مبطلون يطبع الله <قلب>1 يعلمون ف اصبر و عد الله حق يستخفك يوقنون ب
نسل ه سلاله موه مهين سواء و نفخ روح و جعل اذن1 عين2 <قلب>1 قليلا تشكرون و قالوا ضلنا ارض خلق جديد ب
خيبرا و توكل الله و كفى ب الله و كيلا جعل الله رجل1 <قلب>1 جوفه جعل ازواجكم تظاهرون امهاتكم جعل ادعياء
تعملون عين2 جاءوكم ف و فكم اسفل زاغت عين2 و بلغت <قلب>1 الحناجر و تظنون ب الله الظنونا ابتلني
ابتلي المؤمنون زلزل زلزلا شديدا يقول المنافقون <قلب>1 <مرض> و عدنا الله و رسول ه غرورا قالت طائفة اهل
يعرفن يؤدين الله غفورا رحيمًا ينته المنافقون <قلب>1 <مرض> زلزل مدن ل نغربك بجاورونك قليلا ملعونين

22 39 أفمن شرخ الله صدره للإسلام فهو على نور من ربه فوزًا للقاسية فلؤبؤهم من ذكر الله أولئك في ضلال مبين	(39:22) Is he whose breast God has opened to Islam, and becomes guided, so that he follows a light from his Lord ...?, like he whose heart He has sealed (with disbelief)? — this (understanding of the ellipsis) is indicated by (what follows). So woe — an expression indicating ‘chastisement’ — to those whose hearts have been hardened against the remembrance of God, that is, (hardened) against the acceptance of the Qur’ān. Such are in manifest error.	emphasis on the truth of revelation
41 5 أو قالوا فلؤبؤنا في آكله مما تدعونا إليه وفي آذاننا وقر ومن بيننا وبينك حجاب فاعمل إننا عاملون	(41:5) And they say, to the Prophet, ‘Our hearts are veiled, (they are) masked, from that to which you call us, and in our ears there is a deafness and between us and you there is a partition, a variance over religion, so act, according to your religion; indeed we shall be acting!’, according to our religion.	disbelievers

242 Available from: [https://www.almaany.com/ar/dict/ar-en], [Accessed 04 July 2019].

243 In the concordance lines, 1 عين and 1 اذن are two disambiguated terms for the words: vision and hearing respectively. (See also Section 2.3 for disambiguation of terms).

<p>42 24 أَمْ يَقُولُونَ افْتَرَى عَلَى اللَّهِ كَذِبًا فَإِنْ يَشَاءِ اللَّهُ يَخْتَمُ عَلَى قَلْبِكَ وَيَمْحُ اللَّهُ الْبَاطِلَ وَيُحْيِي الْحَقَّ بِكَلِمَاتِهِ إِنَّهُ عَلِيمٌ بِدَاتِ الصُّدُورِ</p>	<p>(42:24) Or (am here is like bal) do they say, 'He has invented a lie against God?', in ascribing the Qur'an to God, exalted be He. For if God will, He can seal, He can fortify, your heart, with patience to endure the hurt they cause (you) by such sayings and otherwise — and God did this. And God will efface the falsehood, which they speak, and vindicate, confirm, the truth with His words, (the ones) revealed to His Prophet. Truly He is Knower of what is in the breasts, of what is in the hearts (of people).</p>	<p>emphasis on the truth of revelation</p>
<p>45 23 أَفَرَأَيْتَ مَنْ اتَّخَذَ إِلَهَهُ هَوَاهُ وَأَصْنَعُ اللَّهُ عَلَى عِلْمٍ وَخَتَمَ عَلَى سَمْعِهِ وَقَلْبِهِ وَجَعَلَ عَلَى بَصَرِهِ عِشَابَةً فَمَنْ يَهْدِيهِ مِنْ بَعْدِ اللَّهِ أَفَلَا تَتَذَكَّرُونَ</p>	<p>(45:23)Have you then seen, inform Me (about), him who has taken as his god his (own) desire, (that is) whatever (new) stone he may desire after (having desired) some other stone, considering this (new one) to be better, and whom God has led astray knowingly, on His part, exalted be He, that is to say, knowing him to be of those who follow misguidance before he was created, and set a seal upon his hearing and his heart, so that he cannot hear guidance or understand it, and laid a covering, a darkness, over his sight?, so that he cannot perceive guidance (here one should understand as implied the second direct object of (the verb) a-ra'ayta, 'have you seen', namely: a-yahtadī, 'can he find guidance?). So who will guide him after God?, that is to say, after His leading him astray? In other words, he will not find guidance. Will you not then remember?, (will you not then) be admonished? (tadhakkarūna: one of the two tā' letters (of the original form tatadhakkarūna) has been assimilated with the dhāl).</p>	<p>disbelievers</p>
<p>47 20 وَيَقُولُ الَّذِينَ آمَنُوا لَوْلَا نُزِّلَتْ سُورَةٌ فَإِذَا أَنْزَلْتَ سُورَةً مُحْكَمَةً وَذَكَرَ فِيهَا الْقِتَالَ رَأَيْتَ الَّذِينَ فِي قُلُوبِهِمْ مَرَضٌ يَنْظُرُونَ إِلَيْكَ نَظَرَ الْمَغْشِيِّ عَلَيْهِ مِنَ الْمَوْتِ فَأُولَئِكَ لَهُمْ</p>	<p>(47:20) And those who believe say, seeking to take part in the struggle: 'Why has a sūra not been revealed?', (one) in which there is mention of the (command to participate in the) struggle. But when a definitive sūra is revealed, one in which nothing is abrogated (naskh), and fighting, that is to say, requirement for it, is mentioned in it, you see those in whose hearts is a sickness, that is, doubt — and they are the hypocrites — looking at you with the look of someone fainting at (the point of) death, because of (their) fear of it and (their) aversion to it; in other words, they fear fighting and are averse to it. Yet more fitting for them (faawlā lahum is a subject, the predicate of which (is the following, tā'atun wa-qawlun ma'rūfun))</p>	<p>hypocrites</p>

To sum up, the first section of this chapter reported the results of the exploration of the SP of nature as a Qur'anic theme. It presented these results as frequencies of nature terms, their collocations and LG patterns, and the evaluative and discourse prosodies. In addition, it interpreted some of their results with the defining features of the theoretical background adopted for this thesis. Similarly, examples and lines of concordance and annotated datasets were provided to illustrate these results. The final point that is worth mentioning is: based on the results of SP of nature in the Qur'an, it generally has the evaluative prosody of being positive; its discourse prosodies include glorifying of God, emphasis on the truth of the message, miracle stories, and the description of believers.

Moreover, it should be noted that since the same tasks were conducted on the Arabic text as on the English translations, the results are bound to be reported similarly.²⁴⁴ However, since the second aim of this research is to evaluate these translations based on their congruency in regard to the Arabic representation of SP of nature in the Qur'an, the English results will not be fully reported in this section. Instead, they are appended as electronic

244 The results of the frequencies, collocations, and SP of natural phenomena in the English translations of the Qur'an were represented based on the results of the Arabic 30 most frequent natural phenomena in the Qur'an. They are available in the electronic directory of Language Resources appended to this thesis.

documents to this thesis as datasets of annotated SP of nature in five translations of the Qur'an.

5.2 Results of the evaluation of the translations

This section will focus on reporting the results of the evaluation of the translations regarding their congruency with the Arabic representation of SP of nature in the Qur'an. The comparison to evaluate the most congruent translation in its representation of nature as a theme in the Qur'an was based on the previously mentioned sequence, which consists of examining frequencies, collocations including LG patterns, and finally semantic prosody. Based on the statistical analysis in this thesis, which revealed both the occurrences and co-occurrences of nature terms in the Arabic Qur'an and its translations, a comparison of the frequencies of the six corpora was conducted via alignment in Microsoft Excel and LancsBox mentioned previously in Section 4.2.2. After generating and disambiguating a list of the 30 most frequent Arabic natural phenomena terms via NLTK in Python and manual root-based disambiguation, the produced list was used as the basis for a comparison of the translations in the second phase of this research. Hence, the results of this section rely on the evaluation of the English translations by utilising the list of Arabic terms, their collocations, and their SPs that were produced in the first stage of this research. Accordingly, the English renderings of the Arabic terms in the five translations were computed and disambiguated (e.g., via the necessary root-based and stem-based disambiguation) to compare their frequencies and collocations (including the LG patterns) with those of the Arabic list, and to uncover the SP of nature in each of these translations and then compare their representations to that of the Arabic Qur'an.

In the comparison of the ratio of the terms in the list of Arabic *frequencies* with their counterparts in the five English translations, it was obvious that they all generally adhered to a similar frequency distribution with variances (See Table 36 and Figure 49).

Table 36: The comparison of the ratios of the most frequent natural phenomena in the Qur'an

<i>Text</i>	<i>The relative frequencies of the 30 most frequent nature terms in the Qur'an</i>
Arabic	62.40
Pickthall	70.25
Ali	73.20
Arberry	72.92
Saheeh	76.76
Haleem	71.25

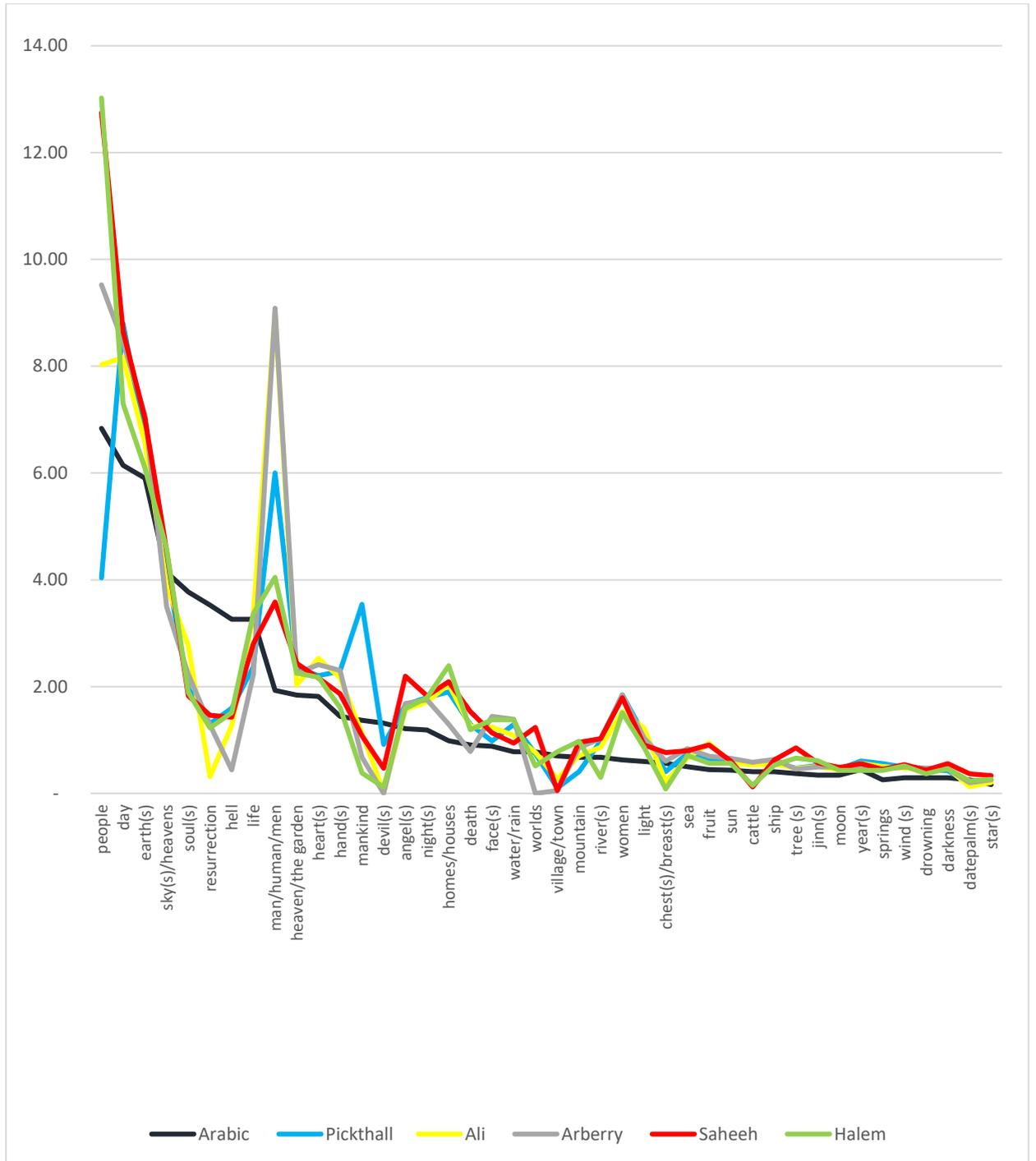


Figure 49: The comparison based on the relative frequencies of natural phenomena in the Arabic text

For example, as seen in Figure 49 which is generated from a spreadsheet of compared frequency results, the exploration of nature terms frequencies in the Qur'an shows that the

word *people* (ratio= 6.84) is the most frequent word from the list of natural phenomena in the Arabic text, whereas the word *day*²⁴⁵ appears most frequently in the Pickthall translation (ratio= 8.80). Ali, Arberry, Saheeh and Haleem, on the other hand, agree with the Arabic text on the ranking of the word *people* (ratios= 8.02; 9.52; 12.75 and 13.02, respectively) as the most frequent in their textual representation of nature. They also confirm Baker's claim that translators tend to use the most frequent words in the ST with more occurrences in the TT:

Studies of corpora of translated texts have shown that they tend to have higher incidences of very frequent words and that they tend to be more explicit in terms of grammar (Baker, 1993 as cited in Hunston, 2006, p.246).

Another noticeable feature emerging from the comparison is the fact that there is an extreme variance between the representation of the word *hell* as an afterlife creation in the Arabic Qur'an and its translations; it is more frequent in the Arabic Qur'an than in its translations. This can be explained by a close inspection of the results of the word *fire*²⁴⁶. The fact that it is shown to appear more repeatedly than its Arabic equivalent in its Arabic reference text illustrates that the translators opted to use the word 'fire' rather than 'hell' in the translation of this word.

Table 37: The frequency distribution of the word 'fire' in the Qur'an and its translations

Term	Arabic	Ratio	Pickthall	Ratio	Ali	Ratio	Arberry	Ratio	Saheeh	Ratio	Haleem	Ratio
Fire	16	0.20	162	2.75	200	3.14	148	2.71	163	2.84	149	2.47

Another noteworthy example is Pickthall's use of the word *mankind* as a translation of both the Arabic words رجل *man* and بشر *mankind*. Hence, in comparison with its frequency in the other texts, it is more prevalent in his translation than either of its correspondents in the Arabic text and the other translations. Finally, Figure 49 shows that the translations are most congruent in the prevalence of natural phenomena terms with the Arabic representation in the use of the words such as *star/s* as an astronomical body and the word *angel/s* as a living creature. Moreover, the sums of the ratios of natural phenomena in the

245 The second most frequent natural phenomena term in the Arabic text.

246 Other synonymous words for the word نار *nār*, which could be translated as 'fire' in the Qur'an, include the following: جهنم/لظى/سقر/جحيم/الحطمة/شهب

Qur'an show a tendency on the part of the translations to vary in their closeness to the Arabic sums of nature conceptual categories (See Figure 50).

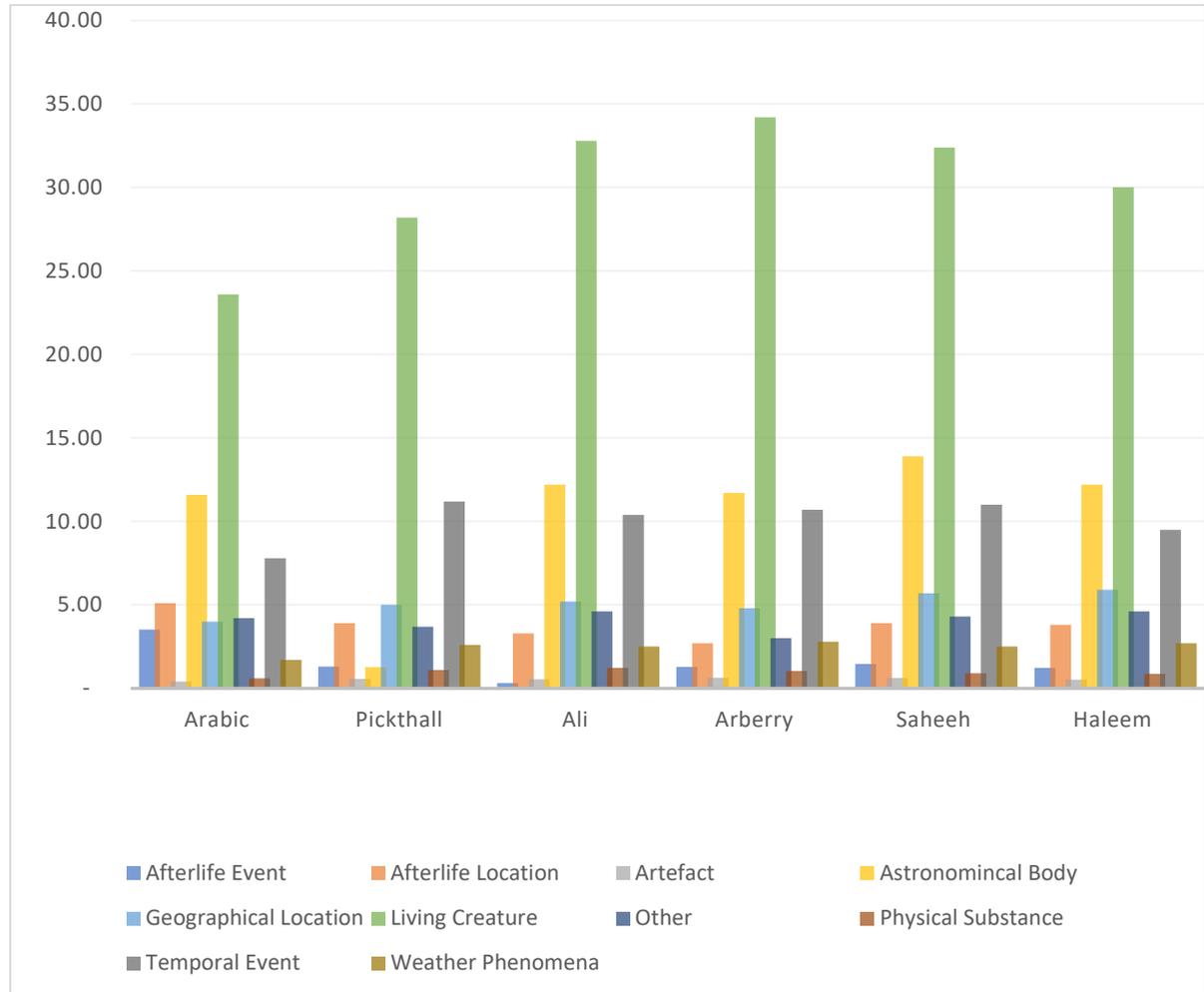


Figure 50: The semantic categories of natural phenomena

Based on the figure of the sums of ratios of the semantic categories of nature in the Qur'an, the following lists were generated to show, in descending order for each, the closest representation of nature in the Qur'an in light of semantic categories of nature.

Afterlife Event

Saheeh

Arberry, Pickthall

Haleem

Ali

Afterlife Location

Pickthall, Saheeh

Haleem

Ali

Arberry

Artefact

Haleem

Ali

Pickthall

Saheeh

Arberry

Astronomical Bodies

Arberry

Ali, Haleem

Pickthall

Saheeh

Geographical Location

Arberry

Pickthall

Ali

Saheeh

Haleem

Living Creatures

Pickthall

Haleem

Saheeh

Ali

Arberry

Other

Arberry

Pickthall

Saheeh

Haleem, Ali

Physical Substances

Haleem, Saheeh

Arberry

Pickthall

Ali

Temporal Event

Haleem

Ali

Arberry

Saheeh

Pickthall

Weather Phenomena

Ali, Saheeh

Pickthall

Haleem

Arberry

The lists show that each of the translations tends to be the closest to the Arabic text in at least one of the categories. However, while Haleem and Ali are the closest to the Arabic text in at least one category, Arberry and Saheeh are closest to the Arabic text in its representation of three semantic categories of nature in the Qur'an.

In addition, the lexico-grammatical patterns of the English texts were identified, and their functions determined both the evaluative and discourse prosodies (See Appendix D for the lexico-grammatical patterns and Appendix E for the English evaluative prosodies results). An overview of their statistical profiling is provided in the following table.

Table 38: The statistical representation of the lexico-grammar of nature in the Qur'an in comparison with its translation

<i>Dataset</i>	<i>N+N</i>	<i>N+V</i>	<i>V+N</i>	<i>ADJ+N</i>	<i>N+ADJ</i>	<i>AV+N</i>	<i>N+AV</i>
Arabic	204	60	79	18	13	2	2
Pickthall	91	58	66	17	14	3	5
Ali	177	42	75	19	13	7	4
Arberry	140	46	42	22	14	6	7
Saheeh	140	40	54	14	10	5	5
Haleem	149	35	32	14	4	1	3

Finally, the comparison of SP of nature in the Qur'an in the evaluation of the five selected translations revealed that Saheeh and Haleem succeeded equally in matching the 36 out of 42 Arabic nature terms in their evaluative prosodies (see Table 39).

Table 39: The comparison of congruency of evaluative SP (as in Ebeling, 2014)

<i>Text</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
congruency score	35/42	35/42	34/42	36/42	36/42

On the other hand, in the discourse prosodies, it was evident that Haleem was more congruent with the representation of nature discourse prosodies to that of the Arabic results. The table below shows the dominance of Haleem's results in the comparison of the discourse prosodies. This concludes that in regard to the representation of SP of nature in the Qur'an, Haleem's translation is evidently the closest to the Arabic original.

Table 40: The comparison of congruency of discourse SP

<i>Text</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
congruency score	33/42	32/42	34/42	34/42	36/42

To recap, this section reported the results of the second stage of this thesis, which is the evaluation of the five translations of the Qur'an concerning their representation of the SP of nature in the Qur'an. It illustrated this evaluation through references to frequencies, collocations and LG patterns, and evaluative and discourse prosodies. In regard to the uncovering of SP, it was shown through a ranking that, out of the five translations examined here, Haleem is the most congruent to the Arabic representation of SP of the theme of nature in the Qur'an. Further discussion of these findings with reference to the criteria of consistency and accuracy in the evaluation of the translations will be provided in the following section of this chapter.

5.3 Discussion of the results

This section discusses the results in light of the research questions, which are the following:

- 1- Is there SP in the representation of nature in the Qur'an and its translations?
- 2- What are the lexico-grammatical patterns, evaluative prosodies, and discourse prosodies of nature in the Qur'an and its translations?

- 3- Which of the five translations is the most congruent and which is the most divergent from the representation of nature in the Qur'an?
- 4- How can variances of the representation of SP of nature in the English translations of the Qur'an be justified in terms of consistency and accuracy?

In response to the first question *Is there SP in the representation of nature in the Qur'an and its translations?* the results will provide evidence for the presence of SP in the analysis of the representation of the theme of nature in the Qur'an. In response to the second question *What are the lexico-grammatical patterns, evaluative prosodies, and discourse prosodies of nature in the Qur'an and its translations?*, there will be a list of the identified LG patterns of words related to nature in the Qur'an with their significance in building a network of meanings with collocations and the relevant evaluative and discourse prosodies that were uncovered through the analysis of these LG patterns. After that, in response to the third question *Which of the five translations is the most congruent and which is the most divergent from the representation of nature in the Qur'an?*, the findings of the evaluation of the translations based on their SP representation of these nature terms in the target texts will be addressed. Finally, in response to the fourth question *How can variances of the representation of SP of nature in the English translations of the Qur'an be justified in terms of consistency and accuracy?*, the researcher will attempt to justify the variances of SP results in the selected English renderings of the Qur'an, and highlight the importance of SP for a translator to achieve accuracy and consistency.

5.3.1 Is there SP in the representation of nature in the Qur'an?

Yes, SP was found to be a characteristic of the language of the Qur'an represented through the collocations of natural phenomena in the Qur'an. For example, Figure 51 shows the evaluative prosodies of nature terms that were transferred to them by habitual co-occurrence (Hunston, 2002, p. 61), unlike their intuitive connotations, which are not collocational like SP. Moreover, the recurrent discourse prosodies that are shown in Figure 52 illustrate that the majority of verses with natural phenomena convey the discourse prosody of the glorifying of God (397 collocations). The five most frequent discourse prosodies of natural phenomena in the Qur'an are discussed in the following sub-section.

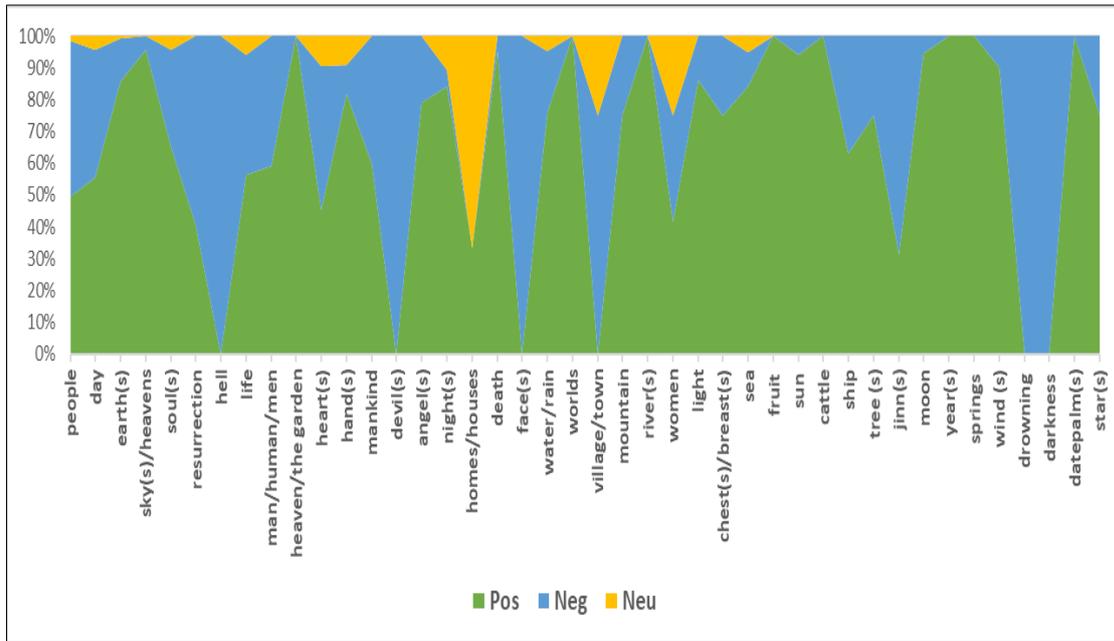


Figure 51: The evaluative SP of nature in the Qur'an

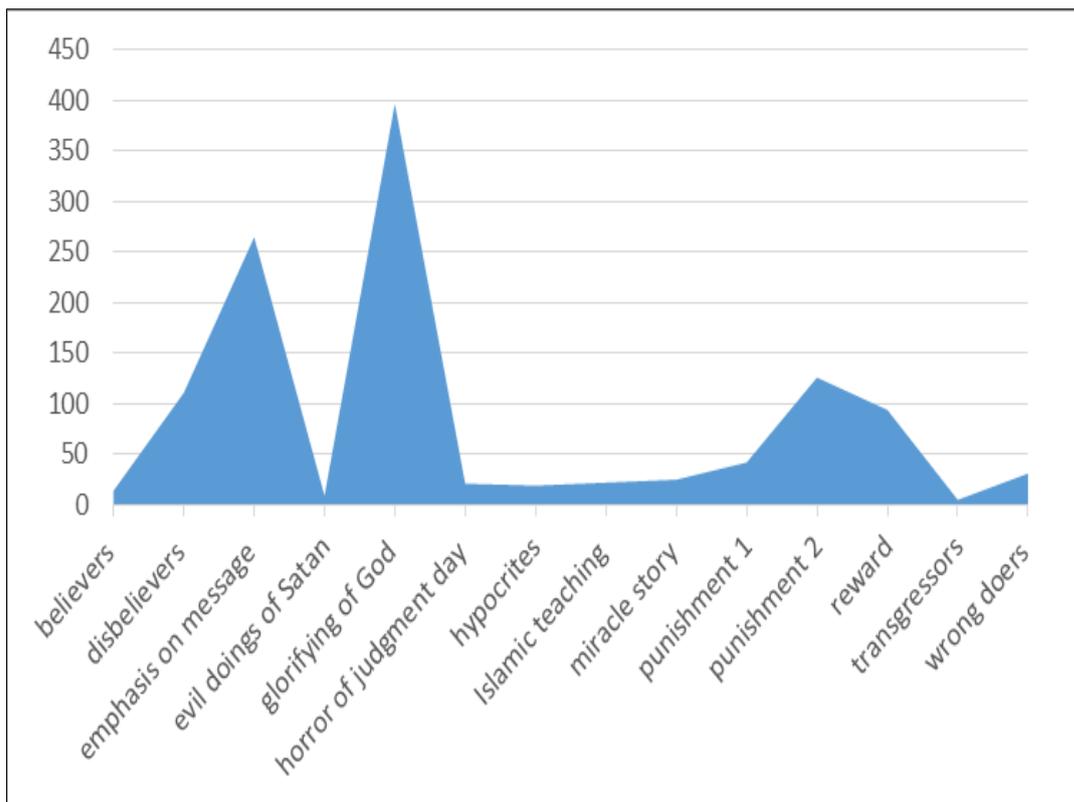


Figure 52: The discourse prosodies of nature in the Qur'an

Another indication of the presence of SP in natural phenomena in the Qur'an can be found in the concordance lines of the word موت mwt 'death'. It was found to violate its intuitive

connotation because of collocation. According to its meaning as per the dictionary, this word has a negative association with grief, whereas in the Qur'an it is shown through SP analysis of its collocate sets as a part of nature's cycle of life and an inevitable beginning of the afterlife, leading to the Hereafter (See also Section 4.1.5). Therefore, it can be said that the intuitive negative connotational meaning of the word 'death' signifying loss and grief is ironically presented in the Qur'an; and this word acquires a positive prosody instead. This is a feature of meaning by collocation which Louw finds to be an indication of SP. With the violation of intuitive connotational meaning (see Louw, 1993, p.137 and p.157), there is a potential SP to achieve an effect of irony on the listener or reader of the word.

To conclude, this sub-section provided evidence that there is SP in the representation of natural phenomena in the Qur'an; this is within the line of established corpus linguistic thinking (e.g., Sinclair, 1996; Louw, 2000; Thompson, 2000; Stubbs, 2001; Hunston, 2002; Partington, 2004b). This research thus asserts the presence of SP as a collocational phenomenon which colours the connotative meanings of nature terms via collocation. To find collocations of words related to nature appearing in the Qur'an, a quantitative analysis was conducted to compute statistical collocations (Evert, 2008), and a qualitative analysis was implemented to interpret the results of collocations by examining the concordance lines of the instances where these statistical (i.e. not occurring by chance) collocations and verified bigrams occur (as in Stubbs, 2001; Sinclair 1996; 2003; 2004). This mixed approach produced datasets of SP annotated texts (Arabic and English) that are electronically appended to this thesis as proof of the presence of SP in the representation of natural phenomena in the Qur'an. More details on how the results of SP were interpreted (e.g., LG patterns and SP) will be provided in the following sub-section.

5.3.2 What are the lexico-grammatical patterns, evaluative prosodies, and discourse prosodies of nature in the Qur'an?

As mentioned previously in this chapter, the eight *lexico-grammatical (LG) patterns* of the 30 most frequent natural phenomena in the Qur'an that were found in this research are the following: Noun (N)+Noun (N); Noun (N) + Verb (V); Verb(V)+ Noun (N); Adjective (ADJ)+Noun (N); Noun (N)+Adjective (ADJ); Noun (N)+ Adverb (AV); and Adverb (AV)+ Noun (N). They were retrieved qualitatively as a means of interpreting the results of the statistical analysis, which was conducted to produce the statistically significant and

statistical collocations of nature (bigrams in this research). In this regard, when these recurrent LG patterns in the bigrams of nature in the Qur'an were used to observe the concordance lines of nature terms, they uncovered the evaluative and discourse prosodies. It was found that the first five of these patterns generally convey the meaning of glorifying God as the most frequent of the discourse prosodies. Examples of each are the following: 'heavens + earth'; 'man+ praise (v.)'; 'sent+ heavens'; 'six+ days'; and 'moon+ subservient'. Similarly, the last two patterns were shown to convey an emphasis on the message of the Qur'an, which was the second most prevalent of the discourse prosodies in this research. Examples of each pattern respectively are the following: 'earth + verily' and 'verily+ hearts'.

These patterns and their prosodies remind us of Halliday and Matthiessen's (2004) and Sinclair's (2004b) view of the lexico-grammar that it is an underlying component of the meaning-making system of a language. Similarly, they confirm the findings of Partington (2004b; 2009) on "semantic –or evaluative- prosody" and Partington et al. (2013, p.8) on SP and lexical grammar, who claim that "lexical items are very largely selected by speakers in batches rather than singly" and that "meanings in utterances, including and especially evaluative meanings, are prosodic-spreading over stretches of language-rather than atomistic in nature".²⁴⁷ Furthermore, the conveyance of meanings through the analysis of the concordances of the LG patterns in this research agrees with Gledhill's (2011, p.73 and p.91) view that lexico-grammar is structured by the expressive and communicative functions it has evolved to display (i.e., discourse prosodies). Moreover, these findings not only support the credibility of the claim that the "examination of concordances generally helps reveal the existence of semantic prosody" (Baker, 2016, p.145), but also add to it the importance of scrutinising this observation by extracting LG patterns and employing them in unveiling the meanings which result from the habitual co-occurrences (i.e. collocations) and cause the connotative colouring (i.e. SP) of natural phenomena in the Qur'an.

To conclude this discussion on the LG patterns of natural phenomena in the Qur'an, a data visualisation is presented, a collocational graph produced in LancsBox²⁴⁸ to help the reader envisage the network of meanings emerging from the SP of nature in the Qur'an shown in

²⁴⁷ See also Partington et.al (2013) and Section 3.1.2.

²⁴⁸ The texts that were read in LancsBox were the Python-processed texts, and the association measure chosen for the purpose of this data visualisation in the software was the Log Likelihood association measure.

Figures 53 and 54 (a guide of the terms and their translations is also provided on the left of the Arabic figure below).

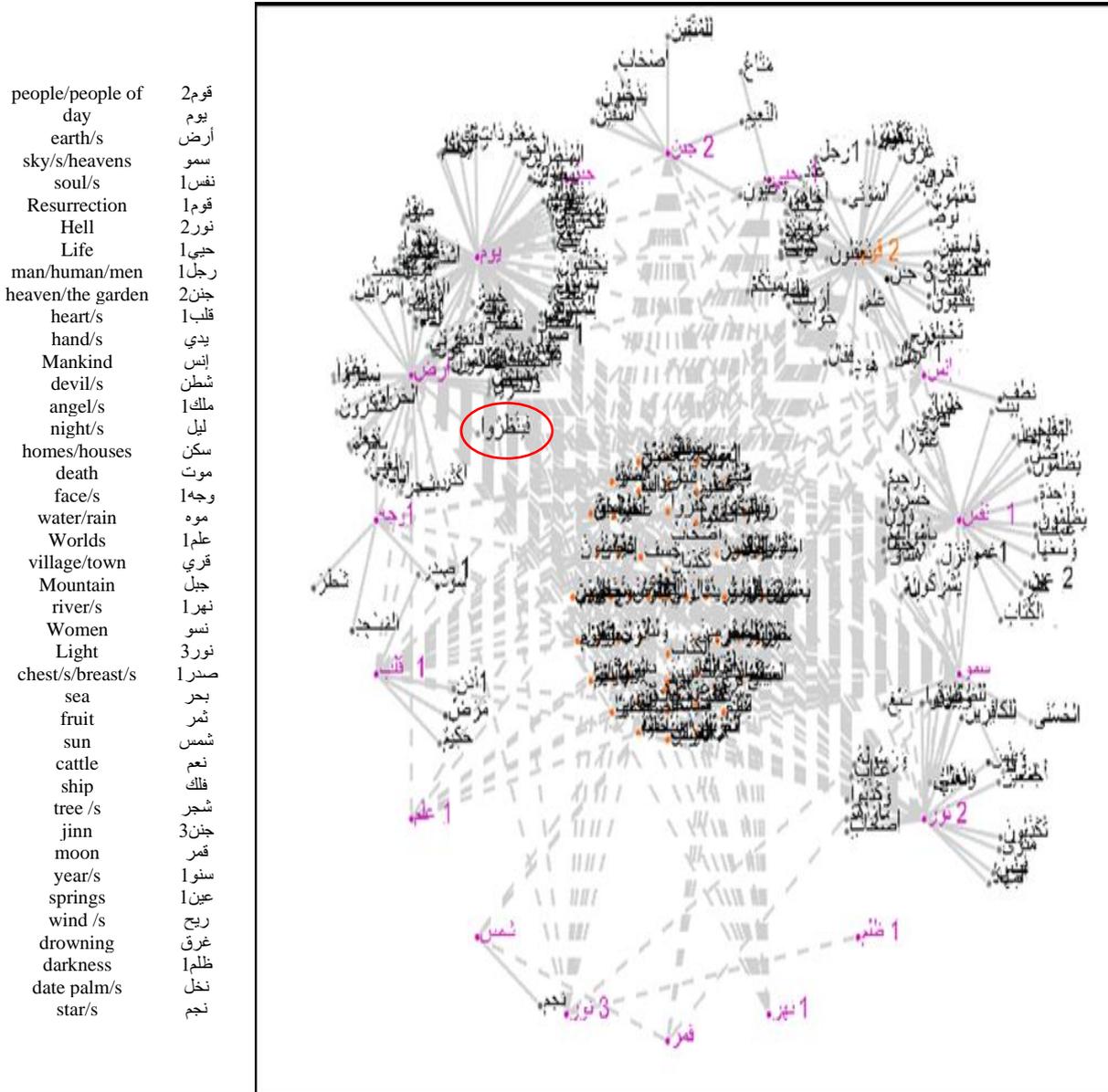


Figure 53: The collocational network of discourse SP of nature nodes in the Qur'an

Figure 53 shows how natural phenomena roots in the Arabic Qur'an are related to one another by virtue of sharing collocate sets. For example, one can see that the roots *rd* 'earth' and *ywm* 'day' are very close to one another by a wordsuch as 'فليظنوا' 'they shall see' (in the red circle in the figure), and a close inspection of their collocates reveals that they share this collocate to reveal the positive SP signifying the glorification of God (See also Appendix C).

affirming the LG cline,²⁴⁹ which states that collocate sets and grammatical patterns function together to reveal the pragmatic side of language (discourse prosody).

A more specific representation of LG patterns in the collocations of natural phenomena in the Qur'an is shown in the alignment of collocation networks illustrated in the figure below.

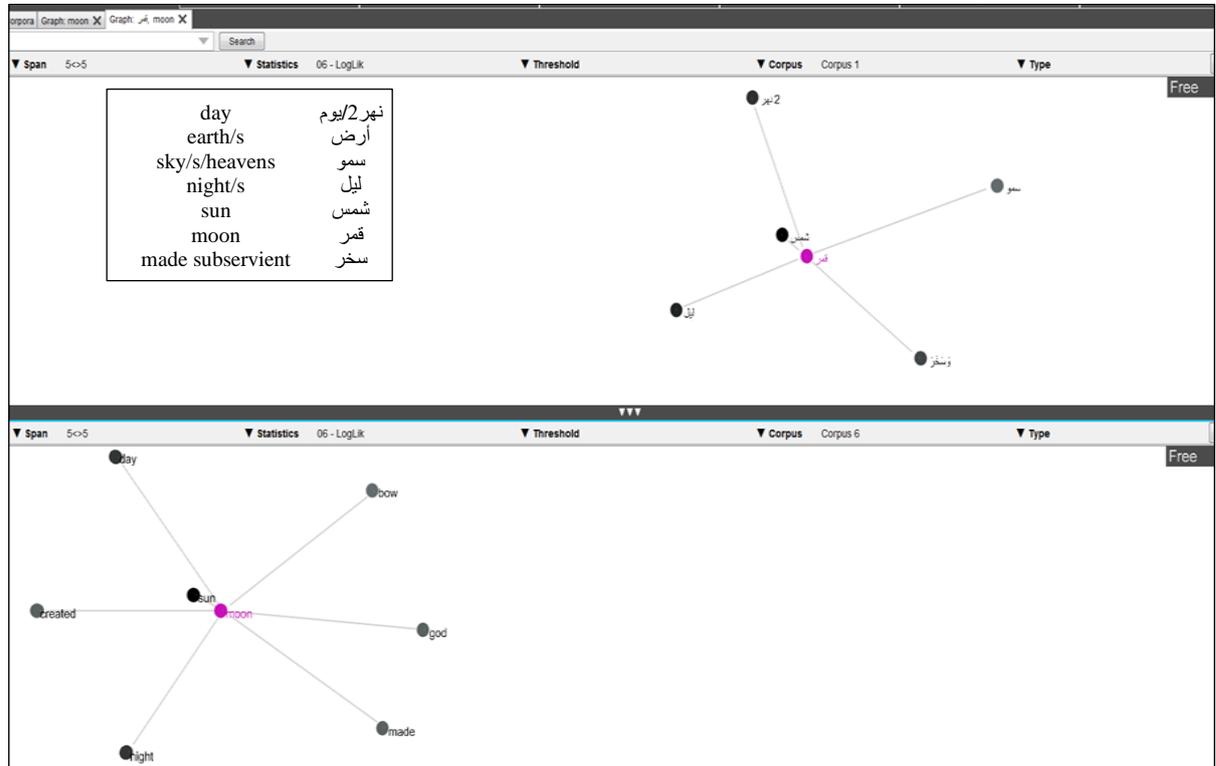


Figure 55: A LancsBox portrait of the alignment of GraphColls of the word قمر *qmr* 'moon' in the Arabic and English corpora (Haleem's translation)

Figure 55, which exemplifies a LancsBox application of the *GraphColl* tool, shows that the LG patterns of the word 'moon' in Arabic and Haleem's English translation are similar. The matching collocates of the node 'moon' in the Arabic and Haleem's translation are the following: شمس *sun*, نهار/يوم *day*, ليل *night*, and سخر *made* [subservient].

As for the *evaluative prosodies*, an analysis of the 30 most frequent natural phenomena terms in the Qur'an showed that the semantic category of living creatures represented the highest number of occurrences with either positive or negative prosodies, while the neutral

249 See also Halliday and Matthiessen (2004) and Section 3.2.1.

evaluative prosody was only displayed in the semantic category of geographical location (see Figure 56).

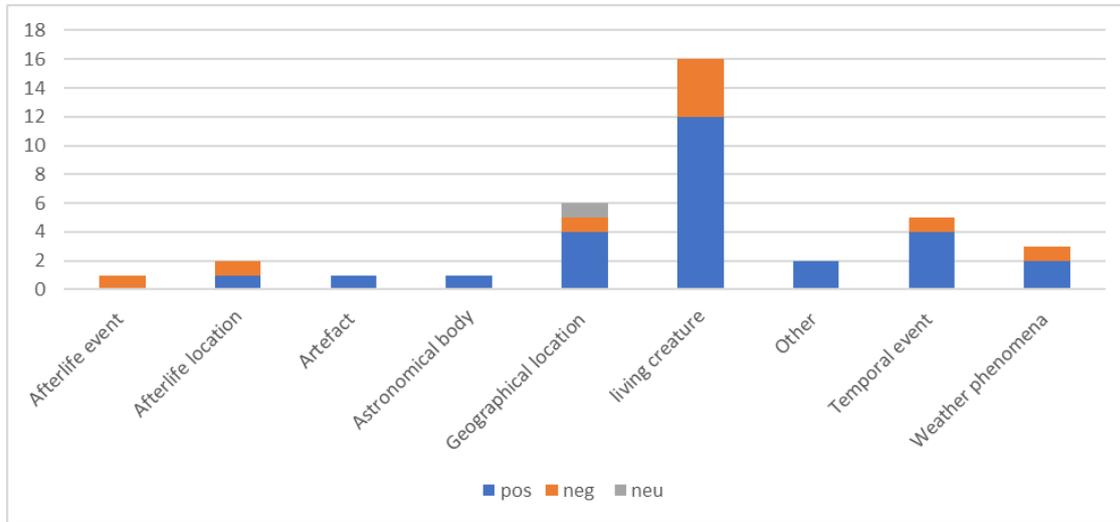


Figure 56: The analysis of evaluative SP of natural phenomena in the Qur'an

Moreover, out of the 13 categories of nature in the Qur'an (Chapter Two), five *discourse prosodies* stand out as the most frequent in the statistical and corpus linguistic analyses of SP of natural phenomena in the Qur'an (see also Figure 52 in Section 5.3.1). A corpus linguistic overview of the five most frequent discourse prosodies of natural phenomena in the Qur'an will be provided in the following lines (for illustrative examples of these discourse prosodies, see the researcher's commentary in Appendix G).

Glorifying of God through nature

Glorifying of God through nature is a discourse prosody, which was uncovered in the analysis of 397 collocations of nature in the Qur'an. Words like *praise, glory, glorify, dominion, All-Mighty, All-Wise, created, creation, Creator, made, knows, subjected, subservient, sustenance, bow, give, gave, establish, grateful, measure, bounties, spread (earth) (v.), etc.* are recurrent neighbouring collocates to natural phenomena terms in the Qur'an, making their evaluative prosody positive and revealing the discourse prosody of glorifying of God.

The emphasis on the message of the Qur'an

The second most frequent SP meaning of natural phenomena in the Qur'an is the emphasis on the message which is the truth of the Qur'an and the unity of God, recurrent discourse prosody occurring in 265 collocations of nature in the Qur'an. Words such as *truth, sign(s),*

understand, submit, remind, witness, remember, say (addressed to Prophet Mohammed) (v.), invite (people to believe) (v.), light (n.) (to mean the message), darkness (to mean disbelief), worship (v.), promised (reward in the afterlife), warning, guidance, guide, faith, see, etc. are neighbouring collocates to natural phenomena terms in the Qur'an, making their evaluative prosodies positive and revealing the discourse prosody of the emphasis on the message of the Qur'an.

The afterlife punishment

The exploration of SP of nature in the Qur'an shows that afterlife punishment, as a discourse prosody of natural phenomena, is the third most frequent of the discourse prosodies of nature in the Qur'an (126 collocations). Words such as *abode (in Hell), fire, enter (Hell), penalty, punishment, black (face), drink (scalding water), woe, punished, companions (of the Fire)*, etc. are neighbouring collocates to natural phenomena terms in the Qur'an, rendering their evaluative prosodies negative and revealing the discourse prosody of punishment in the afterlife.

The shunning of disbelievers

The shunning of disbelievers is the fourth most frequent motif in the Qur'an (111 collocations). Words such as *rejection, reject, deny, disbelief, disbelieve*, etc. are neighbouring collocates to natural phenomena terms in the Qur'an, changing their evaluative prosodies to negative and revealing the discourse prosody of shunning disbelievers.

The afterlife reward

Reward in the afterlife (94 collocations) is the fifth most recurrent of the discourse prosodies of nature in the Qur'an. Words like *abode (in Heaven), rivers, flow, enter (the Garden), garden, triumph, bliss, Eden, grapevines, reward, drink (water from the springs in the Garden), companions (of the Garden), righteous*, etc. are neighbouring collocates to natural phenomena terms in the Qur'an which make their evaluative prosodies positive and reveal the discourse prosody of reward in the afterlife.

To sum up, the previous two sub-sections addressed the research question on the LG patterns, and evaluative and discourse prosodies of natural phenomena in the Qur'an. They provided a corpus-linguistic overview of the exploration of SP of natural phenomena in the Qur'an; this also fulfilled the primary aim of this thesis. In this regard, evidence was

provided of the presence of SP in natural phenomena in the Qur'an; that is, the 30 most frequent nature terms in the Qur'an were shown to have positive, negative, and neutral evaluative prosodies based on their collocations. In addition, their discourse prosodies were identified based on LG patterns and observation of concordance lines. These LG patterns not only convey meanings of natural phenomena but are also the building blocks for a network of meanings that can be visualised, which proves that meaning is typically dispersed over several word-forms which habitually co-occur in the text (Stubbs, 2001, p.63). To put it differently, the word forms (i.e., LG patterns) habitually co-occur, and the distribution of the theme of nature is dispersed throughout the Qur'an. This emphasises the importance of an awareness of these habitual co-occurrences and their representations, as they can aid the understanding of the Qur'an.

To achieve the secondary aim of this thesis, the following two sub-sections will address the remaining two questions of this research about the evaluation of the translations of the Qur'an.

5.3.3 Which of the five translations is most congruent, and which is the most divergent from the representation of nature in the Qur'an?

Based on the results of this research, it can be said that the most congruent translation in regard to the representation of nature in the Qur'an is Haleem's translation. An alignment of the results of SP (evaluative and discourse prosodies) of the Arabic Qur'an with those of the English translations to mark the congruence revealed that Haleem scored the highest out of the translations, with a total of 72 out of 84 (42 Evaluative SPs +42 Discourse SPs for each natural phenomenon term in this research) in its SP congruency score. On the other hand, the most divergent from the representation of nature in the Qur'an is Ali's translation with a total of 67 out of 84 in its SP congruency score (see Figure 57).

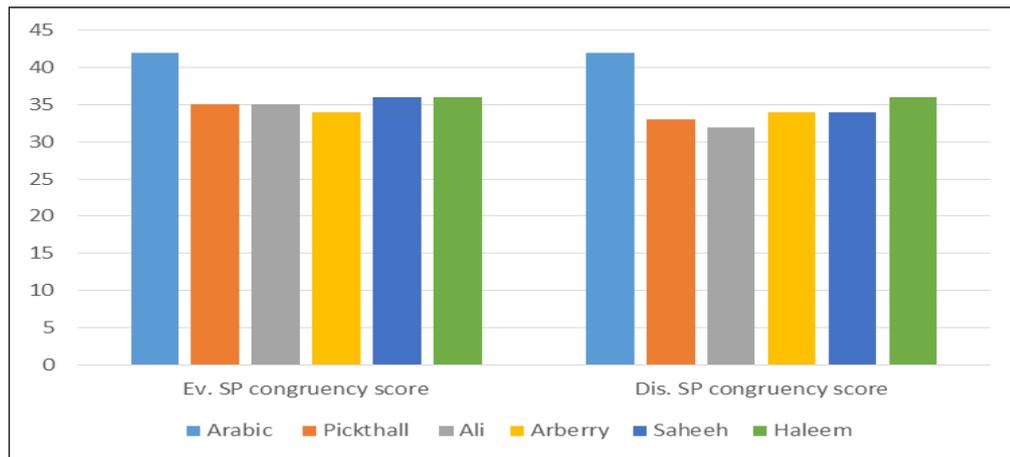


Figure 57: The SP congruency scores of the English translations

These findings above confirm concerns expressed by Sardinha (2000), Dam-Jensen and Zethsen (2008), Morley and Partington (2009), and Stewart (2009) about translators missing out on the importance of SP in the translation process. The fact that there are discrepancies, although not abundant, in the SP results among the datasets proves that there could be a loss of consistency and perhaps accuracy in the representation of natural phenomena in the Qur'an. Hence, this research highlights this issue by reporting the results of the evaluation and attempts to explain the SP variances between the source text and its translations in the parallel corpora without making judgments on the quality of the translations.

5.3.4 How can variances in the representation of SP of nature in the English renderings of the Qur'an be explained in terms of consistency and accuracy?

This section presents the researcher's discussion of the potential reasons for the variances in the representation of SP of natural phenomena in the translations of the Qur'an, which affect the accuracy and consistency of the SP translation. One possible explanation is that these discrepancies are rooted in the word level and caused the variance in the counts and in turn the collocation extraction and SP allocation. In this regard, it has been claimed that one of the problems of loss of translation in the Qur'an is the loss of meaning because of lexical gaps between English and Arabic, which are two languages that frequently express reality in different ways. Lexical gaps have implications at different levels including the semantic and the morphological levels (Abdul-Raof, 2005, p.118), or they may be due to differences in the denotative meanings between the source language and the target language (Bentivogli and Pianta, 2005, p.256).

SP and Accuracy

This research highlights Mona Baker’s definition of *accuracy* in translation, which is defined as being precise in rendering from one language into another, i.e., making the correct choices of equivalent terms in the target language text to convey the meaning of the source language text (Baker, 1993, pp.233-52). The following table shows examples of possible translation loss in relation to the nature terms *plant*, *wind*, and *camel* in at least one of the translations. The renderings of the verses display a possible lexical gap caused by a variance in the choice of translation of natural phenomena.

Table 41: Examples of translations with lexical gaps

<i>Verse</i>	<i>Tafsīr</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
6 55 وَالنَّجْمِ وَالشَّجَرِ يَسْجُدَانِ	and the grass, all stalkless vegetation, and the trees, that (vegetation) which possesses stalk, prostrate, comply with what is required of them.	The stars and the trees prostrate.	And the herbs and the trees - both (alike) prostrate in adoration.	and the stars and the trees bow themselves;	And the stars and trees prostrate.	the plants and the trees submit to His designs;
1 77 وَالْمُرْسَلَاتِ عُرْفًا	By those sent in succession (‘urfan), that is, (by) the winds that follow (one another) in succession, like the mane (‘urf) of a horse, one part (of hair) coming after the other (‘urfan is in the accusative because it is a circumstantial qualifier);	By the emissary winds, (sent) one after another	By the (Winds) sent forth one after another (to man's profit);	By the loosed ones successively	By those (winds) sent forth in gusts	'By the (winds) sent forth in swift succession
33 77 جَمَالَتٌ صُفْرًا	as if they were (dark) yellow camels (jimālātun is the plural of jimalatun, the plural of jamalun; a variant reading has jimālātun), in terms of their shape and colour. In hadīth (it is stated), ‘The sparks of humans (thrown into the air by the Fire) are black as pitch (qīr)’; the Arabs call dark camels sufr, ‘yellow’, because a touch of yellow is mixed with their blackish colour, and therefore it is said that ‘yellow’ in this verse (actually) means ‘black’, on account of what has been mentioned; but some say no (to this	(Or) as it might be camels of bright yellow hue.	’As if there were (a string of) yellow camels (marching swiftly).’	sparks like to golden herds.	As if they were yellowish (black) camels.	and as bright as copper.’

<i>Verse</i>	<i>Tafsīr</i>	<i>Pickthall</i>	<i>Ali</i>	<i>Arberry</i>	<i>Saheeh</i>	<i>Haleem</i>
	interpretation); (sharar is the plural of sharāra; qīr is qār, 'pitch').					

The lexical gap in the first verse is because the word نجم *najm* 'star' has two possible meanings²⁵⁰: it can mean 'star in the sky' (A-Tabari and Al-Saadi) or 'plants of the earth' (Al-Qortoby, Ibn Kathir, and Baghaway). Hence, when computing the frequencies of the word 'star' in the translations, there is a variance caused by some translators consulting exegeses which say that the word means 'plants' in this particular verse. The lexical gap in the second verse is arguably due to a translator's oversight regarding the importance of explicitly rendering the Arabic word رِيَّاح *rīyāḥ* as the English word 'winds'. As the first verse, the third verse shows a variety of interpretations that could cause a semantic lexical gap. The word جمالت *ġimālatun*, typically meaning 'camels',²⁵¹ can have several interpretations: 'black camels' (Al-Qortoby and Ibn Kathir), 'ship anchors' (Al-Tabari), 'copper' (Ibn Kathir), 'flaming fires' (Al-Saadi), and 'stones' (Baghaway).

These examples indicate that the variances in the translators' choice of English renderings can be attributed to their choices of exegeses of the Qur'an. Therefore, the counts of the frequency of the natural phenomena words by the NLTK in Python would be compromised. This, in turn, affects the exploration of SP in the translations of the Qur'an. This is an issue of accuracy of SP caused by word-level variances of the English renderings of some of the natural phenomena in the Qur'an.

Moreover, an investigation of the comparison of some of the LG patterns in the corpora via Log-Likelihood scores reveals that the LG patterns of the word 'sun' in the Quran, for example, appear to be similar in the in English renderings, except for a difference in the LG pattern [V+N] in the bigram 'subjected sun' (See Figure 58).

250 Available from: [<http://quran.ksu.edu.sa/tafseer>], [Accessed 18 May 2019].

251 Although this word is not within the list of the 30 most frequent nature terms, it demonstrates a possible example of a lexical gap in the translation of the Quran.

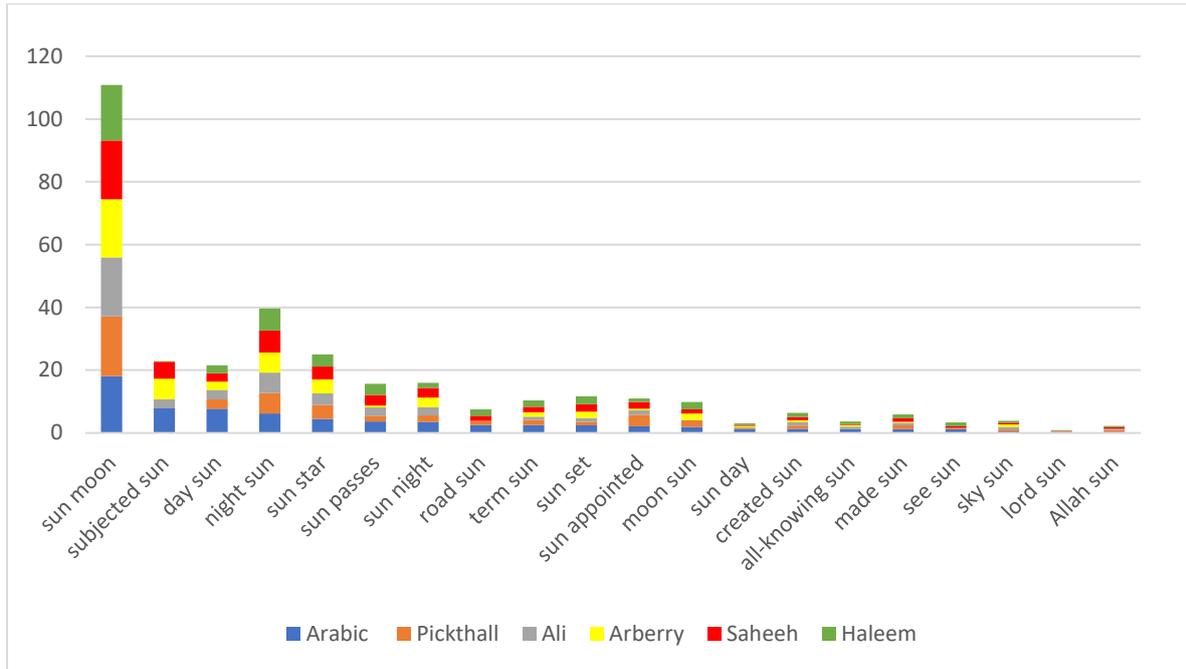


Figure 58: A comparison based on the LLR scores of the collocations of the word ‘sun’

A closer look at the collocate lists of this term in the different corpora of this research (as seen below) reveals that both Pickthall and Haleem did not opt for the exact choice of words or otherwise the program did not extract the word ‘subjected’ as a near or highly ranked collocate to the node ‘sun’ in these translations; hence, the difference in the collocate lists emerged.

Arabic	Pickthall	Ali	Arberry	Saheeh	Haleem
moon	moon	moon	moon	moon	moon
day	night	night	night	night	night
night	day	rising	subjected	subjected	rising
subjected	appointed	day	day	day	day
star(s)	service	subjected	stars	rising	stars
	made	stars	rising	stars	made
	stars	course	setting	made	created
	maketh	law	bow	course	god
	created	prostrate	praise	created	
	Allah	made	god	said	
		created		Allah	
		people			
		one			
		lord			
		said			
		Allah			

Another example of a variance the researcher examined is in the SP results of the term ‘people’ (as seen below).

	Arabic		Pickthall		Ali		Arberry		Saheeh		Haleem	
Nature Term	Arabic Dis SP	Arabic Ev.SP	Dis. SP	Ev. SP	Ali Dis SP	Ali Ev SP	Arberry Dis SP	Arberry Ev SP	Saheeh Dis SP	Saheeh Ev SP	Haleem Dis SP	Haleem Ev SP
people	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos	emphasis on message	pos/neg	disbelievers	neg	emphasis on message	pos

In accordance with the comparison method and results of this research, Saheeh is evidently the most divergent of the translations in the representation of the SP of the term ‘people’ in the Quran. While in the Arabic text and other translations it was found to have a positive SP with a meaning related to the emphasis on the message of the Quran, in Saheeh’s translation, it was found to acquire a negative SP through association with collocates that refer to the shunning of disbelievers. Hence, to further examine this difference, the researcher compared the LG patterns of the word ‘people’ in this research via the Log-likelihood scores of their collocations (See figure below).

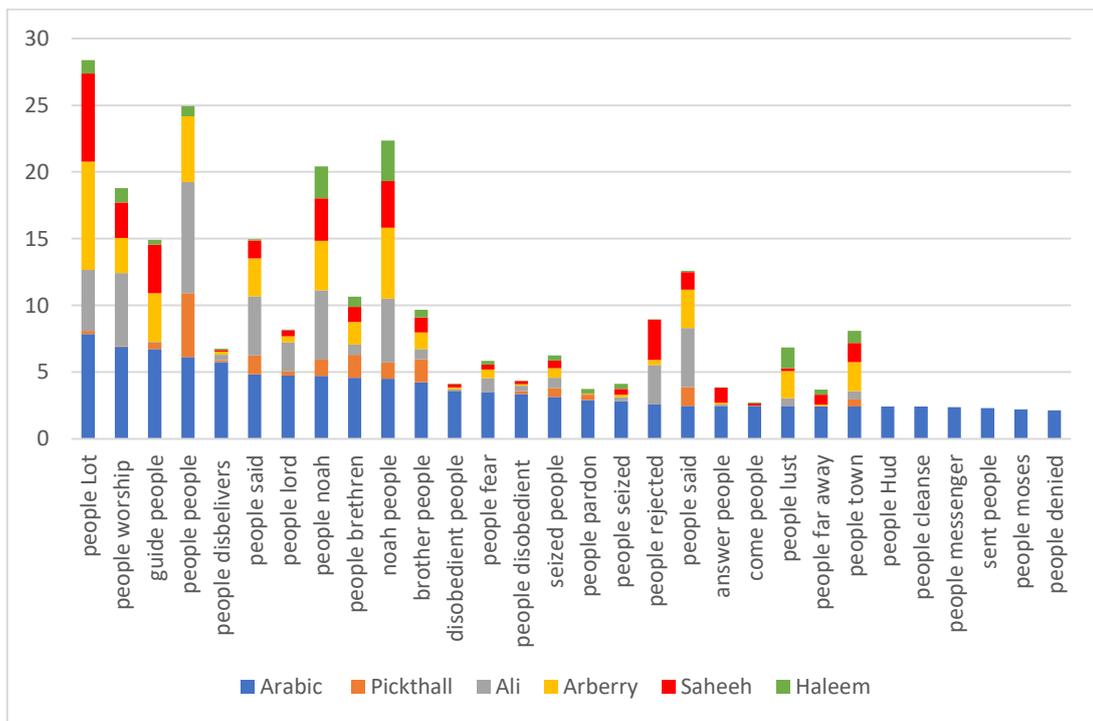


Figure 59: A comparison based on the LLR scores of the collocations of the word ‘people’

The figure above reveals that the LLR association measure highlighted more associations to the word ‘people’ in Saheeh than it did for the other translations in this research, which resulted in a difference in the SP results. This can be attributed to two obvious reasons: the first of which is that this term has the highest relative frequency in Saheeh’s translation

(12.75 while in the Arabic text, it is 6.84) compared to the other corpora²⁵², and it has a greater tendency in Saheeh’s corpus than it does in other corpora to be associated with the words which appear in the description of the previous people who denied and disbelieved the message of God delivered to them by the His messengers (e.g., people of Lot and people of Noah). Hence, it can be said that a higher or lower relative frequency can potentially be an underlying and covert factor which may affect the criterion of accuracy in the representation of SP in translation.

SP and Consistency

This research also portrays SP as a cohesive lexical device that is employed as a linguistic tool through *consistency* of terminology in translation. In this regard, it agrees with Newmark (1988, p.192), who asserts that consistency or uniformity of terminology is one of the characteristics of any good translation. Therefore, discrepancies of SP of natural phenomena in the Qur’an can be said to be caused by the presence of polysemous words in the translation of natural phenomena. Therefore, discrepancies in the translations of SP of natural phenomena in the Qur’an can be said to be caused by the presence of polysemous words in the translation of nature terms. For example, the same SP can be represented by two variants of the same nature term. The example below illustrates the possible inconsistency of terminology in the translation of several nature terms with the same SP (evaluative and discourse prosody). The words that are imprinted in bold are the nodes, while the underlined ones are the collocates based on the results of this research.²⁵³

Table 42: Examples of the inconsistency of terminology in the English translation of nature terms in the Qur’an

<i>Translation</i>	<i>Nature term</i>	<i>Translated verses</i>	<i>SP</i>
Haleem	earth	29 20 Say, ‘ <u>Travel</u> throughout the earth and <u>see</u> how He brings life into being: and He will bring the next life into being. God has power over all things.	Emphasis on the message of the Qur’an, positive
		30 42 Say, ‘ <u>Travel</u> through the land , and <u>see</u> how those before you met their end— most of them were idolaters.	
		47 10 Have they not <u>travelled</u> the earth and <u>seen</u> how those before them met their end? God destroyed them utterly: a similar fate awaits the disbelievers.	Glorifying of God, positive

252 The results in Section 5.2.

253 Results of the collocations of the Python-processed texts and their statistical analysis based on Evert (2008).

<i>Translation</i>	<i>Nature term</i>	<i>Translated verses</i>	<i>SP</i>
		40-82 Have they not <u>travelled</u> through the land and <u>seen</u> how those who lived before them met their end? They were more numerous than them, stronger than them, and made a more impressive mark on the land, yet what they achieved was of no use to them at all.	
Pickthall	mountains	15 19 And the earth have We spread out, and placed therein <u>firm hills</u> , and caused each seemingly thing to grow therein. 77 27 And made therein mountains standing <u>firm</u> , lofty (in stature); and provided for you water sweet (and wholesome)?	Glorifying of God, positive
Haleem	fountains	47 12 God will <u>admit</u> those who believe and do good deeds to <u>Gardens</u> graced with <u>flowing streams</u> ; the disbelievers may take their fill of pleasure in this world, and eat as cattle do, but the Fire will be their home 51 15 The <u>righteous</u> will be in <u>Gardens</u> with <u>flowing springs</u> .	Afterlife reward, positive
Ali	sea	31 31 <u>Seest</u> thou not that the <u>ships sail</u> through the ocean by the Grace of Allah?- that He may show you of His <u>Signs</u> ? Verily in this are Signs for all who constantly persevere and give thanks. 45 12 It is Allah Who has <u>subjected</u> the sea to you, that <u>ships</u> may <u>sail</u> through it by His command, that ye may seek of his Bounty, and that ye may be grateful. 42 32 And among His <u>Signs</u> are the <u>ships</u> , smooth-running through the ocean , (tall) as mountains.	Glorifying of God, positive
Haleem	township	15 4 Never have We <u>destroyed</u> a community that did not have a set time; 22 45 How many towns steeped in <u>wrongdoing</u> We have <u>destroyed</u> and left in total ruin; how many deserted wells; how many lofty palaces!	Present life punishment , negative

Table 42 illustrates examples of the translators' inconsistent use of the terminology of nature terms which causes variances in the counts of frequency, and henceforth collocation and the SP of natural phenomena in the Qur'an. The variances between the ST and the five TTs can be attributed to the use of semantic variances of some nature terms that are

polysemous or synonymous yet have the same SP. The translators' use of these variants challenged the statistical analysis in this research to find the collocations of nature in these corpora and confirmed the claim that the ambiguity of a polysemous word in translation studies is also a corpus-linguistic problem that could affect the counts and the exploration of textual, linguistic features (Olohan, 2004, p. 16). Despite the efforts of the researcher to disambiguate the nature terms before the statistical analysis, there were still collocations which Python did not capture due to the polysemous choice of words which the researcher did not observe at the time. The fact that this affects the results of SP as a cohesive lexical device, which when taken into consideration maintains consistency in the text, means it affects the consistency of the representation of the theme of natural phenomena in the Qur'an.

Finally, it should be stressed that this research does not assume that the translators' choices in the translation of nature terms are inaccurate; interpretations of the Qur'an are known to vary, and it is understood that a particular translator may choose one interpretation over another. Moreover, although inconsistency in some of the translations of nature terms can be problematic in the judgment of uniformity, it does not affect the judgment of this research on the quality of the translation. It is merely a suggestion that a translator should observe SP as a vital tool for achieving the closest congruency to the source text. With the understanding that SP is a hidden meaning that is uncovered as the text unfolds and collocations occur, the translator will acquire the broader picture of the most frequent words in the text, just as a corpus linguist can decide on the importance of concepts through frequency and collocation. This discussion shows that SP is just as important to a translator as it is to a corpus linguist. If a translator is to achieve the representation of the meaning of a text which is closest to the original, he/she should bear in mind that hidden meanings of a text can be uncovered via collocation and SP analysis. Therefore, he/she can in practice, analyse the text in this fashion before commencing the translation to achieve the desired effect that the source text unveils through these stylistic features.

5.4 Summary

This chapter reported the results of this corpus-based research to explore natural phenomena in the Qur'an and evaluate five translations of it in regard to their congruency with the representation of SP of nature in the Qur'an. The findings are based on the quantitative and qualitative analyses of datasets of SP of nature in the Qur'an. The results

of the quantitative analysis covered the frequencies of natural phenomena terms; verified statistical collocations of natural phenomena in the Qur'an; and finally, the evaluative and discourse prosodies of nature in the six texts –the Qur'an and its five chosen translations. On the other hand, samples of the qualitative findings covered the annotated datasets of SP in the Qur'an and the quantitative (but not computational) results of the evaluation of the translations based on the congruency with SP of nature in the Arabic Qur'an. The final section of this chapter presented the answers to the research questions and shed light on SP as a linguistic feature with two roles in two fields of study. In corpus linguistics, SP is a collocational phenomenon and a tool for cohesion to convey the network of meanings in the exploration of literary and religious texts. In translation studies, SP analysis can be used as a tool that can safeguard accuracy and consistency during the process of translation. To conclude, this chapter has put forward the importance of SP in conveying meaning in the Qur'an, on the one hand, and illustrated its implementation with LG patterns in reflecting differences in the evaluation of translation, on the other.

Chapter 6 Conclusion

This chapter concludes this research, which explored the SP of natural phenomena in the Qur'an with an attempt to evaluate the translations of the Qur'an based on congruency with the Arabic representation of the SP of natural phenomena as it appears in the Qur'an. Firstly, it provides some insights into its results as a corpus-based study. It then discusses the challenges of and reflections on the experience of this research, and, lastly, presents suggestions for future studies.

6.1 Findings and implications

This section will review the results discussed in the previous chapter to highlight the implications that each of these findings contributes to several fields of study: corpus linguistics, translation studies, and Qur'anic studies. On the theoretical level, this research has found that it is useful to align the methodology of this study to a theoretical framework that can account for collocation and SP as a collocational phenomenon. The theoretical framework applied (i.e., a combination of theories of collocation, functional lexico-grammar, and semantic prosody), was useful for both the quantitative and qualitative analysis of this research. On the quantitative level, the syntagmatic (the company a word keeps - i.e., collocational) examination of natural phenomena in the Qur'an produced a list of bigrams, which were analysed within the context of lexico-grammar. Hence, the contribution to corpus linguistic research made by this thesis involved bringing to light the importance of LG patterns and relating these to the exploration of SP, a finding which confirms claims by several previous researchers such as Partington et al. (2013) that it is useful to examine LG patterns when exploring SP. This theoretical framework can be utilised for future corpus-based studies on SP and is very useful in describing collocations. In essence, studies on SP can adopt the use of LG patterns to account for collocation in their qualitative analysis to interpret the quantitative analysis of collocations (e.g., bigrams).

Moreover, the use of concordance to observe SP was further expanded in this research in two ways: in the exclusive use in the investigation of the first 30 lines of only statistical collocations, and the use of the eight LG patterns to uncover the SP of natural phenomena in the Qur'an. This was found to be very efficient. The fact that the search was focused on statistical collocates (i.e., not combinations occurring by chance) meant that there are not any highlighted coincidental co-occurrences of natural phenomena in the Qur'an. The use

of state-of-the-art visual representations of collocations in concordances in Sketch Engine and the addition of *GraphColl* in LancsBox was also useful for exploring collocations of nature in the Qur'an. Once concordance lines were copied into Microsoft Excel and filtered to include only statistical collocates of nature in the Qur'an, the collocate sets were visually depicted, and the allocation of SP was facilitated.

In addition, the mixed approach of quantitative and qualitative analyses in this research was shown to be useful in the investigation of SP in natural phenomena. The statistical analysis of collocation with its NLP and mathematical operations provided reliable results of statistical collocations, which were in turn used in the qualitative analysis to explore SP of nature in the Qur'an and reveal its communicative purpose. Finally, the use of quantitative measurements in allocating SP (evaluative and discourse) to nature terms based on the highest percentage of meaning in their collocations provided a means to measure congruency of the five chosen translations of the Qur'an in this research.

Its shedding of light on the importance of SP in corpus linguistics and translation studies was another merit of this research; it confirmed findings of previous research which explored SP as a collocational phenomenon and cohesive device, and that can uncover subliminal meaning and potential irony. To gain an understanding of SP, this research stresses the importance of using a mixed approach to find collocations statistically; this should be followed by qualitative interpretations of concordance and LG patterns to highlight the covert meanings of the text.

Beyond this, the significance of SP in translation studies was demonstrated in the evaluation of the English renderings of the Qur'an. The discrepancies in the representation of SP of natural phenomena in English translations of the Qur'an, as seen in the parallel corpora in this research, are evidence of the importance of the need to be aware of this collocational phenomenon. To put it differently, without judging the quality of the translations, this research showed how accuracy and consistency could affect the representation of the hidden meanings of the source text. In this regard, this research has indicated the importance of SP for translation studies, based on which it is recommended that a translator should be aware of the SP of the critical concepts in a source text before translating it. This can be achieved in practice through a corpus-based exploration of the terms that represent these major themes to observe their frequencies and collocations and

taking note of the SPs that each one of them acquires through collocation. In this manner, SP could be utilised as a tool to achieve congruency in translation.

In relation to Qur'anic studies, this study has produced a commentary (See Appendix G), which discusses the five most prominent discourse prosodies of natural phenomena in the Qur'an. It provides examples of verses in which these meanings were uncovered and linguistically analysed via frequency, collocation (including LG patterns), and SP. It also refers to Islamic traditions and commentaries from existing literature to support the researcher's observations which describe the presence of natural phenomena in the Qur'an in the light of their discourse prosodies. This overview of the five most frequent discourse prosodies in the Qur'an can be of benefit to Qur'anic studies in that it proposes a novel approach to analysing the meanings of the Qur'an using state-of-the-art NLP and a corpus-based methodology.

On another practical level, this research has produced several linguistic resources that can be useful for future research. They are:

1. An ontology of natural phenomena in the Qur'an designed with the aid of the Arabic Qur'anic Corpus and existing literature on nature as a Qur'anic theme.
2. A list of Arabic stop-words (e.g. function words) that is exclusive to the Qur'an. None of the available lists of Arabic stop-words, although these are very useful for Arabic NLP research on the analysis of tweets on Twitter, for example, was found to be sufficient for the use of exploratory research on the Qur'an. Hence, this research presents a compiled list of stop-words which includes all the function words in the Qur'an.
3. Six SP tagged datasets of the Qur'an and its translations (where each line comprises a verse with nature term or terms and its evaluative and discourse prosodies).
4. The LG patterns of nature in the Qur'an (e.g., *N+N* (noun +noun), *N+V* (noun +verb), etc.), which can be used for further exploration on the theme of nature in the Qur'an.
5. A machine-readable English-language Tafsīr (one of the traditional exegeses of interpretation of the Qur'an) by Al-Jalalayn, which was not previously before, and can be utilised in future studies on the meanings of the Qur'an using a corpus-based methodology.

To reiterate, this research demonstrated the effectiveness of the implementation of a corpus-based method combined with an NLP approach via the NLTK libraries in Python for finding collocations and exploring SP as a collocational phenomenon. It provided insights which can be of use in corpus linguistics, translation studies, and Qur'anic studies.

6.2 Challenges and reflections

The challenges of this research were related to the nature of the study, as well as the application of the methodology. This research, being interdisciplinary in nature, posed a challenge in maintaining a balance between two fields of study: computational and corpus linguistics and descriptive translation studies. Another challenge was learning the Python programming language and understanding and using a variety of language processing toolkits to implement the NLP approach in the first phase. Having a background mainly in applied linguistics, the researcher required training in state-of-the-art NLP with Python for automated textual analysis and data visualisation. The experience of learning this type of approach has been invaluable and will benefit the researcher in her future research and professional life. Moreover, the journey of searching for a theory which accounts for collocation and the methodology which best suits the aims of this research was a fruitful one for it opened the researcher's mind to developments in the field of corpus linguistics and possible further avenues of research in this domain. Finally, reading the Qur'an in particular and generally any literary text with a new eye which navigates through collocations to find SP was a beautiful addition to the researcher's development and mental growth throughout her PhD study.

6.3 Suggestions for future studies

The aforementioned linguistic resources and insights produced by this research can be useful for future studies. For example, the evidence provided for the importance of SP as a measure of accuracy and consistency both for corpus-linguistic and translation studies can open up new avenues of corpus-based research in both fields. To be more specific, this evidence can be further explored, developed and applied in the context of prescriptive translation studies for the translation of the Qur'an. In practical terms, this research suggests that translators should conduct a corpus-based exploration of a source text before translating it to achieve better accuracy and consistency. Another suggestion for future studies is the application of the theoretical framework of this thesis to other corpus linguistic studies to explore collocations generally and account for their presence via the

use of a linguistic theory to uncover SP. Also, this study could be revisited with the alteration of some of the choices in this research, such as aggregation, the employed association measure, and stopword-removal. A further suggestion for future studies is to add more natural phenomena to the proposed list of nature related words, since the scope of this study covered an exhaustive list which can be considered a working list too. Therefore, when one would find that a natural phenomenon word such as *خمر* *khamr* 'wine' is missing, he or she can add it to the list that is proposed in this research. Then, he or she would apply the same methodology to yield up more patterns of nature in the Qur'an and provide more insights on the SP of this theme.

Similarly, the mixed approach to the analysis of collocation can also be of use to future studies on collocation and SP as two components of Sinclair's Extended Lexical Unit (2004a). The same pre-methodology and methodology tasks can also be applied to other themes (e.g., Fazular Rahman's major themes, 2009) prevalent in the Qur'an to build conceptual ontologies and to explore their SP and uncover the subliminal meanings relevant to them. Likewise, the LG patterns of the theme of nature in the Qur'an and its five translations can be used in future evaluative studies on the translation of the Qur'an. Just as the analysis of SP of natural phenomena in the Qur'an was the basis on which the ranking of the most congruent of the translations in this research, the LG features of natural phenomena-related words could be compared for the same corpora, and judgments could be made regarding their similarities and differences. Finally, Haleem's translation, which was found to be the most congruent of the translations in its representation of the hidden meanings of nature (i.e., semantic prosody of nature in the Arabic Qur'an), can be employed in a further study to produce a bilingual dataset for SP of nature as a Qur'anic theme, and the dataset can be used to develop a bilingual project that explores nature as a Qur'anic theme which can, in turn, benefit the everyday reader of the Qur'an and help him/her to understand its meanings.

Bibliography

- Aarts, B., & McMahon, A. (Eds.). 2008. *The handbook of English linguistics*. John Wiley & Sons.
- Aarts, J. 2002. Does corpus linguistics exist? Some old and new issues. *Language and Computers*. **40**, pp.1-18.
- Aarts, J. and Meijs, W. 1984. *Corpus linguistics: recent developments in the use of computer corpora in English language research*. Amsterdam: Rodopi.
- Aarts, J. and Meijs, W. 1986. *Corpus linguistics II: new studies in the analysis and exploitation of computer corpora (Vol. 56)*. Amsterdam: Rodopi.
- Aarts, J. and T. 1982. Grammars and intuitions in corpus linguistics. In: Johansson, S. ed. *Computer corpora in English language research*. Bergen: Norwegian Computing Centre for Humanities, pp.66–84.
- Abbas, A. 1991. *Tafsir Ibnu Abbas*. Beirut: Mu'asasah Al Kutub Ats Tsaqafiyah.
- Abbas, I. and Al-Firuzabadi, M. 2007. *Tanwīr al-Miqbās min Tafsīr Ibn'Abbās*. Abu Thahir ibn Ya'qub al-Fayruzzabadi (ed). Beirut: Daar el-Fikr.
- Abbas, N. 2009a. *Quran 'search for a concept' tool and website*. MSc Research Thesis, School of Computing, University of Leeds.
- Abbas, N. 2009b. *Qur'any: a tool to search for concepts in the Qur'an*. [Online]. [Accessed 10 July 2017]. Available from: [<http://Qur'anytopics.appspot.com/>]
- Abbas, N., Luluh, A., Al-Khalifa, H., Alqassem, Z., Atwell, E., Dukes, K., Sawalha, M. and Sharaf, A. 2013. Unifying linguistic annotations and ontologies for the Arabic Qur'an. In: *Proceedings of WACL-2 Second Workshop on Arabic Corpus Linguistics: The University of Leeds*.
- Abdul Wadud, S. 1971. *Phenomena of nature and the Qur'an*. Lahore: Sayed Khalid Publishers.
- Abdul Wadud, S. 1996. *Gateway to the Qur'an*. Lahore. Sayed Khalid Publishers.
- Abdul Wadud, S. 1998. *The heavens, the earth, and the Qur'an*. Lahore: Sayed Khalid Publishers
- Abdullah, Y. 2010. *Assessing the translations of collocation in the glorious Qur'an into English*. Adab AL Rafidayn. 57, pp.724-51.
- Abdul-Rahman, S. 2009. *Tafsir Ibn Kathir [the translation]*. MSA Publication Limited.
- Abdul-Raof, H. 2001. *The Qur'an outlined: outline, theme and text*. Lincom Publishing Press.

- Abdul-Raof, H. 2005. Pragmalinguistic forms in cross-cultural communication: contributions from Qur'an translation. *Intercultural Communication Studies*, **14**(4), pp. 115-29.
- Abed, Q. 2015. *Ontology-based approach for retrieving knowledge in Al-Qur'an*. Thesis, Universiti Utara Malaysia.
- Abu-Chacra, F. 2007. *Arabic: an essential grammar*. Routledge.
- Adolphs, S. 2006. *Introducing electronic text analysis: a practical guide for language and literary studies*. London: Routledge.
- Adolphs, S. and Carter, R. 2002. Points of view and semantic prosodies in Virginia Woolf's *To the Lighthouse*. *Poetica*. **58**, pp.7-20.
- Ahmad, N., Bennett, B., and Atwell, E. 2016. Semantic-based ontology for Malay Qur'an reader. In: *IMAN'2016 4th International Conference on Islamic Applications in Computer Science and Technologies*. IMAN'2016. Khartoum, Sudan. pp. 20-2
- Ahmad, O., Hyder, I., Iqbal, R., Murad, M., Mustapha, A., Sharef, N. and Mansoor, M. 2013. A survey of searching and information extraction on a classical text using ontology-based semantics modelling: a case of Qur'an. *Life Science Journal*. **10**(4), pp.1370-77.
- Aijmer, K., and Stenström, A. 2004. *Discourse patterns in spoken and written corpora (Vol. 120)*. Amsterdam: John Benjamins.
- Al Ghamdi, S. 2015. *Critical and comparative evaluation of the English translations of the near-synonymous divine in the Qur'an*. Ph.D. thesis, University of Leeds.
- Al-Fairuz, M. 1991. *Tanwīr al-miqbās min tafsīr Ibn'Abbās*. Lebanon, Beirut: Dar Al-Kutub Al-Ilmiyah.
- Al-Hilali, M. and Khan, M. 1983. *Translation of the meanings of the noble Qur'an in the English language*. Madina, Saudi Arabia: King Fahad Complex.
- Ali, A. 1939. *The glorious Kur'an: translation and commentary*. Call of Islam Society.
- Ali, A. 1987. *The holy Qur-ān: English translation of the meanings and commentary*. King Fahd Holy Qur-ān Printing Complex.
- Al-Jabari, A. 2008. *Reasons for the possible incomprehensibility of some verses of three translations of the meaning of the holy Qur'an into English*. Ph.D. thesis, University of Salford.
- Al-Kabi, M., Kanaan, G., Al-Shalabi, R., Nahar, K. and Bani-Ismail, B. 2005. Statistical classifier of the holy Qur'an verses (Fatihah and Yaseen chapters). *Journal of Applied Sciences*. **5**(3), pp.580-83.
- Al-Khalidi, S. 1996. *Al- Tafsīr wa al-Ta 'wīl fi al-Qur'an*. Amman: Dar al-Nafais

- Al-Khalifa, H., Al-Yahya, M., Bahanshal, A. and Al-Odah, I. 2009. SemQ: a proposed framework for representing semantic opposition in the holy Qur'an using semantic web technologies. In: *2009 International Conference on the Current Trends in Information Technology (CTIT): IEEE*, pp.1-4.
- Al-Nasser, S.A. and Khashan, R.A. 2008. The collocation of mubeen in the holy Qur'an. *Journal of the College of Arts in the University of Basrah*. **45**, pp.1-16.
- Alonso, A., Millon, C., & Williams, G. 2011. Collocational networks and their application to an E-Advanced Learner's Dictionary of Verbs in Science (DicSci). In: Kosem, I. & Kosem, K. (Eds.) *Electronic Lexicography in the 21st Century: New Applications for New Users: Proceedings of eLex 2011, Bled, 10-12 November 2011* (pp. 12–22).
- Alqahtani, M. and Atwell, E. 2018. Developing bilingual Arabic-English ontologies of Al-Qur'an. In: *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR): IEEE*, pp.96-101.
- Alqahtani, M., and Atwell, E. 2015. A review of semantic search methods to retrieve information from the Qur'an corpus. *Corpus Linguistics 2015*, pp.7–9.
- Al-Qinai, J. 2012. Convergence and divergence in the interpretation of Qur'anic polysemy and lexical recurrence. *Acta Orientalia*. **65**(1), pp.71-92.
- Alrehaili, S. and Atwell, E. 2014. Computational ontologies for semantic tagging of the Qur'an: a survey of past approaches. In: *LREC 2014 Proceedings. Ninth International Conference on Language Resources and Evaluation (LREC'14), May 2014, Reykjavik, Iceland*. European Language Resources Association. pp. 26-31. ISBN 978-2-9517408-8-4
- Alrehaili, S. M. 2012. *The chronology of the texts in the holy Quran according to NLP*. Unpublished MSc Dissertation, School of Computing, University of Leeds.
- Alromima, W., Moawad, I., Elgohary, R. and Aref, M. 2015. Ontology-based model for Arabic lexicons: an application of the place nouns in the holy Qur'an. In: *Computer Engineering Conference (ICENCO), 2015 11th International: IEEE*, pp.137-43.
- Al-Sa'ad, A. 1995. *Dilalat al-'ansaq albina'iah fil tarkeeb al qurāni (The semantics of the structural patterns in the holy Qur'an)*. Ph.D. thesis, University of Basrah.
- Alserougij, M. and Mutawalli Sha'rawi, M. 2009. *The miracles of the Qur'ān*. United Kingdom: Dar Al Taqwa.
- Alshahrani, H. and Brierley, C. 2017. An overview of natural phenomena in the Qur'an. In: *Postgraduate Research Conference May 5th, 2017. University of Leeds, Leeds*
- Alshahrani, H. and Brierley, C. 2018. A corpus-based study of semantic prosody in the Qur'an. In: *The 4th College of Languages and Translation (COLT) Scientific Research Symposium, King Saud University, 3rd of October 2018, Riyadh*.

- Alshahrani, H. and Brierley, C. 2018. Semantic prosody in nature as a Qur'anic theme. In: *Corpora and Discourse International Conference, 22nd to 24th June 2018, Lancaster University, Lancaster.*
- Alshahrani, H. and Brierley, C. 2018. The root-based disambiguation of words in the annotation of semantic prosody of nature as a Qur'anic theme. In: *The Workshop on Computational Approaches to Morphologically Rich Languages, 3 July 2018, University of Leeds, Leeds.*
- Alshahrani, H., Brierley, C. and Dickins, J. 2018. The functional lexico-grammar of semantic prosody in the Qur'an. In: *Symposium on Corpus Approaches to Lexicogrammar, 16 June 2018, Ormskirk. Edge Hill University.*
- Al-Sofi, K., Maros, M. and Bakr, K. 2014. Semantic prosody in the Qur'an. *International Journal of Linguistics*. **6**(2), pp.120-31.
- Altenberg, B., & Eeg-Olofsson, M. 1990. Phraseology in spoken English: Presentation of a project. In: Aarts, J., and Meijs, W. Eds. *Theory and practice in corpus linguistics* (No. 4). Rodopi, pp.1-26.
- Al-Timen, F. A. A. 2015. An analysis of translation methods for English proverbs: literal, literary or substitution. *European Academic Research II*, **11**, pp. 14016- 14026.
- Al-Ubaidi, M. 2013. Semantic prosody in the Qur'an. *QUR'ANICA-International Journal of Qur'anic Research*. **5**(2), pp.91-108.
- Al-Yahya, M., Al-Khalifa, H., Bahanshal, A., Al-Odah, I. and Al-Helwah, N. 2010. An ontological model for representing semantic lexicons: an application on time nouns in the holy Qur'an. *Arabian Journal for Science and Engineering*. **35**(2), p.21.
- Atwell, E., Brierley, C., Dukes, K., Sawalha, M. and Sharaf, A. 2011. An artificial intelligence approach to Arabic and Islamic content on the internet. In: *Proceedings of NITS 3rd National Information Technology Symposium: Leeds*, pp.1-8.
- Ayetiran, E. and Agbele, K. 2016. An optimized Lesk-based algorithm for word sense disambiguation. *Open Computer Science*. **8**(1), pp.165-172.
- Badawi, J., 2012. *Muslims and non-Muslims relations: reflections on some Qur'anic texts*. [Online]. Available from: <http://www.islamawareness.net>
- Bahumaid, S., 2006. Collocation in English-Arabic translation. *Babel*. **52**(2), pp.133-52.
- Baker, M. 1993. Corpus linguistics and translation studies: Implications and applications. In: Francis, G. Baker, M., and Tognini Bonelli, El. eds. *Text and technology: in honour of John Sinclair*, pp. 233-52. Amsterdam: John Benjamins.
- Baker, M. 1996. Corpus-based translation studies - the challenges that lie ahead. In: Somers, H. ed. *Terminology, LSP and translation*. Amsterdam/Philadelphia: John Benjamins, pp 175-86.
- Baker, M. 2004. The treatment of variation in corpus-based translation studies. *Language matters*, **35**(1), pp. 28-38.

- Baker, P. 2016. The shapes of collocation. *International Journal of Corpus Linguistics*, **21**(2), pp.139-64.
- Baker, P., Hardie, A., and McEnery, A. 2006. *A glossary of corpus linguistics*. Edinburgh, Scotland: Edinburgh University Press.
- Barlow, M. 2011. Corpus linguistics and theoretical linguistics. *International Journal of Corpus Linguistics*, **16**(1), pp. 3-44.
- Barnbrook, G., Mason, O., & Krishnamurthy, R. 2013. *Collocation: applications and implications*. Springer.
- Bartsch, S. 2004. *Structural and functional properties of collocations in English: a corpus study of lexical and pragmatic constraints of lexical co-occurrence*. Gunter Narr Verlag Tübingen.
- Bartsch, S. and Evert, S. 2014. Towards a Firthian notion of collocation. *Vernetzungsstrategien, Zugriffsstrukturen und automatisch ermittelte Angaben in Internetwörterbüchern*. **2**, pp.48-61.
- Basile, P., Caputo, A. and Semeraro, G. 2014. An enhanced lesk word sense disambiguation algorithm through a distributional semantic model. In: *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pp.1591-1600.
- Bauer, L. 1983. *English word-formation*. Cambridge: Cambridge University Press
- Begagić, M. 2013. Semantic preference and semantic prosody of the collocation make sense. *Jezikoslovlje*, **14**(2-3), pp. 403-16.
- Bell, R., 1987. *The Qur'an, T. and T.* Edinburgh: Clark.
- Bell, R., Bosworth, C. and Richardson, M. 1991. *A commentary on the Qur'an*. The University of Manchester.
- Benson, M. 1985. Collocations and idioms. In Ilson, R. ed. *Dictionaries, lexicography and language learning*. Oxford: Pergamon Press.
- Benson, M., Benson, E. and Ilson, R., 1986. *The BBI combinatory dictionary of English: A guide to word combinations*. Amsterdam: John Benjamins.
- Bentivogli, L. and Pianta, E., 2005. Exploiting parallel texts in the creation of multilingual semantically annotated resources: the MultiSemCor Corpus. *Natural Language Engineering*. **11**(3), pp.247-61.
- Berber Sardinha, T. 2000. Semantic prosodies in English and Portuguese: a contrastive study. *Cuadernos de filología inglesa*, vol. **9**, 2000.
- Bernard, H. 1996. Qualitative data, quantitative analysis. *CAM Journal*. **8**(1), pp.9-11.

- Bernard, H. and Ryan, G. 1998. Text analysis. In: Bernard, H. R., & Gravlee, C. C. eds. *Handbook of methods in cultural anthropology*. Rowman & Littlefield, pp.595-645.
- Biber, D. and Conrad, S. 1999. Lexical bundles in conversation and academic prose. In: Hasselgard, H. and Oksefjell, S. eds. *Out of corpora*. Amsterdam-Atlanta GA: Rodopi, pp.181-90.
- Biber, D., Conrad, S. and Cortes, V. 2004: If you look at ...: lexical bundles in university teaching and textbooks. *Applied Linguistics* **25**, 371–405.
- Biber, D., Conrad, S. and Reppen, R. 1998. *Corpus linguistics: investigating language structure and use*. Cambridge University Press.
- Biber, D., Conrad, S., Johansson, S. and Leech, G. 1999. *Longman grammar of spoken and written English*. London: Pearson Education Limited.
- Bin Tyeer, S. R. 2016. *The Qur'an and the aesthetics of premodern Arabic prose*. Springer
- Bird, S., Klein, E. and Loper, E. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. O'Reilly Media, Inc.
- Bong, S., 2002. Debunking myths in qualitative data analysis. In: *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*. **3**(2)
- Brashi, A., 2005. *Arabic collocations: implications for translations*. The University of Western Sydney.
- Brezina, V. 2018. *Statistics in corpus linguistics: a practical guide*. Cambridge University Press.
- Brezina, V., McEnery, T. and Wattam, S. 2015. Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*. **20**(2), pp.139-73.
- Brezina, V., Timperley, M., and McEnery, T. 2018. #LancsBox v. 4.x [software]. [Online]. [Accessed 10 July 2018]. Available from: <http://corpora.lancs.ac.uk/lancsbox>
- Brierley, C., Sawalha, M., Islam, T., Dickins, J. and Atwell, E. 2018. Automatic extraction of Qur'anic lexis representing two different notions of linguistic salience: keyness and prosodic prominence. *Journal of Semitic Studies*. **63**(2), pp.407-56.
- Brown, P., Lai, J. and Mercer, R. 1991. Aligning sentences in parallel corpora. In: *the Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pp. 169-76. Berkeley, California.
- Bublitz, W. 1996. Semantic prosody and cohesive company: somewhat, somewhat predictable. *Leuvense Bijdragen: Tijdschrift voor Germaanse Filologie*, **85**(1–2), pp.1–32.
- Butler, C., 2004. Corpus studies and functional linguistic theories. *Functions of language*. **11**(2), pp.147-86.

- Campanini, M. 2013. *The Qur'an: the basics*. Routledge.
- Cartoni, B., Zufferey, S., Meyer, T., and Popescu-Belis, A. 2011. How comparable are parallel corpora? Measuring the distribution of general vocabulary and connectives. In: *Proceedings of the 4th Workshop on Building and Using Comparable Corpora: Comparable Corpora and the Web* (pp. 78-86). Association for Computational Linguistics.
- Casagrande, J., 1954. The ends of translation. *International Journal of American Linguistics*. **20**(4), pp.335-40.
- Cha, J. 2017. *Interplay of prosodic features and discourse functions in Korean conversation*. Thesis, University of California.
- Cheng, W. 2006. Describing the extended meanings of lexical cohesion in a corpus of SARS spoken discourse. In: Flowerdew, J. and Mahlberg, M. eds. Corpus linguistics and lexical cohesion (special issue). *International Journal of Corpus Linguistics*, **11**(3), pp. 325–44.
- Cheng, W. 2013. Semantic prosody. *The Encyclopaedia of Applied Linguistics*. Blackwell Publishing Ltd. DOI: 10.1002/9781405198431.wbeal1062
- Chesterman, A. 2000. A causal model for translation studies. In Olohan, M. 2017. *Intercultural faultlines*. pp.15-27.
- Coffin, C., Hewings, A. and O'Halloran, K. 2014. *Applying English grammar: corpus and functional approaches*. Routledge.
- Croft, W. 2001. *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Cruse, D. A. 1986. *Lexical semantics*. Cambridge University Press.
- Cruse, D., 2002. Aspects of the micro-structure of word meanings. In: Ravin, Y. and Leacock, C. eds. *Polysemy: theoretical and computational approaches*. Oxford: Oxford University Press, pp.30-51.
- Crystal, D., 2011. *A Dictionary of linguistics and phonetics*. John Wiley and sons.
- Daille, B., Gaussier, E. and Langé, J. 1994. Towards automatic extraction of monolingual and bilingual terminology. In: *Proceedings of the 15th conference on Computational Linguistics-Volume 1: Association for Computational Linguistics*, pp. 515-21.
- Dam-Jensen, H. and Zethsen, K., 2008. Translator awareness of semantic prosodies. Target. *International Journal of Translation Studies*. **20**(2), pp.203-12.
- Dam-Jensen, H., and Zethsen, K. 2007. Pragmatic patterns and the lexical system: a reassessment of evaluation in language. *Journal of Pragmatics*, **39**(9), 1608–1623.
- Deedat, A. 2012. *The Choice: Islam and Christianity*. Adam Publishers.

- Dickins, J. 2014. Associative meaning and scalar implicature: a linguistic-semiotic account. *Linguistica Online*. **16**, pp.1-38.
- Dickins, J. 2019. Types of connotative meaning, and their significance for translation. In: Faiq, S., ed. *Discourse in translation*. London and New York: Routledge, pp. 135-62.
- Dickins, J., Hervey, S., and Higgins, I. 2017. *Thinking Arabic translation: a course in translation method: Arabic to English*. Routledge.
- Djamil, A. 2012. Oceanic verses of the Qur'an: their semantics, ontology and categorization. In: *The Proceeding of The 2nd World-ISLAM; World Congress on Islamic Systems 2012*. Malaysia.
- Douglas, A. 1973. RASC Papers-Al-Biruni, Persian scholar, 973-1048. *Journal of the Royal Astronomical Society of Canada*. **67**, pp. 209-10.
- Downe-Wamboldt, B. 1992. Content analysis: method, applications, and issues. *Health care for women international*. **13**(3), pp.313-21.
- Duhainah, H. 2007. *Linguistic collocations and their significance in determining the semantics of the holy Qur'an: a theoretical and applied study*. Arabic, PhD dissertation, Al-Azhar University, Cairo, Egypt.
- Dukes, K. 2011. *The Qur'anic Arabic Corpus*. Online. Accessed: October 2016 onwards. Available from: <http://corpus.Qur'an.com>
- Dukes, K., and Habash, N. 2010. Morphological annotation of Qur'anic Arabic. In: *Proceedings of the Language Resources and Evaluation Conference (LREC), Malta, 2010*, pp.2530-2536.
- Dukes, K., Atwell, E. and Habash, N. 2013. Supervised collaboration for syntactic annotation of Qur'anic Arabic. *Language Resources and Evaluation*. **47**(1), pp.33-62.
- Dunning, T. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, pp.61-74.
- Dweik, B. and Abu Shakra, M. 2011. Problems in translating collocations in religious texts from Arabic into English. *Linguistics Journal*. **5**(1).
- Ebeling, J. and Ebeling, S. O. 2013. Semantic prosody in a cross-linguistic perspective. In: Huber, M. and Mukherjee, J. eds. *Corpus linguistics and variation in English: focus on non-native Englishes*, **13**, [online]. Available from: [<http://www.helsinki.fi/varieng/series/volumes/13/ebeling/>].
- Ebeling, S. 2013. *Patterns in contrast*. Amsterdam: John Benjamins.
- Ebeling, S. 2014. Cross-linguistic semantic prosody: the case of 'commit', 'signs of', and 'utterly' and their Norwegian correspondences. *Oslo Studies in Language*. **6**(1).

- Elewa, A. H. 2004. *Collocation and synonymy in classical Arabic: A corpus-based study*. Ph.D. thesis, University of Manchester: UMIST.
- Embong, A. and Hanapi, M. 2017. The Application methods of the holy Qur'an's interpretations in Islamic-related research. *International Journal of Academic Research in Business and Social Sciences*. 7(2), pp.521-35.
- Evert, S. 2005. *The statistics of word cooccurrences: word pairs and collocations*. PhD thesis, Institut für maschinelle Sprachverarbeitung Universität Stuttgart.
- Evert, S. 2008. Corpora and collocations, In: Lüdeling, A. and Kytö, M. eds. *Corpus linguistics: an international handbook*. 2. Berlin: Mouton de Gruyter, pp.1212-1248.
- Faizal, P. R. M., Ridhwan, A. A. M., & Kalsom, A. W. 2013. The entrepreneurs characteristic from al-Quran and al-Hadis. *International Journal of Trade, Economics and Finance*, 4(4), p. 191.
- Farghaly, A. and Shaalan, K. 2009. Arabic natural language processing: challenges and solutions. *ACM Transactions on Asian Language Information Processing (TALIP)*. 8(4).
- Farooqi, M. 1992. *Plants of the Qur'an*. Sidrah Publishers.
- Faruqi, Y. M. 2007. Islamic view of nature and values: could these be the answer to building bridges between modern science and Islamic science. *International Education Journal*, 8(2), pp.461-69.
- Fazlur Rahman, F. 2009. *Major themes of the Qur'an (2nd ed.)*, with a new forward by Ebrahim Moosa. Chicago: IL: University of Chicago.
- Fellbaum, C. 1998. *WordNet: an electronic lexical database*. Cambridge. MA: MIT Press.
- Fernandes, L. 2006. Corpora in translation studies: revisiting Baker's typology. *Fragmentos: Revista de Língua e Literatura Estrangeiras*, 30.
- Firth, J.R. 1951. Modes of meaning. In: Firth, J.R. ed. *Papers in linguistics 1934-1951*. London: Oxford University Press, pp. 118-49.
- Firth, J.R. 1957. *Papers in linguistics, 1934-1951*. London: Oxford University Press.
- Flowerdew, J., and Mahlberg, M. 2009. *Lexical cohesion and corpus linguistics (Vol. 17)*. John Benjamins Publishing.
- Gablasova, D., Brezina, V. and McEnery, T. 2017. Collocations in corpus-based language learning research: identifying, comparing, and interpreting the evidence. *Language Learning*. 67(S1), pp.155-79.
- Gabrielatos, C. 1994. Collocations: pedagogical implications, and their treatment in pedagogical materials. Unpublished.

- Gabrielatos, C., Davies, M., Rayson, P., Hunston, S., and Danielsson, P. eds. 2007. If-conditionals as modal colligations: a corpus-based investigation. *Paper presented at 4th Corpus Linguistics Conference, United Kingdom, 28/07/07 - 30/07/07.*
- Garside, R., Leech, G., and McEnery, A. 1997. *Corpus annotation – linguistic information from computer text corpora.* London and New York: Longman.
- Gavioli, L. 2005. *Exploring corpora for esp learning.* John Benjamins Amsterdam.
- Giaretta, P. and Guarino, N. 1995. Ontologies and knowledge bases towards a terminological clarification. *Towards very large knowledge bases: knowledge building and knowledge sharing.* **25**(32), pp.307-17.
- Gilquin, G. and Gries, S. 2009. Corpora and experimental methods: a state-of-the-art review. *Corpus linguistics and linguistic theory.* **5**(1), pp.1-26.
- Gledhill, C. 2000. The discourse function of collocation in research article introductions. *English for Specific Purposes.* **19**(2), pp.115-35.
- Gledhill, C. 2011. A lexicogrammar approach to checking quality: Looking at one or two cases of comparative translation. *Perspectives on translation quality,* 9, 71-97.
- Goldberg, A.E. 2006. *Constructions at work: the nature of generalization in language.* Oxford University Press.
- Golshani, M. 2003. *The holy Qur'an and the sciences of nature: a theological reflection.* New York: Global Scholarly Publications.
- Greenbaum, S. 1974. Some verb-intensifier collocations in American and British English. *American Speech.* **49**(1-2), pp.79-89.
- Gries, S. 2006a. Corpus-based methods and cognitive semantics: the many senses of to run, In: Gries, St. and Stefanowitsch, Th. eds. *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis.* Berlin and New York: Mouton de Gruyter, pp.1-17.
- Gries, S. 2006b. Exploring variability within and between corpora: some methodological considerations. *Corpora,* **1**(2), pp. 109-51.
- Gries, S. 2008. Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics,* 13(4), pp. 403–37. DOI: 10.1075/ijcl.13.4.02gri
- Gries, S. 2009. What is corpus linguistics? *Language and linguistics compass.* **3**(5), pp.1225-1241.
- Gries, S. 2010a. Corpus linguistics and theoretical linguistics: A love–hate relationship? Not necessarily.... *International Journal of Corpus Linguistics,* **15**(3), pp. 327-43.
- Gries, S. 2010b. Dispersions and adjusted frequencies in corpora: further explorations. In: Gries, S. T., Wulff, S., and Davies, M. eds. *Corpus-linguistic applications.* Brill Rodopi, pp.197-212.

- Gries, S. 2010c. Useful statistics for corpus linguistics. In: Sánchez, A. and Almela, M. eds. *A mosaic of corpus linguistics: selected approaches*. Frankfurt am Main: Peter Lang, pp. 269-91.
- Gries, S. 2013. 50-something years of work on collocations. *International Journal of Corpus Linguistics*. **18**(1), pp.137-66.
- Gries, S. and Durrant, P., 2019. *Analyzing co-occurrence data*. Springer.
- Gruber, T. R. 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), pp. 199–220. Available from: <http://doi.org/10.1.1.101.7493>
- Guezzou, M. 2007. *Tanwīr al-miqbās min tafsīr Ibn'Abbās*. Amman, Jordan: Royal Aal al-Bayt Institute for Islamic Thought.
- Habash, M. 2001. *Mushaf al tajweed*. Syria: Dar-Al-Maarifah.
- Haleem, M. A., & Haleem, M. A. 2010. *Understanding the Qur'an: themes and style*. Bloomsbury Publishing
- Haleem, M.A. 2004. *The Qurān: a new translation*. Oxford, UK, Oxford University Press.
- Halimah, A. 2014. Translation of the holy Qur'an: a call for standardization. *Advances in Language and Literary Studies*. **5**(1), pp.122-33.
- Halliday, M. 1961. Categories of a theory of grammar. *Word*. **17**(3).
- Halliday, M. 1966. Lexis as a linguistic level. In: Halliday, M.A.K. and Robins, R.H., eds. *In memory of JR Firth*, pp. 148-62.
- Halliday, M. and Hasan, R. 1976. *Cohesion in English*. London: Longman.
- Halliday, M. and Matthiessen, C. 2004. *An introduction to functional grammar (3rd ed)*. London: Arnold.
- Hammo, B., Sleit, A., Baarah, A. and Abu-Salem, H. 2012. A computational approach for identifying Qur'anic themes. *International Journal of Computer Processing of Languages*. **24**(02), pp.189-206.
- Hamza, F. 2008. *Tafsir Al-Jalalayn* (an English translation). [Online]. Pdf available from: <http://altafsir.com>
- Hanks, P. 1996. Contextual dependency and lexical sets. *International Journal of Corpus Linguistics*, **1**(1), pp. 75-98.
- Harris, Z. 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.
- Hatim, B. 1998. Translation quality assessment: Setting and maintaining a trend. *The Translator*, **4**(1), pp. 91-100.
- Hatim, B. 2020. *Communication across cultures: the linguistics of texts in translation*. University of Exeter Press.

- Hatim, B. and Munday, J. 2004. *Translation: an advanced resource book*. Routledge.
- Hauser, D. and Schwarz, N. 2016. Semantic prosody and judgment. *Journal of Experimental Psychology: General*. **145**(7), p.882.
- Herbst, T. 1996. What are collocations: sandy beaches or false teeth?. *English Studies*, **4**, pp. 379-93.
- Herbst, T. 2011. Choosing sandy beaches—collocations, probabemes and the idiom principle. In: Herbst, T., Faulhaber, S., and Uhrig, P. eds. *The phraseological view of language. A tribute to John Sinclair*. pp.27-57.
- Hewson, L. 2011. *An approach to translation criticism: Emma and Madame Bovary in translation*. John Benjamins Publishing.
- Hoey, M. 1991. *Patterns of lexis in text*. Oxford: Oxford University Press.
- Hoey, M. 1997. From concordance to text structure: new uses for computer corpora. In *Palc. Applications in Language Corpora Proceedings*. **97**, pp. 2-23.
- Hoey, M. 2000. *A world beyond collocation: new perspectives on vocabulary teaching. Collocations*. pp.224-45.
- Hoey, M. 2004. The textual priming of lexis. In: Aston, G. and Bernardini, S. eds. *Corpora and language learners*. Amsterdam: John Benjamin, pp.21-42.
- Hoey, M. 2005. *Lexical priming: a new theory of words and language*. London, England: Routledge.
- Hoey, M. Mahlberg, M. Stubbs, M. and Teubert, E. eds. 2007. *Text, discourse and corpora: theory and analysis*, London: Continuum.
- Hofland, K. and Johansson, S. 1998. The translation corpus aligner: a program for automatic alignment of parallel texts. *Language and Computers*. **24**, pp.87-100.
- House, J. 1981. *A model for translation quality assessment*, 2nd ed. Gunter Narr Verlag.
- House, J. 1997. *Translation quality assessment: a model revisited*. Gunter Narr Verlag.
- Hsieh, H. F., and Shannon, S. E. 2005. Three approaches to qualitative content analysis. *Qualitative Health Research*, **15**(9), pp.1277-1288.
- Hunston, S. 2001. Colligation, lexis, pattern and text. In: Scott, M. and Thompson, G. eds., *Patterns of text: in honour of Michael Hoey*. Amsterdam, Netherlands: John Benjamins, pp. 13–33.
- Hunston, S. 2002. *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hunston, S. 2006. *Corpus linguistics*. The University of Birmingham Elsevier, pp.234-48.
- Hunston, S. 2007. Semantic prosody revisited. *International Journal of Corpus Linguistics*. **12**(2), pp.249-68.

- Hunston, S. 2008. Starting with the small words: patterns, lexis and semantic sequences. *International journal of corpus linguistics*. **13**(3), pp.271-95.
- Hunston, S. and Francis, G. 2000. *Pattern grammar: a corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Hunston, S. and Laviosa, S. 2006. Corpus linguistics. *Linguistics*. **7**(2), pp. 215-44.
- Hunston, S. and Thompson, G. 2000. *Evaluation in text: authorial stance and the construction of discourse*. Oxford University Press, UK.
- Ibn Kathīr, I.1983. *TafsīrAl-Qurān. Al-cAzīm: vols.1-4*. Beirut: DarAlqalam.
- Iqbal, M. O. 2013. *The Qur'an as a literary masterpiece within its historical and religious milieus*. Ph.D. thesis, University of Washington.
- Irving, T. 1985. *The Qur'an: first American version. translation and commentary*. Brattleboro, VT: Amana Books.
- Ismail, R., Rahman, N. and Bakar, Z. 2016. Identifying concept from English translated Qur'an. In: *Open Systems (ICOS), 2016 IEEE Conference on: IEEE*, pp.48-52.
- Jain, S. 1991. Plants of the Qur'an. JSTOR, p.291.
- Jaliliyan, S. and Haji-Khani, A. 2017. An analysis of spiritual factors of rain in the Qur'an. *Āfāq-i Dīn*. **7**(25), pp.33-54.
- Janssen, S. 1990. Automatic sense disambiguation with LDOCE: Enriching syntactically analyzed corpora with semantic data. In: Aarts, J., and Meijs, W. Eds. *Theory and practice in corpus linguistics* (No. 4). Rodopi, pp.105-35.
- Jassem, Z. A., & Jassem, A. J. 2001. Abdullah Yusuf Ali's translation of the Quran: an evaluation. *Issues in Education*, **24**, 29-52.
- Jivani, A. 2011. A comparative study of stemming algorithms. *Int. J. Comp. Tech. Appl.* **2**(6), pp.1930-1938.
- Jones, C. and Waller, D. 2015. *Corpus linguistics for grammar: a guide for research*. Routledge.
- Jurafsky, D., and Martin, J. 2009. *Speech and language processing*. Pearson Prentice Hall. ISBN 978-0-13-187321-6
- Kahrizi, D., Molsaghi, M., Faramarzi, A., Yari, K., Kazemi, E., Farhadzadeh, A., Hemati, S., Hozhabri, F., Asgari, H. and Chaghmirza, K. 2012. Medicinal plants in the holy Qur'an. *American Journal of Scientific Research*. **42**, pp.62-71.
- Kamali, M. 2008. *Shari'ah law: an introduction*. One World Publications.
- Kamdar, I. 2018. *Themes of the Quran*. Independently Published. ISBN13 9781730905858. e-book available from: <https://www.bookdepository.com/Themes-Quran-Ismail-Kamdar/9781730905858?ref=grid-view&qid=1582282546757&sr=1-1>

- Karanikolas, N. 2015. Supervised learning for building stemmers. *Journal of Information Science*. **41**(3), pp.315-28.
- Keane, A. 2017. *Python: python programming language for beginners (learn how to code in python, computer programming, python crash course, python cookbook)*. E-book Available from: <https://www.scribd.com/book/354087267/Python-Programming>"
- Kennedy, G. 2014. *An introduction to corpus linguistics*. London and New York: Routledge.
- Kenning, M. 2010. What are parallel and comparable corpora, and how can we use them. In O'Keeffe, A., & McCarthy, M. edss. *The Routledge handbook of corpus linguistics*. Routledge. pp. 487-500.
- Kenny, D. 2001. *Lexis and creativity in translation: a corpus-based study*. St. Jerome Publishing.
- Kenny, D. 2006. Translators at play: exploitations of collocational norms in German English. In: Dodd, B. ed. *Working with German Corpora: with a foreword by John Sinclair*. A&C. p.143.
- Khan, H., Saqlain, S., Shoaib, M. and Sher, M. 2013. Ontology based semantic search in holy Qur'an. *International Journal of Future Computer and Communication*. **2**(6), p.570.
- Kidwai, A. 1987. Translating the untranslatable: a survey of English translations of the Qur'an. *The Muslim World Book Review*. **7**(4), pp.66-71.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. and Suchomel, V. 2014. The sketch engine: ten years on. *Lexicography*. **1**(1), pp.7-36.
- Kilgarriff, A., Rychly, P., Smrz, P., and Tugwell, D. 2004. The sketch engine, In: Williams, G. and Vessier, S. eds. *Proceedings of the Eleventh International Congress of Euralex 2004. Bretagne, France: Université de Bretagne Sud*. pp.105-16.
- Krippendorff, K. 2018. *Content analysis: an introduction to its methodology*. Sage publications.
- Kübler, N. and Volanschi, A. 2012. *Semantic prosody and specialised translation, or how a lexico-grammatical theory of language can help with specialised translation*. Amsterdam/Philadelphia: John Benjamins.
- Kübler, S., McDonald, R., & Nivre, J. 2009. Dependency parsing. *Synthesis lectures on human language technologies*, **1**(1), 1-127.
- Kucera, H., and Francis, W. N. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Kyngas, H. and Vanhanen, L. 1999. Content analysis. *Hoitotiede*. **11**, pp. 3-12.

- Kytö, M., Ihalainen, O., and Rissanen, M. 1988. Corpus linguistics, hard and soft. In: *Proceedings of the Eighth International Conference on English Language Research on Computerized Corpora (Vol. 2)*. Rodopi.
- Lane, E. 1863. *Arabic-English Lexicon, derived from the best and most copious sources, 2 books in 4 vols.* London and Edinburgh: Williams and Norgate. 85.
- Laviosa, S 1997. How comparable can 'comparable corpora' be?. *Target* **9**(2), pp 289-319.
- Laviosa, S. 2002. *Corpus-based translation studies: theory, findings, applications (Vol. 17)*. Amsterdam: Rodopi.
- Lawson, T. 2017. *The Quran, epic and apocalypse*. Oneworld Publications.
- Leech, G. 1992. Corpora and theories of linguistic performance. In: Svartvik, J. ed.. 2011. *Directions in Corpus Linguistics*, pp. 105-22.
- Lehrer, A. 1974. *Semantic fields and lexical structure*. Amsterdam: North Holland Publishing Company.
- Lenci, A. 2008. Distributional semantics in linguistic and cognitive research. *Italian journal of linguistics*, **20**(1), pp.1-31.
- Léon, J. 2005. Claimed and unclaimed sources of corpus linguistics. *Henry Sweet Society for the History of Linguistic Ideas Bulletin*, **44**(1), pp. 36-50.
- Léon, J. 2007. Meaning by collocation. In *History of Linguistics 2005: Selected Papers from the Tenth International Conference on History of the Language Sciences (ICHOLS X), 1-5 September 2005, Urbana-Champaign, Illinois (Vol. 112, p. 404)*. John Benjamins Publishing.
- Lesk, M. 1986. Automatic sense disambiguation using machine-readable dictionaries: how to tell a pine cone from an ice cream cone. In: *Proceedings of the 5th annual international conference on Systems documentation: Citeseer*, pp.624-26.
- Lewandowska-Tomaszczyk, B. 1996. Cross-linguistic and language-specific aspects of semantic prosody. *Language Sciences*. **18** (1-2), pp.153-78.
- Lewis, J. and Ritchie, J. 2003. Generalising from qualitative research. In: Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. eds. *Qualitative research practice: A guide for social science students and researchers*. Sage.
- Lewis, M. 1993. *The lexical approach (Vol. 1, p. 993)*. Hove: Language teaching publications.
- Louw, B. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In: Baker, M. Francis, G. and Tognini-Bonelli, El. eds. *Text and technology: in honour of John Sinclair*. Amsterdam: John Benjamins, pp.157-76.

- Louw, B. 1997. The role of corpora in critical literary appreciation. In: Wichmann, A., Fligelstone, S., McEnery, A. and Knowles, G. eds. *Teaching and language corpora*. London: Addison Wesley Longman, pp.240-51.
- Louw, B. 2000. Contextual prosodic theory: bringing semantic prosodies to life. In: Heffer, C. ed. *Words in context: A tribute to John Sinclair on his retirement*. Birmingham: University of Birmingham.
- Louw, B. 2007a. Collocation as the determinant of verbal art. In: Miller, D. and Turci, M. eds. *Language and verbal art revisited: Linguistic approaches to the study of literature*, pp.149-80
- Louw, B. 2007b. Literary worlds as collocation. In: Watson, G. and Zyngier, S. eds. *Literature and stylistics for language learners*. London: Palgrave Macmillan, pp. 91-105.
- Louw, B. 2007c. Truth, literary worlds and devices as collocation. In: *Corpora in the Foreign Language Classroom*. Brill Rodopi, pp. 327-62.
- Louw, B. 2010. Collocation as instrumentation for meaning: a scientific fact. In: *Literary education and digital learning: methods and technologies for humanities studies*. IGI Global, pp. 79-101.
- Louw, B. and Chateau, C. 2010. Semantic prosody for the 21st century: are prosodies smoothed in academic contexts? A contextual prosodic theoretical perspective. In: *Proceedings of 10th International Conference on Statistical Analysis of Corpus Study*: Citeseer, pp.755-64.
- Louw, B., and Milojkovic, M. 2016. *Corpus stylistics as contextual prosodic theory and subtext (Vol. 23)*. John Benjamins Publishing Company.
- Lüdeling, A. and Kytö, M. 2009. *Corpus linguistics: an international handbook*. De Gruyter Mouton.
- Lyons, J. 1977. *Semantics. Volume I*. Cambridge: Cambridge UP.
- MacCarthy, M. 2001. *Issues in applied linguistics*. Ernst Klett Sprachen.
- Mahlberg, M. 2005. English general nouns: *A corpus theoretical approach*. Amsterdam/ Philadelphia: Benjamins
- Mahlberg, M. 2006. Lexical cohesion: corpus linguistic theory and its application in English language teaching. *International Journal of Corpus Linguistics*. **11**(3), pp.363-83.
- Mahlberg, M. 2009. Local textual functions of move-in newspaper story patterns. In: Römer, U. and Schulze, R. eds. *Exploring the lexis-grammar interface*. Amsterdam: John Benjamins, pp. 265–87.
- Mahlberg, M. 2013a. Corpus analysis of literary texts. In: Chapelle, C. A. ed. *The Encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell.

- Mahlberg, M. 2013b. *Corpus stylistics and Dickens's fiction*. New York /London: Routledge.
- Mahlberg, M. 2014a. Corpus stylistics. In: Burke M. ed. *The Routledge handbook of stylistics*. London: Routledge, pp. 378–92.
- Mahlberg, M. 2014b. Corpus linguistics and discourse analysis. *Pragmatics of Discourse: Handbooks of Pragmatics*, **3**, pp. 215-38.
- Mahlberg, M., 2010. Corpus linguistics and the study of nineteenth-century fiction. *Journal of Victorian Culture*, **15**(2), pp. 292-98, DOI: 10.1080/13555502.2010.491667
- Malik, M. 1997. *Al-Qur'an, the guidance for mankind-English translation of the meanings of Al-Qur'an with Arabic*. The Institute of Islamic Knowledge.
- Manning, C. and Schütze, H. 1999. *Foundations of statistical natural language processing*. MIT press.
- Manning, C. D., Raghavan, P., & Schütze, H. 2008. *Introduction to information retrieval*. Cambridge university press.
- Mason, I. 2001. Translator behaviour and language usage: some constraints on contrastive studies. *HERMES-Journal of Language and Communication in Business*, **26**, pp.65-80.
- Mason, O. 2000. Parameters of collocation: the word in the centre of gravity. *Language and Computers*, **30**, pp. 267-80.
- Matthes, E. 2015. *Python crash course: a hands-on, project-based introduction to programming*. No Starch Press. Available from: https://767ab07c-a-62cb3a1a-sites.googlegroups.com/site/books184pdf876/PDF-Python-Crash-Course-A-Hands-On-Project-Based-Introduction-to-Programming.pdf?attachauth=ANoY7crMYnwGEZt6WhLG6PVFy-OjOfY4FIISoR9xd8d8ffl6Hnyysk8RjvKZVoP6Sp6mg2mJ0u9AuM3CmLR3gounvlZvwLMg-I-kPVpEPvaCYbkjeg1hQe5cEt6P4ImWUj82MgqhA0UCis85U0OUKBwGDPdL8b4C2phx4D1sTIRiMxrUiN0GEOWnBmN_mQaQg0qZ4_I5Ofgyb6CF3UkBf08rl3C3EKo7_owPmS_XY6L-Q7C6907WysZCEsn1EpYcmMxtKvq1IKqvz5-VFsxf3QRAiMOwDze-OOi7gm4TgXse5F8Tr3M2hg%3D&attredirects=0
- McAlester, G. 1999. The source text in translation assessment. In: Anderman, G. and Rogers, M. eds. *Words, words, words. The translator and the language learner*. Clevedon, Philadelphia and Adelaide: Multilingual Matters, pp. 169–78.
- McArthur, T. 1998. *Concise Oxford companion to the English language*. Oxford University Press. [Online]. [Accessed December 2017]. Available from: <https://www.oxfordreference.com/view/10.1093/acref/9780192800619.001.0001/acref-9780192800619-e-319>.
- McEnery, A. and Xiao, R. 2004. The use of corpora in the study of collocation and semantic prosody exemplified through corpus data from English and Mandarin Chinese. In: *BAAL Annual Conference 2004*.

- McEnery, A. Xiao, R. and Tono, Y. 2006. *Corpus-based language studies: an advanced resource book*. Taylor and Francis.
- McEnery, A., and Baker, P. eds. 2015. *Corpora and discourse studies: integrating discourse and corpora*. Springer.
- McEnery, A., and Oakes, M. P. 1996. Sentence and word alignment in the CRATER project: Methods and assessment. In: Thomas, J. and Short, M. eds. *Using corpora for language research*. pp. 211-31
- McEnery, T. and Gabrielatos, C. 2006. English corpus linguistics. In: Aarts, B. and McMahon, A. eds. *The handbook of English linguistics*. pp.33-71.
- McEnery, T. and Hardie, A. 2012. *Corpus linguistics: method, theory and practice*. Cambridge: Cambridge University Press.
- McEnery, T. and Wilson, A. 2001. *Corpus linguistics: an introduction (2nd ed.)*. Edinburgh U.P.
- McEnery, T., and Wilson, A. 2005. Corpus linguistics. In: Mitkov, R. ed. *The Oxford handbook of computational linguistics*. Oxford University Press, pp. 448-63.
- McGuinness, D. L., and Van Harmelen, F. 2004. OWL web ontology language overview. *W3C Recommendation*, **10**(10).
- Meijs, W. 1987. Corpus linguistics and beyond. In: *Proceedings of the seventh international conference on English language research on computerized corpora (Vol. 59)*. Rodopi.
- Mel'čuk, I. 1998. Collocations and lexical functions. In: Cowie, A. ed. *Phraseology. Theory, analysis, and applications*, pp. 23-53.
- Mel'čuk, I. 2012. Phraseology in the language, in the dictionary, and in the computer. *Yearbook of Phraseology*, **3**(1), pp. 31-56.
- Menai, M. 2014. Word sense disambiguation using evolutionary algorithms—Application to the Arabic language. *Computers in Human Behavior*. **41**, pp.92-103.
- Meyer, C. 2002. *English corpus linguistics: an introduction*. Cambridge University Press.
- Mitkov, R. 2005. *The Oxford handbook of computational linguistics*. Oxford University Press.
- Mogotsi, I. C., Manning, C. D., Raghavan, P., & Schütze, H. 2010. Introduction to information retrieval. *Information Retrieval*, **13**(2), 192-195.
- Mohamed, M. 2014. The metaphor of nature in the Holy Qur'an: A critical metaphor analysis (CMA). *Journal of Arabic and Human Sciences*. **270**(1688), pp.1-37.
- Mohammed, K. 2005. Assessing English translations of the Qur'an. *Middle East Quarterly*. **12**(2), pp.58-71.

- Mohammed, S. 2012 *A corpus-based study of the collocational behaviour and ideological usage of political terms in the Arabic news discourse of pre-revolutionary Egypt*. Ph.D. thesis, Lancaster University.
- Morley, J. and Partington, A. 2009. A few frequently asked questions about semantic or evaluative prosody. *International Journal of Corpus Linguistics*. **14**(2), pp.139-58.
- Muhammad, A. 2012. *Annotation of conceptual co-reference and text mining the Qur'an*. PhD thesis, University of Leeds.
- Muhammad, A., Ul Qayyum, Z., Tanveer, S., Martinez-Enriquez, A. and Syed, A. 2012. E-hafiz: Intelligent system to help Muslims in recitation and memorization of Qur'an. *Life Science Journal*. **9**(1), pp.534-41.
- Mukhtar, T., Afzal, H. and Majeed, A. 2012. Vocabulary of Qur'anic concepts: a semi-automatically created terminology of holy Qur'an. In: *Multitopic Conference (INMIC), 2012 15th International: IEEE*, pp.43-46.
- Munday, J. 2012. *Evaluation in translation*. Routledge.
- Munday, J. 2013. *Introducing translation studies: theories and applications*. London and New York: Routledge.
- Munday, J. 2014. Systems in translation: A systemic model for descriptive translation studies. *Crosscultural Transgressions: Research Models in Translation Studies II. Historical and Ideological Issues*, pp. 76-92.
- Nawi, N. 2018. A group-based feature selection approach to improve classification of holy Qur'an verses. In: *Recent Advances on Soft Computing and Data Mining: Proceedings of the Third International Conference on Soft Computing and Data Mining (SCDM 2018), Johor, Malaysia, February 06-07, 2018*: Springer, p.282.
- Nelson, M. 2006. Semantic associations in business English: a corpus-based analysis. *English for Specific Purposes*, **25**(2), pp.217-34.
- Nesselhauf, N. 2004a. *Collocations in a learner corpus*. Amsterdam; Philadelphia: J. Benjamins Pub. Co.
- Nesselhauf, N. 2004b. What are collocations? In Allerton, D. Nesselhauf, N., and Skandera, P.eds., *Phraseological units: basic concepts and their application* (Vol. 8, pp. 1 - 21). Basel: Schwabe.
- Newmark, P. 1988. *A textbook of translation*. Prentice Hall New York.
- Nida, E. 1964. *Toward a science of translating: with special reference to principles and procedures involved in Bible translating*. Brill Archive.
- Nöldeke, T., Schwally, F., Bergsträßer, G. and Pretzl, O. 2013. *The history of the Qur'ān*. Brill.
- O'Keefe, A. and McCarthy, M. 2010. *The Routledge handbook of corpus linguistics*. Routledge.

- Olohan, M. 2004. *Introducing corpora in translation studies*. Routledge.
- Oostdijk, N. and Haan, P. 1994 *Corpus-based research into language*. Amsterdam: Rodopi.
- Ouda, K. 2015. *Qur'an analysis: a semantic search and intelligence system for the Qur'an*. [Online]. [Accessed December 2016]. DOI: 10.13140/RG.2.1.3165.7681.
- Pace-Sigge, M. 2013. The concept of lexical priming in the context of language use. *ICAME Journal*. **37**, pp.149-73.
- Palmer, E. H. 2017. *The Koran: the holy book of Islam with introduction and notes*. Watkins Media Limited.
- Palmer, F. R. 1976. *Semantics: a new outline*. Cambridge: Cambridge University Press.
- Parekh, A. 2000. *Complete easy dictionary of Qur'an*. [Online]. [Accessed 01 November 2016] Available from:
[https://understandQur'an.com/fileadmin/user_upload/vocabulary/dictionary/english/Easy_Dictionary_of_Qur'an_English_AK%20Parekh.pdf]
- Parkinson, D. 2011. Under the hood of arabiCorpus. byu. edu. In: *Arabic Corpus Linguistics Workshop, Brigham Young University, Provo, UT*.
- Partington, A. 1998. *Patterns and meanings: using corpora for English language research and teaching*. Amsterdam: Netherlands: John Benjamins Publishing.
- Partington, A. 2004a. Corpora and discourse, a most congruous beast. In: Morley, J. and Haarman, L. *Corpora and discourse*, pp. 11-20.
- Partington, A. 2004b. Utterly content in each other's company: semantic prosody and semantic preference. *International Journal of Corpus Linguistics*. **9**(1), pp.131-56.
- Partington, A., Duguid, A., and Taylor, C. 2013. *Patterns and meanings in discourse: theory and practice in corpus-assisted discourse studies (CADS) (Vol. 55)*. John Benjamins Publishing.
- Perkins, J. 2010. *Python text processing with NLTK 2.0 cookbook*. Packt Publishing Ltd.
- Philip, G. 2011. *Colouring meaning: collocation and connotation in figurative language*. John Benjamins Publishing.
- Phillips, M. 1985. *Aspects of text structure: An investigation of the lexical organisation of text*. Amsterdam, Netherlands: North-Holland.
- Phillips, M. 1989. *Lexical structure of text* [Discourse Analysis Monograph 12]. University of Birmingham, Birmingham, UK.
- Phillips, M. K. 1983. *Lexical macrostructure in science text* (Unpublished doctoral dissertation). University of Birmingham, Birmingham, UK.

- Phillis, M. 1985. *Aspects of text structure: an investigation of the lexical organisation of text*. North Holland.
- Pickthall, M. 1930. *The Koran*. WH Allen.
- Polanyi, L., and Zaenen, A. 2006. Contextual valence shifters. In: Shanahan, J. and Wiebe, J. eds. *Computing attitude and affect in text: theory and applications*. Dordrecht: Springer, pp. 1-10.
- Pooja, P., Dhiraj, A. and Sharvari, G. 2016. Rule-based stemmer using Marathi WordNet for Marathi Language. *International Journal of Advanced Research in Computer and Communication Engineering*. **5**, pp.278-82.
- Porter, M. 2001. *A language for stemming algorithms*. DOI: <https://snowball.tartarus.org/texts/introduction.html>.
- Poulsen, S. 2005. Collocations as a language resource. A functional and cognitive study in English phraseology. *Unpublished PhD dissertation, University of Southern Denmark*
- Rabadi, R. 2014. The use of corpora in linguistics. *Amarabac Magazin*. **39**(2232), pp.1-35.
- Rajaraman, A. and Ullman, J. 2011. *Mining of massive datasets: data mining*. Cambridge: Cambridge University Press, pp.1-17.
- Renouf, A. 2003. WebCorp: providing a renewable data source for corpus linguists. In: Granger, S. and Petch-Tyson, S. eds. *Extending the scope of corpus-based research; new applications*. Amsterdam: Rodopi, pp. 39-58.
- Renouf, A. and Sinclair, J. 1991. Collocational frameworks in English. In: Ajimer, K. and Altenberg, B. eds. *English corpus linguistics. studies in honour of Jan Svartvik*. Harlow: Longman, pp.128-43.
- Rippin, A. 2006. *The Blackwell companion to the Qur'aⁿ*. (Blackwell Companions to Religion). MA: Blackwell. ISBN 1 4051 1752 4
- Rippin, A. 2010. *Qur'an: Oxford bibliographies online research guide*. Oxford University Press. Retrieved from: https://books.google.co.uk/books/about/Qur_an_Oxford_Bibliographies_Online_Rese.html?id=e21Tdq6MiWEC&redir_esc=y
- Rippin, A. 2016. *The formation of the classical Islamic world: the Qur'an, style and contents*. Cambridge: Cambridge University Press.
- Ritchie, J., Lewis, J., Nicholls, C. M., and Ormston, R. eds. 2013. *Qualitative research practice: a guide for social science students and researchers*. Sage.
- Robinson, N. 1999. *Discovering the Qur'an: a contemporary approach to a veiled text*. Washington DC: Georgetown University Press.
- Römer, U. 2005. *Progressives, patterns, pedagogy: a corpus-driven approach to English progressive forms, functions, contexts and didactics*. John Benjamins Publishing.

- Römer, U. 2009. The inseparability of lexis and grammar: corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*. 7(1), pp.140-62.
- Rundell, M. 2000. These things happen. In: *Humanising Language Teaching, Year 2, Issue 4, July 2000*. Available from: <http://www.hltmag.co.uk/jul00/idea.htm>
- Saad, S., Salim, N., Zainal, H. and Noah, S. 2010. A framework for Islamic knowledge via ontology representation. In: *Information Retrieval and Knowledge Management, (CAMP), 2010 International Conference on: IEEE*, pp.310-14.
- Sadi, A., Anam, T., Abdirazak, M., Adnan, A., Khan, S., Rahman, M. and Samara, G. 2016. Applying ontological modelling on Qur'anic 'nature' domain. In: *2016 7th International Conference on Information and Communication Systems (ICICS): IEEE*, pp.151-55.
- Sadiq, A. 2010. *A comparative study of four English translations of Surat Ad-Dukhān on the semantic level*. Newcastle: UK: Cambridge Scholars Publishing.
- Saheeh International (1997). *The Qur'an: Arabic text with corresponding English meanings*. Riyadh: Abul Qasim Publishing House.
- Sarkar, D. 2018. Text analytics with python: A practical real-world approach to gaining actionable insights from your data. ISBN 978-1-4842-2388-8
- Schmitt, N. and Carter, R. 2004. Formulaic sequences in action: an introduction. In: Schmitt, N. ed. *Formulaic sequences: acquisition, processing, and use*. John Benjamins Press, pp.2-22
- Schütze, H. 1998. Automatic word sense discrimination. *Computational Linguistics*. 24(1), pp.97-123.
- Scott, M. 2007. John Sinclair (1933-2007). Language pioneer and explorer. *Language Awareness*, 16(2), 79-80.
- Seretan, V. and Wehrli, E. 2007. Collocation translation based on sentence alignment and parsing. In: *Actes de la 14e conference sur le Traitement Automatique des Langues Naturelles (TALN 2007)*. IRIT Press, 2007. pp. 401-10
- Shah, M. 2010. *A critical study of Abdel Haleem's new translation of the holy Qur'an*. Al Qalam. pp.1-15.
- Sharaf, A, Atwell, ES, Dukes, K et al. (8 more authors) 2010. Arabic and Quranic computational linguistics projects at the University of Leeds *اللغة على الحاسوبية*. In: *Proceedings of the workshop of Increasing Arabic Contents on the Web, organized by Arab League Educational, Cultural and Scientific Organization (ALECSO). Workshop of Increasing Arabic Contents on the Web, 16 Oct 2010, Damascus, Syria*. (Unpublished)
- Sharaf, A. and Atwell, E. 2012. QurSim: a corpus for evaluation of relatedness in short texts. *LREC Conference Proceedings*.

- Sharoff, S., Babych, B. and Hartley, A. 2006. Using collocations from comparable corpora to find translation equivalents. *LREC06 (2006)*. pp.465-70.
- Sherif, M. and Ngonga Ngomo, A. 2015. Semantic Qur'an. *Semantic Web*. 6(4), pp.339-45.
- Sinclair, J. 1966. Beginning the study of lexis. In: Bazell, C. E. Catford, J. C., Halliday, M. A. K and Robins R. H. eds. *In memory of J. R. Firth*. London, England: Longman. pp. 410–30
- Sinclair, J. 1987a. *Looking up: an account of the COBUILD project in lexical computing and the development of the Collins COBUILD English language dictionary*. London, England: Collins.
- Sinclair, J. 1987b. Collocation: A progress report. In: Steel, R. and Threadgold, T. eds. *Language topics: Essays in honour of Michael Halliday* PA: John Benjamins, pp. 319-31. (Chapter 8, 'Collocation', was first published in 1987 in R. Steele & T. Threadgold (Eds.), *Language Topics: Essays in Honour of Michael Halliday*, Amsterdam: John Benjamins.)
- Sinclair, J. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. 1992. The automatic analysis of corpora. In: Svartvik, J. ed. *Directions in corpus linguistics: Proceedings of the nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin and New York: Mouton de Gruyter, pp. 379-97.
- Sinclair, J. 1994. Trust the text. In: Coulthard, M. ed. *Advances in written text analysis*. London: Routledge, pp.12-25.
- Sinclair, J. 1996. The search for units of meaning. *Textus* 9(1), pp.75-106
- Sinclair, J. 1998. The lexical item. In: Weigand, E. ed. *Contrastive lexical semantics*, Amsterdam: Benjamins, pp.1-24
- Sinclair, J. 1999. *Concordance tasks*. [Online]. [Accessed 15 December 2017]. Available from: <http://www.twc.it/happen.html>. The Tuscan Word Centre.
- Sinclair, J. 2003. *Reading concordances: an introduction*. Pearson Longman.
- Sinclair, J. 2004a. The lexical item. In: Sinclair, J. McH. (ed.) *Trust the text. Language, corpus and discourse*. London/New York: Routledge, pp. 131–84.
- Sinclair, J. 2004b. The search for units of meaning. In: Sinclair, J. McH. (ed.) *Trust the text. Language, corpus and discourse*. London/New York: Routledge, pp. 24–48
- Sinclair, J. 2004c. Trust the text. In: Sinclair, J. McH. (ed.) *Trust the text. Language, corpus and discourse*. London/New York: Routledge, pp. 9-23
- Sinclair, J. 2005. Corpus and Text - Basic Principles. In: Wynne, M. (ed.) *Developing linguistic corpora: a guide to good practice*. Oxford: Oxbow Books, pp. 1-16.
- Sinclair, J. and Carter, R. 2004. *Trust the text. Language, corpus and discourse*. Routledge.

- Sinclair, J. M. 2006. Small words make big meanings. *Plenary, AAAL (American Association for Applied Linguistics)*
- Sinclair, J., Jones, S., and Daley, R. 2004. *English collocation studies: the OSTI report*. London: Continuum
- Skillitee, S. 1970. Arthur John Arberry. *Bulletin of the School of Oriental and African Studies*. **33**(2), pp.363-67.
- Smadja, F. 1993. Retrieving collocations from text: xtract. *Computational Linguistics*. **19**(1), pp.143-77.
- Stefanowitsch, A. 2017. *Corpus linguistics: a guide to the methodology*. Book Manuscript, Freie Universität Berlin: Book manuscript. [Online]. [Accessed 15 December 2017]. Available from: <http://stefanowitsch.net/clm/clmbookdraft>. Pdf.
- Stefanowitsch, A. and Gries, S. 2003. Collostructions: Investigating the Interaction of Words and Constructions. *International Journal of Corpus Linguistics* **8**(2), pp.209-24.
- Stefanowitsch, D. 2003. *Semantic prosody*. Available from: https://www.db-thueringen.de/servlets/MCRFileNodeServlet/dbt_derivate_00010791/corp_sempros.pdf
- Stewart, D. 2000. Conventionality, creativity and translated text: The implications of electronic corpora in translation. *Intercultural Faultlines: Research Models in Translation Studies I*. pp.73-91.
- Stewart, D. 2003. The corpus of revelation. The BNC, semantic prosodies and syntactic shells. In: *Lewandowska-Tomaszczyk, B. ed. PALC 2001: Practical Applications in Language Corpora. Proceedings of the conference Practical Applications in Language Corpora (PALC), Lodz, 7–9 September 2001*. Frankfurt: Peter Lang, pp.329–41.
- Stewart, D. 2009. Safeguarding the lexicogrammatical environment: translating semantic prosody. In: Beeby, A., Inés, P. R., & Sánchez-Gijón, P. eds. *Corpus use and translating: corpus use for learning to translate and learning corpus use to translate*. pp.29-46.
- Stewart, D. 2010. *Semantic prosody: a critical evaluation*. Routledge.
- Stubbs, M. 1993. British traditions in text analysis: From Firth to Sinclair. In: Baker, M., Francis, G and Tognini-Bonelli, E. eds. *Text and technology*. Philadelphia, PA: John Benjamins, pp. 1–33.
- Stubbs, M. 1994. Grammar, text, and ideology: computer-assisted methods in the linguistics of representation. *Applied Linguistics*, **15**(2), pp. 201-23.
- Stubbs, M. 1995. Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language*. **2**(1), pp.23-55.

- Stubbs, M. 1996. *Text and corpus analysis: computer-assisted studies of language and culture*. Oxford: Blackwell.
- Stubbs, M. 2001. *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell.
- Stubbs, M. 2005. Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature*, **14**(1), pp. 5-24.
- Stubbs, M. 2006. Corpus analysis: The state of the art and three types of unanswered questions. In: Thompson, G. and Hunston, S. eds. *System and corpus: exploring connections*, pp.15–36.
- Stubbs, M. 2007. On texts, corpora and models of language. In Hoey, M. Mahlberg, M. Stubbs, M. and Teubert, E. eds. *Text, discourse and corpora: theory and analysis*, London: Continuum, pp. 127-61.
- Stubbs, M. 2009. Memorial article: John Sinclair (1933–2007) the search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*. **30**(1), pp.115-37.
- Szabo, Z. G. 2006. The distinction between semantics and pragmatics. *The Oxford handbook of philosophy of language*, pp. 361-90.
- Ta'a, A., Abed, Q., and Ahmad, M. 2017. Al-Qur'an ontology based on knowledge themes. *Journal of Fundamental and Applied Sciences*, **9**(5), pp. 800-17.
- Ta'a, A., Zainal Abidin, S., Abdullah, M., Ali, M., Bashah, A. and Ahmad, M. 2013. Al-Qur'an themes classification using ontology. *Icoci. Cms. Net. My*, **74**, pp. 383-89.
- Taghva, K., Elkhoury, R. and Coombs, J. 2005. Arabic stemming without a root dictionary. In: *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on IEEE*, pp.152-57.
- Tang, X. and Liu, G. 2018. Solving contradictions in semantic prosody analysis with prosody concord. *International Journal of Corpus Linguistics*, **23**(4), pp. 437-66.
- Tao, H. 2003. Towards an emergent view of lexical semantics. *Language and Linguistics*, **4** (4), pp. 837–56
- Taylor, C. 2008. What is corpus linguistics? what the data says. *ICAME Journal*. **32**, pp.179-200.
- Taylor, C., and Marchi, A. 2018. *Corpus approaches to discourse: a critical review*. Routledge.
- Teubert, W. 1996. Comparable or parallel corpora?. *International Journal of Lexicography*, **9**(3), pp. 238-64.
- Teubert, W. 2005. My version of corpus linguistics. *International Journal of Corpus Linguistics*. **10**(1), pp.1-13.

- Teubert, W. 2007. Sinclair, pattern grammar and the question of hatred. *International Journal of Corpus Linguistics*, **12**(2), pp. 223-48.
- Teubert, W. and Čermáková, A. 2004a. Language and corpus linguistics. In: Halliday, M. Teubert, W. Yallop, C. and Čermáková, A. eds. *Lexicology and corpus linguistics*, London: Continuum. pp.73-112.
- Teubert, W. and Čermáková, A. 2004b. Directions in corpus linguistics. In Halliday, M. Teubert, W. Yallop, C. and Čermáková, A. eds. *Lexicology and corpus linguistics*, London: Continuum., pp.113-66.
- Teubert, W., Čermáková, A. and Wang, H. 2007. *Corpus linguistics: a short introduction*. London: Continuum.
- Thanopoulos, A., Fakotakis, N. and Kokkinakis, G. 2002. Comparative evaluation of collocation extraction metrics. In: *LREC: Citeseer*, pp.620-25.
- Thomas, D. 2006. A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, **27**(2), pp. 237-46. DOI: 10.1177/1098214005283748.
- Thompson, G. and Hunston, S., 2006. System and corpus: Two traditions with a common ground. In Thompson, G. and Hunston, S. eds. *System and corpus: exploring connections*, London: Equinox, pp. 1-14.
- Todd, R. 2019. Exploring the direction of collocations in eight languages. *Canadian Journal of Linguistics/Revue canadienne de Linguistique*. 64(1), pp.146-54.
- Tognini-Bonelli, E. 2001. *Corpus linguistics at work*. Amsterdam, Netherlands: John Benjamins Publishing.
- Toivonen, H., Gross, O., Toivanen, J. and Valitutti, A. 2013. On creative uses of word associations. In: Kruse, R., Berthold, M. R., Moewes, C., Gil, M. Á., Grzegorzewski, P., & Hryniewicz, O. eds. *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*. Springer, pp.17-24.
- Tribble, C. 2010. What are concordances and how are they used. In: O'Keeffe, A. and McCarthy, M. eds. *The Routledge handbook of corpus linguistics*, pp. 167-83.
- Tymoczko, M. 1998. Computerized corpora and the future of translation studies. *Meta: journal des traducteurs/Meta: Translators' Journal*. **43**(4), pp.652-60.
- Ullmann, S. 1962. *Semantics: an introduction to the science of meaning*. Oxford: Basil Blackwell.
- Veronis, J (ed) 2000. *Parallel text processing: alignment and use of translation corpora*. Dordrecht, Kluwer Academic Publishers.
- Wahidi, A. A. 1989. *Asbab al-Nuzul*. Beirut: Maktabah al-Tsaqafah. [Online]. [Accessed December 2017]. Available from: <https://www.altafsir.com/Books/Asbab%20Al-Nuzul%20by%20Al-Wahidi.pdf>.

- Weber, R. 1990. *Basic content analysis*. USA: Sage Publications.
- Weisser, M. 2016. *Practical corpus linguistics: an introduction to corpus-based language analysis (Vol. 43)*. John Wiley and Sons.
- Whitsitt, S. 2005. A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics*. **10**(3), pp.283-305.
- Wiese, A., Ho, V. and Hill, E. 2011. A comparison of stemmers on source code identifiers for software search. In: *2011 27th IEEE International Conference on Software Maintenance (ICSM), 25-30 Sept. 2011*, pp.496-99.
- Wild, S. 2015. Muslim translators and translations of the Qur'an into English. *Journal of Qur'anic Studies* **17**(3), pp.158-82.
- Willett, P. 2006. The Porter stemming algorithm: then and now. *Program*. **40**(3), pp.219-23.
- Williams, G. 1998. Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, **3**(1), 151–171. DOI: 10.1075/ijcl.3.1.07wil
- Wynne, M. 2005 (ed.) *Developing linguistic corpora a guide to good practice*. Oxbow Books Limited
- Xiao, R. and McEnery, T. 2006. Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics*. **27**(1), pp.103-29.
- Xiao, R., and Yue, M. 2009. Using corpora in translation studies: state of the art. In Baker, P. ed. *Contemporary Approaches to Corpus Linguistics*. London: Continuum, pp.237-62.
- Yahya, H. 2003. *General knowledge from the Qur'an*. New Delhi: Millat Book Centre.
- Yarowsky, D. 1993. One sense per collocation. In: *Proceedings of the Workshop on Human Language Technology: Association for Computational Linguistics*, pp.266-71.
- Yauri, A., Kadir, R., Azman, A. and Murad, M. 2012. Qur'anic-based concepts: verse relations extraction using Manchester OWL syntax. In: *Information Retrieval and knowledge Management (CAMP), 2012 International Conference on IEEE*, pp.317-21.
- Younis, N. 2018. Semantic prosody as a tool for translating prepositions in the holy Qur'an: A corpus-based analysis. In: McEnery, T. ed. *Arabic corpus linguistics*. Edinburgh University Press, pp.120-42
- Zaabalawi, R. 2019. English collocations versus Arabic collocations: A pedagogical dimension. *The Reading Matrix: An International Online Journal*. **19**(1).
- Zaidi, S., Laskri, M. and Abdelali, A. 2010. Arabic collocations extraction using Gate. In: *Machine and Web Intelligence (ICMWI), 2010 International Conference on: IEEE*, pp.473-75.

- Zanettin, F. 2002. Corpora in translation practice. *Language Resources for Translation Work and Research*, pp.10-4.
- Zethsen, K. 2006. Semantic prosody: creating awareness about a versatile tool. *Tidsskrift for Sprogforskning*, **4**(1-2), pp. 275-94.
- Zhang, C. 2010. An overview of corpus-based studies of semantic prosody. *Asian Social Science*. **6**, pp.190-94.
- Zhang, R. 2013. A corpus-based study of semantic prosody change: the case of the adverbial intensifier. *Concentric: Studies in Linguistics*. **39**(2), pp.61-82.
- Zhang, Y., Jin, R., and Zhou, Z. H. 2010. Understanding the bag-of-words model: a statistical framework. *International Journal of Machine Learning and Cybernetics*, **1**(1-4), 1, pp.43-52.

Appendix A: The Arabic Qur'an List of Stop-Words

Table 1: The Arabic Qur'an stop-words and their English transliteration and translation

<i>Arabic Stop word</i>	<i>Transliteration</i>	<i>English Translation</i> [https://www.almaany.com/en/dict/ar-en]
ذَلِكَ	<i>Dalika</i>	that
هَذَا	<i>haḍā</i>	this
هؤُلَاءِ	<i>ha`ulā`i</i>	these
تِلْكَ	<i>tilka</i>	that
أُولَئِكَ	<i>`ūla`ika</i>	those
أُولَآءِ	<i>`ulā`i</i>	these
هَاتَيْنِ	<i>hātaini</i>	these two
هَذِهِ	<i>haḍihi</i>	this is
أَلَمْ	<i>`lam</i>	did it not? has it not?
هَلْ	<i>hal</i>	do/does...?
كَيْفَ	<i>kaifa</i>	how do/does?
أَلَا	<i>`lā</i>	isn't/aren't? doesn't /don't?
مَاذَا	<i>māḍā</i>	what
أَنَّى	<i>`annae</i>	where/when
لِمَ	<i>lima</i>	why
كَمْ	<i>kam</i>	How many/much
مَتَى	<i>matae</i>	when
مَنْ	<i>man</i>	who/whom
أَيَّ	<i>`y</i>	which
أَيَّانَ	<i>`ayāna</i>	when/where
أَيْنَ	<i>`ayna</i>	where
بِمَ	<i>bima</i>	with what..
مِنْ	<i>mn</i>	from
فِي	<i>fī</i>	in
عَلَى	<i>`alae</i>	on
إِلَى	<i>`ilae</i>	to/ towards
الْبَاءِ	<i>bi</i>	the syllabic sound /bi/
الْلَامِ	<i>li</i>	the syllabic sound /la/
عَنْ	<i>`an</i>	about
مَعَ	<i>ma`</i>	with
حَتَّى	<i>ḥattae</i>	until
الْكَافِ	<i>ka</i>	the syllabic sound /ka/
الَّذِينَ	<i>aladīna</i>	the relative pronoun who/whom
مَا	<i>mā</i>	what
مِنْ	<i>mn</i>	from
الَّذِي	<i>alḍī</i>	who/ which (singular masculine)
الَّتِي	<i>allatī</i>	who/which (singular feminine)
الَّذَاتِي	<i>alllātī</i>	who/which (plural feminine)
الَّذِي	<i>alllā`ī</i>	who/which (plural masculine)
إِلَّا	<i>`lā</i>	except
دُونَ	<i>dūn</i>	without
غَيْرَ	<i>ḡair</i>	without
هُوَ	<i>huwa</i>	he

Arabic Stop word	Transliteration	English Translation [https://www.almaany.com/en/dict/ar-en]
هُمْ	hum	they
نَحْنُ	naḥnu	we
انت	'ant	you (singular)
أنا	'nā	I/me
هي	hiya	she
أَنْتُمْ	'ntm	you (plural)
هُنَّ	hunnā	they (plural feminine)
هما	hmā	they (dual)
إِيا	'yā	whoever/whatever
أَمْ	'am	or
تَم	tuma	then
أَوْ	'w	or
ولو	wlw	though
فإما	f'mā	either
لأن	l'n	because
كأين	k'yn	where
لكي	likay	in order to
ليتا	lyt	would
كي	ky	in order to
أن	'n	that
لكن	lakin	but
لعل	la'alla	might
لن	lan	will not
إن	'in	that
إن	'nna	that
لا	lā	No (negation)
أن	'n	if
إن	'n	if
لم	lm	did not
لو	lw	if
لئن	l'n	while
مهما	mhmā	whatever
لولا	lwlā	if not
لما	lmā	when (pronoun)
يا	yā	oh/ o/ alas
ذو	dū	possessor of (singular/plural masculine)
ذات	dāt	possessor of (singular feminine)
ذوات	dawāt	possessor of (plural/dual feminine)
ذوي	dwy	possessor of (plural/dual masculine)
كان	kāna	was
سوف	sawfa	will
ليس	laysa	an expression for negation (not)
لقد	laqad	has/have
قد	qad	may
لدي	ldy	in front of
هاهنا	hahnā	here
خلف	ḥlf	behind
الآن	'alān	right now

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation	
أَلَمْ	أَلَمْ	'alam	didn't...?	
هَلْ	أَوَلَمْ	'awalam	and didn't?	
كَيْفَ	هَلْ	hal	do?	
	فَهَلْ	fahal	is it?	
	هَلِ	hali	do you?	
	كَيْفَ	kaifa	how do?	
أَلَا	فَكَيْفَ	fakaiifa	so how?	
	وَكَيْفَ	wakaiifa	and how	
	أَلَا	'alā	except	
	مَاذَا	mādā	what	
أَنَّى	فَمَاذَا	famādā	what then?	
	وَمَاذَا	wamādā	and what	
	أَمَّاذَا	'ammādā	what?	
	أَنَّى	'annae	when/where?	
لِمَ	فَأَنَّى	fa'annae	ao when?/where?	
	لِمَ	lima	why	
كَمْ	فَلِمَ	falima	so why	
	وَكَمْ	wakam	and how much/many	
مَتَى	كَمْ	kam	how much/many	
	مَتَى	matae	when	
مَنْ	لِمَنْ	limani	whose	
	وَمَنْ	wamani	and from	
	وَلِمَنْ	waliman	and to whom	
	أَوْمَنْ	'awaman	or whom	
	بِمَنْ	bimani	who	
	لِمَنْ	lamani	whose	
	وَبِمَنْ	wabiman	and who	
	أَيُّ	أَيُّهُمْ	'aiyuhum	which one
		أَيُّكُمْ	'aiyukum	which one of you
		بِأَيِّ	bi'aiyi	with any
		أَيِّ	'aiyi	which
		فَأَيُّ	fa'aiyu	what?
أَيْنَ		'aina	where	
أَنَا		'annakum	I am ye	
أَيِّ		'aiya	which	
فَبِأَيِّ		fabi'aiyi	so what?	
أَيَّ		fa'aiya	what?	
أَيَّ		'aiyāna	whatever	
بِمَنْ		أَيَّنْ	'inna	isn't/aren't
	أَيَّنْ	'a'in	isn't/aren't	
مِنْ	بِمَنْ	bima	with what/whom	
	مِنْ	min	from	
	مِنْ	mina	from	
	مِنْهُمْ	minhum	who are they	

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	مِمَّا	mimmā	than
	مِنْكُمْ	minkum	of you
	مِنْهُ	minhu	from him
	مِنْهَا	minhā	of which
	وَمِنْ	wamin	and from
	وَمِنْ	wamina	and from
	مِنَّا	minnā	from us
	وَمِنْهُمْ	waminhum	and from them
	مِنْهُمَا	minhumā	who are they
	مِنْهُمْ	minhumu	who are they
	وَمِنْهَا	waminhā	such as
	بِمَنْ	biman	who
	مِنْهُمْ	minhunna	of them
	وَمِنْكُمْ	waminkum	and of you
	مِنْكَ	minka	from you
	وَمِمَّنْ	wamimman	and who
	مِنْكُمْ	minkunna	from you
	فِمَنْ	famina	who is it?
	فِمِنْ	famin	who
	مِنْكُمْ	minkumu	of you
	مِنَّا	mannan	from us
	وَمِنَّا	waminnā	and from us
	لَمِنْكُمْ	laminkum	for you
فِي	فِي	fī	in a
	فِيهَا	fīhā	In which
	فِيهِ	fīhi	in it
	وَفِي	wafī	and in me
	لَفِي	lafī	and in
	فِيمَا	fīmā	whatever
	فِيهِمْ	fīhim	in them (plural masculine)
	فِيهِنَّ	fīhinna	in them (plural feminine)
	فِيهِمَا	fīhimā	in them (dual)
	فَفِي	fafī	and in
	وَفِيهَا	wafīhā	and in it
	فِيمَ	fīma	in what
	أَفِي	'afī	is in ...?
عَلَى	عَلَى	'alae	on
	عَلَيْهِمْ	'alaihim	on them
	عَلَيْكُمْ	'alaikum	on you (plural)
	عَلَيْهِ	'alaihi	on him
	عَلَيْكَ	'alaika	on you (singular)

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	عَلَيْنَا	'alainā	on us
	وَعَلَى	wa'ala	and on
	عَلَيْكُمْ	'alaikumu	on you (plural)
	عَلَيَّ	'alaiya	on me
	عَلَيْهِنَّ	'alaihinna	On them (plural feminine)
	عَلَيْهِمَا	'alaihimā	on them (dual)
	فَعَلَيْهِ	fa'alaihi	and on him/it
	فَعَلَيْهَا	fa'alaihā	and on her/it
	فَعَلَيْهِمْ	fa'alaihim	so on them (plural masculine)
	عَلَيْهِمْ	'alaihimu	on them (plural masculine)
	عَلَيْهَا	'alaihā	on her/it
	عَلَيْهِ	'alaihu	on him/it
إِلَى	إِلَيَّ	'ilae	to me
	إِلَيْهِ	'ilaihi	to him/it
	إِلَيْكَ	'ilaika	to you (singular)
	وَإِلَى	wa'ilae	and to
	إِلَيْكُمْ	'ilaikum	to you (plural)
	إِلَيَّ	'ilaiya	to me
	وَإِلَيْهِ	wa'ilaihi	and to him/it
	إِلَيْكُمْ	'ilaikumu	to you (plural)
	إِلَيْهِمْ	'ilaihimu	to them (plural masculine)
	لِإِلَى	la'ilae	and to
	فَإِلَيْنَا	fa'ilainā	to us
	وَإِلَيَّ	wa'ilaiya	and to me
إِلَيْهِ	بِهِ	bihi	with it (masculine)
	بِهَا	bihā	with it (feminine)
	بِهِمْ	bihim	with them (plural masculine)
	بِكُمْ	bikum	with you (plural masculine)
	بِي	bī	with
	بِكَ	bika	with you (singular masculine)
	وَبِهِ	wabihi	and by him/it
إِلَيْهِمْ	لَهُمْ	lahum	for them (plural masculine)
	لَكُمْ	lakum	to you (plural masculine)
	لَهُ	lahu	for him/it (singular masculine)
	لَكَ	laka	is yours (singular masculine)
	لَنَا	lanā	for us
	لِي	lī	for me

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
		لَهُمْ <i>lahumu</i>	for them (plural masculine)
		لَكُمْ <i>lakumu</i>	for you (plural masculine)
		وَلَهُمْ <i>walahum</i>	and for them (plural masculine)
		لَهَا <i>lahā</i>	for her/hers
		وَلَهُ <i>walahu</i>	and for him/ and his
		وَلَكُمْ <i>walakum</i>	and for you (plural masculine)
		فَلَهُ <i>falahu</i>	for it /him
		لَكُمْآ <i>lakumā</i>	for you (dual)
		لِي <i>liya</i>	For me
		فَلَهَا <i>falahā</i>	and for her (singular feminine)
		فَلَكُمْ <i>alakum</i>	for you (plural masculine)
		لَكَ <i>laki</i>	is yours (singular feminine)
		وَلَهَا <i>walahā</i>	And for her
عَنْ		عَنْ <i>'an</i>	about
		عَنْهُمْ <i>'anhum</i>	about them (plural masculine)
		عَمَّا <i>'ammā</i>	what
		عَنْهُ <i>'anhu</i>	about him
		عَنْهَا <i>'anhā</i>	about her
		عَنْكُمْ <i>'ankum</i>	about you (plural masculine)
		عَنَّا <i>'annā</i>	about us
		عَنْهُمْ <i>'anhumu</i>	about them (plural masculine)
		عَنْكَ <i>'anka</i>	about you (singular masculine)
		عَنِّي <i>'annī</i>	about me
		وَعَنْ <i>wa'ani</i>	and about
		وَعَنْ <i>wa'an</i>	and about
مَعَ		مَعَ <i>ma'a</i>	with
		مَعَهُ <i>ma'ahu</i>	with him
		مَعَكُمْ <i>ma'akum</i>	with you (plural masculine)
		مَعَهُمْ <i>ma'ahum</i>	with them (plural masculine)
		مَعَكَ <i>ma'aka</i>	with you (singular masculine)
		مَعِي <i>ma'iya</i>	with me
		مَعَكُمْآ <i>ma'akumā</i>	with you
		مَعَهَا <i>ma'ahā</i>	with her
حَتَّى		حَتَّى <i>hattae</i>	until

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation	
الكاف الَّذِينَ	كَمَا	<i>kamā</i>	as such	
	الَّذِينَ	<i>!llaḏīna</i>	Whose (plural masculine)	
	وَالَّذِينَ	<i>wāllaḏīna</i>	and those (plural masculine)	
	لِلَّذِينَ	<i>lillaḏīna</i>	for those who are (plural masculine)	
	فَالَّذِينَ	<i>fāllaḏīna</i>	and those who are (plural masculine)	
	كَالَّذِينَ	<i>kāllaḏīna</i>	as those	
	بِالَّذِينَ	<i>bi!llaḏīna</i>	with those (plural masculine)	
	لِلَّذِينَ	<i>lallaḏīna</i>	for those who are	
	الَّذِينَ	<i>alllaḏīni</i>	who (plural masculine)	
	ما	وَمَا	<i>wamā</i>	and who/which
بِمَا		<i>bimā</i>	with what/whom	
وَمِنْ		<i>waman</i>	and from	
فَمَا		<i>famā</i>	what?	
لِذَا		<i>limā</i>	when	
وَمِمَّا		<i>wamimmā</i>	and what	
وَبِمَا		<i>wabimā</i>	and what	
فَبِمَا		<i>fabimā</i>	so what?	
من		مِمَّنْ	<i>mimmani</i>	of whom
		مَنْ	<i>man</i>	from
	فَمَنْ	<i>faman</i>	who?	
	لِمَنْ	<i>liman</i>	whose	
	لِمَنْ	<i>lamina</i>	whose	
	مِمَّنْ	<i>mimman</i>	of whom	
	مَنْ	<i>mani</i>	who	
	فَمَنْ	<i>famani</i>	who is	
	أَمَّنْ	<i>`amman</i>	as whom?	
	كَمَنْ	<i>kaman</i>	as whom	
الَّذِي	الَّذِي	<i>!llaḏī</i>	ahich/whom (singular/masculine)	
	وَالَّذِي	<i>wāllaḏī</i>	and which /whom (singular masculine)	
	بِالَّذِي	<i>bi!llaḏī</i>	with him	
	كَالَّذِي	<i>kāllaḏī</i>	as the one	
	لِلَّذِي	<i>lillaḏī</i>	for those who are	
	وَبِالَّذِي	<i>wabi!llaḏī</i>	and by him	
	الَّتِي	الَّتِي	<i>!llaṭī</i>	whom/which (singular feminine)
		بِالَّتِي	<i>bi!llaṭī</i>	with whom/which (singular feminine)
	الَّتِي الَّتِي	الَّتِي	<i>!llaṭī</i>	those (feminine)
		الَّتِي	<i>!llaṭī</i>	those

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	وَاللَّائِي	wāllā`ī	and those
إِلَّا	إِلَّا	`illā	except
	وَالْآ	wa`illā	and otherwise
دُونَ	دُونِهِ	dūnihi	without him/it
	دُونَ	dūna	without
	دُونِ	dūni	without
غَيْرِ	بِغَيْرِ	bigairi	otherwise
	غَيْرِ	gaira	non
	غَيْرِ	gairi	non
	غَيْرُهُ	gairuhu	not him
	لِغَيْرِ	ligairi	for not
	وَغَيْرِ	wāgaira	and not
	وَغَيْرِ	wāgairu	and not
	هُوَ	huwa	he is
هُوَ	وَهُوَ	wahuwa	and he
	فَهُوَ	fahuwa	it is
	لَهُوَ	lahuwa	and he/it is
	هُمْ	hum	they
هُمْ	وَهُمْ	wahum	and they are
	هُمْ	humu	they
	فَهُمْ	fahum	they are
	فَلَهُمْ	falahum	have them
	بِهِمْ	bihimu	with them
	أَهُمْ	`ahum	are they
	فَهُمْ	fahumu	and they are
	أَفَهُمْ	`afahumu	are they
	نَحْنُ	naḥnu	we
نَحْنُ	وَنَحْنُ	wanaḥnu	and we
	لَنَحْنُ	lanahnu	and we
	أَنْتَ	`anta	You (singular)
انت	وَأَنْتَ	wa`anta	and you (singular)
	أَأَنْتَ	`a`anta	are you (singular)
	لَأَنْتَ	la`anta	you are (singular)
	فَأَنْتَ	fa`anta	you are (singular)
	أَنَا	`anā	I
أنا	وَأَنَا	wa`anā	and I am
	هِيَ	hiya	she
هي	فَهِیَ	fahiya	they are
	لَهِیَ	lahiya	and it is
	وَأَنْتُمْ	wa`antum	and you (plural masculine)
أنتم	أَأَنْتُمْ	`a`antum	are you? (plural masculine)
	فَأَنْتُمْ	fa`antum	and you are (plural masculine)

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation	
		لَا أَنْتُمْ	<i>la`antum</i>	you are (plural masculine)
		لَهُنَّ	<i>lahunna</i>	for them (plural feminine)
هُنَّ		هُنَّ	<i>hunna</i>	them (plural feminine)
		بِهِنَّ	<i>bihinna</i>	with them (plural feminine)
		وَلَهُنَّ	<i>walahunna</i>	and to them (plural feminine)
		لَهُمَا	<i>lahumā</i>	to them (dual)
هما		بِهِمَا	<i>bihimā</i>	With them (dual)
إيا		إِيَّاهُ	<i>`iyāhu</i>	and him
		وَإِيَّاكُمْ	<i>wa`iyākum</i>	and you (plural masculine)
		وَإِيَّايَ	<i>wa`iyāya</i>	and to me
		إِيَّانَا	<i>`iyānā</i>	us
		فِيَّايَ	<i>fa`iyāya</i>	so that I may
		إِيَّاكَ	<i>`iyāka</i>	you (singular)
		وَإِيَّاكَ	<i>wa`iyāka</i>	and you (singular)
		وَإِيَّاهُمْ	<i>wa`iyāhum</i>	and them (plural masculine)
		أَمْ	<i>`am</i>	or
أَمْ		أُمِّ	<i>`ummi</i>	or
		أُمِّي	<i>`ami</i>	or
		ثُمَّ	<i>tumma</i>	then
ثم		أَوْ	<i>`au</i>	or
أو		وَلَوْ	<i>walau</i>	and if
ولو		فَمَا	<i>fa`immā</i>	either
فأما		لِأَنَّ	<i>li`an</i>	because
لأن		فَكَأَيِّنْ	<i>faka`aiyin</i>	so as
كأين		وَكأَيِّنْ	<i>waka`aiyin</i>	and like
		لِكَيْ	<i>likai</i>	in order to
لكي		لَيَتَّ	<i>laita</i>	would
ليت		لَيَتَّنَا	<i>laitanā</i>	we would
		كَيْ	<i>kai</i>	in order to
		وَأَنْ	<i>wa`an</i>	and now
مي		أَنْ	<i>`ani</i>	that
أن		أَنَا	<i>`annamā</i>	I am
		وَلَكِنْ	<i>walakin</i>	but
		وَلَكِنَّ	<i>walakinna</i>	and but
لكن		لَكِنِّي	<i>lakini</i>	but
		وَلَكِنِّي	<i>walakinnī</i>	but I do
		لَكُنَّا	<i>lakunnā</i>	but we
		وَلَكُنَّا	<i>walakinnā</i>	but we
		وَلَكِنَهُمْ	<i>walakinnahum</i>	but they (masculine)

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation	
		لَكِنَّا	lakinnā	but
		لَعَلَّكُمْ	la'allakum	perhaps you
لَعَلَّ		لَعَلَّهُمْ	la'allahum	perhaps they are
		لَعَلِّي	la'allī	perhaps
		لَعَلِّي	la'alae	perhaps
		لَعَلَّ	la'alla	might
		فَلَعَلَّكَ	fala'allaka	perhaps you (singular masculine)
		لَعَلَّكَ	la'allaka	perhaps you (singular masculine)
		لَنْ	lan	will not
لَنْ		وَلَنْ	walan	will not
		فَلَنْ	falan	not
		فَإِنْ	fa'in	so if
إِنْ		كَأَنَّ	ka'an	it was
		بِأَنَّ	bi'an	that
		أَنَّ	'anna	that
		وَأَنَّ	wa'anna	and that
		أَنَّهُمْ	'annahum	that they
		أَنَا	'annā	I
		أَنْتُمْ	'annakum	you are (plural)
		وَأَنَّهُ	wa'annahu	and it is
		وَأَنَا	wa'annā	and I am
		وَأَنَّهُمْ	wa'annahum	and that they are
		كَأَنَّهُ	ka'annahu	it was like
		فَكَأَنَّمَا	faka'annamā	and as though you were
		كَأَنَّمَا	ka'annamā	as if
		كَأَنَّهُهَا	ka'annahā	as if it were
		وَأَنِّي	wa'annī	and I
		فَأَنَّهُ	fa'annahu	for it is
		فَأَنَّ	fa'anna	it is
		بِأَنَّهُمْ	bi'annahumu	that they are
		وَأِنِّي	wa'innanī	and I am
		كَأَنَّكَ	ka'annaka	it's like you
		وَأَنْتُمْ	wa'annakum	and you
		كَأَنَّ	ka'anna	it was
		أَنَّهُ	'annahu	it's a
		إِنِّي	'innī	I am
إِنَّ		إِنَّ	'inna	that
		إِنَّا	'innā	I
		إِنَّهُ	'innahu	it's a
		إِنَّمَا	'innamā	only
		وَإِنَّ	wa'inna	though
		فَإِنَّ	fa'inna	it is

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	إِنَّهُمْ	'innahum	that they
	إِنَّكَ	'innaka	you are (singular)
	وَإِنَّا	wa'innā	and we are
	إِنَّكُمْ	'innakum	you are (plural)
	فَإِنَّمَا	fa'innamā	but it is
	بِأَنَّهُمْ	bi'annahum	that they are
	إِنَّهَا	'innahā	it
	فَإِنَّهُ	fa'innahu	for it is
	بِأَنَّ	bi'anna	that
	وَإِنِّي	wa'innī	and I am
	إِن	'ini	that
	وَإِنَّهُمْ	wa'innahum	and they are (masculine)
	فَإِنَّهُمْ	fa'innahum	for they are (masculine)
	فَإِنَّا	fa'innā	I am
	فَإِنَّكَ	fa'innaka	for you
	أَإِنَّا	'a'innā	I am
	فَإِن	fa'ini	it is
	فَإِنِّي	fa'innī	for I am
	وَإِنَّكَ	wa'innaka	and you (singular)
	إِنِّي	'innanī	I am
	إِنْنَا	'innanā	we are
	وَلَمَّا	wala'ini	while
	وَإِنَّمَا	wa'innamā	but only
	وَإِن	wa'ini	and now
	وَأَنَّ	wa'ani	and that
	لَمَّا	la'ini	whilst
	وَإِنَّكُمْ	wa'innakum	and you are (plural)
	وَإِنَّهَا	wa'innahā	and it is
	وَلَكِن	walakini	and but
	فَإِنَّكُمْ	fa'innakum	for you are (plural)
	أَأِنَّكَ	'a'innaka	tou are (singular)
	أَإِنَّا	'a'innā	are we
	بِأَنَّنَا	bi'annanā	that we are
	وَإِنَّهُمَا	wa'innahumā	and they are
	وَإِنَّهُ	wa'innahu	and it is
	لَا	lā	no
(negation) لَا	وَلَا	walā	nor
	فَلَا	falā	let it not
	أَنَّ	'an	that
أَنَّ	إِنَّ	'in	that
إِنَّ	وَلَوْ	walawi	and if
	وَإِنَّ	wa'in	and now
	لَمْ	lam	did not

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	وَلَمْ	walam	he did not
	فَلَمْ	falam	so when
لو	لَوْ	lau	if
	لَوْلَا	laulā	without it
لئن	لَئِنْ	la`in	while
	مَهْمَا	mahmā	whatever
مهما	وَلَوْلَا	walaulā	without it
لولا	لَمَّا	lamman	when
لما	وَيَا	wayā	hey
يا	أَيُّهَا	`aiyuhā	hey
	أَيُّهَا	`aiyatuhā	o you
	ذُو	ḏū	himself (singular nominative)
ذو	لِذِي	liḏī	himself (singular genitive)
	وَبِذِي	wabiḏī	and himself (singular genitive)
ذات	ذَا	ḏā	himself (accusative)
	ذَاتِ	ḏāti	herself
	ذَاتِ	ḏāta	herself
	وَذَاتِ	waḏāta	and herself
ذوات	ذَاتِ	ḏātu	herself
	ذَوَاتَا	ḏawātā	themselves (dual)
	ذَوِي	ḏawī	themselves (plural)
نوي	ذَوِي	ḏawai	themselves (plural)
	وَكَانُوا	wakānū!	and they were (plural masculine)
كَانَ	كُنْتُ	kunta	I was
	كَانَ	kāna	It was
	كَانُوا	kānū!	they were (plural masculine)
	كُنْتُمْ	kuntum	you were (plural masculine)
	كُنَّا	kunnā	we were
سَوْفَ	فَسَوْفَ	fasaufa	and will
	سَوْفَ	saufa	will
	وَأَسَوْفَ	walasaufa	and will
	لَسَوْفَ	lasaufa	and will
لَيْسَ	وَلَيْسَ	walaisa	It is not
	لَسْتُنَّ	lastunna	are not (plural feminine)
	وَلَقَدْ	walaqad	and... has/have
لَقَدْ	لَقَدْ	laqad	has/have
	وَلَقَدْ	walaqadi	and ...has/have
	لَقَدْ	laqadi	has/have
	قَدْ	qad	may

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	قَدْ	faqad	and may
	وَقَدْ	waqad	and may
	فَقَدْ	faqadi	and may
	قَدْ	qadi	may
	لَدَيْهِمْ	ladaihim	in front of them (plural masculine)
لدي	لَدَيَّ	ladaiya	in front of me
	هَاهُنَا	hāhunā	here
ها هنا	خَلْفَهُمْ	ḥalfihim	behind them
خلف	الآنَ	!lāna	right now
الآن	الآنَ	ālāna	right now
	فَالآنَ	fālāna	so now
	حَوْلَهُ	ḥaulahu	around it
حول	حَوْلَهَا	ḥaulahā	around it
	حَوْلَ	ḥaula	about
	وَرَاءَ	warā'a	behind
وراء	وَرَاءِهِمْ	warā'ihim	behind them
	هُنَاكَ	hunālika	there
هناك	فَوْقَ	fauqa	above
فوق	فَوْقَهُمْ	fauqahum	above them
	فَوْقَكُمْ	fauqakumu	above you (plural masculine)
	فَوْقَهُمْ	fauqahumu	above them (plural masculine)
	حِينَ	hīnin	when
حين	وَحِينَ	waḥīna	and then
	أَبَدًا	'abadan	never
أبدا	حَيْثُ	ḥaiṭu	where
حيثُ	وَحَيْثُ	waḥaiṭu	and where
	تَحْتِهَا	taḥtihā	underneath it
تحت	تَحْتَ	taḥta	under
	تَحْتِهِمْ	taḥtihimu	under them
	تَحْتَهُ	taḥtahu	underneath it
	بَيْنَ	baina	between
بين	بَيْنَهُمْ	bainahum	between them (plural masculine)
	بَيْنَهُمَا	bainahumā	between them (dual)
	بَيْنَنَا	bainanā	between us
	بَيْنِي	bainī	between me
	وَبَيْنَ	wabaina	and between
	بَيْنَ	baini	between
	وَبَيْنَكُمْ	wabainakum	and among you
	وَبَيْنَهُمْ	wabainahum	and between them
	بَيْنَكُمْ	bainakumu	among you
	بَيْنَكُمْ	bainikum	among you

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation	
		وَبَيْنِكَ	wabainika	and between you
		قَبْلُ	qablu	before
قبل		قَبْلِهِمْ	qablihim	before them
		قَبْلِكَ	qabluka	before you (singular masculine)
		قَبْلِي	qabli	before
		قَبْلِكُمْ	qablikum	before you (plural masculine)
		قَبْلِهِ	qablihi	before him
		قَبْلَهُمْ	qablahum	before them
		قَبْلِي	qablī	before me
		قَبْلِنَا	qablinā	before us
		قَبْلِهَا	qablihā	before her
		وَقَبْلَ	waqabla	and before
		قَبْلِهِمْ	qablihimu	before them
		بَعْدَ	ba'di	after
بعد		بَعْدَ	ba'da	after
		بَعْدَهُ	ba'dihi	after him
		بَعْدَهُمْ	ba'dihim	after them
		بَعْدَ	ba'du	after
		بَعْدَمَا	ba'damā	after
		فَعِنْدَ	fa'inda	and then
عند		عِنْدَ	'inda	at
		عِنْدِ	'indi	at
		عِنْدَهُ	'indahu	has
		عِنْدَنَا	'indanā	we have
		عِنْدِنَا	'indinā	we have
		عِنْدِي	'indī	I have
		عِنْدِكَ	'indika	you have
		عِنْدِهِ	'indihi	has
		عِنْدَكَ	'indaka	you have
		عِنْدَهَا	'indahā	then
		وَعِنْدَ	wa'inda	and then
		فَلَمَّا	falammā	and when he was
لَمَّا		لَمَّا	lammā	when
		وَلَمَّا	walammā	and when
		إِذْ	'id	then
إِذْ / إِذَا		وَإِذْ	wa'id	and then
		إِذِ	'idi	then
		وَإِذِ	wa'idī	and then
		وَإِذَا	wadā	and then
		إِذَا	'idā	if
		وَإِذَا	wa'idā	and If
		إِذَا	'idan	then
		أِذَا	'a'idā	then

Arabic Stop-word	Morpho-syntactic Forms	Transliteration	English Translation
	وَإِذَا	wa`īdan	if
	فَإِذَا	fa`īdan	then
	كَلَّا	kallā	both
كُلُّ	كُلُّ	kullun	each
	كُلَّمَا	kullamā	whenever
	وَكُلُّ	wakullun	and all
	كُلِّهَا	kullahā	all of which
	وَلِكُلِّ	walikullin	and for all
	كُلِّهِ	kullihī	all of it
	وَلِكُلِّ	walikulli	and for all
	وَكُلِّ	wakulla	and all
	كُلُّهُمْ	kulluhum	all of them (plural masculine)
	كُلِّ	kullin	each
	لِكُلِّ	likullin	for every
	كُلُّهُ	kulluhu	all of it
	أَوْكُلَّمَا	`awakullamā	or have they spoken? (dual)
	كِلْتَا	kiltā	both
	وَكُلُّهُمْ	wakulluhum	and all of them (plural masculine)
	كُلُّهُنَّ	kulluhunna	all of them (plural feminine)
	كُلِّهَا	kullihā	all of which
	بِكُلِّ	bikulli	with all
بَعْضُ	بَعْضُهُمْ	ba`duhum	some of them
	بَعْضِ	ba`da	some
	بَعْضِكُمْ	ba`dukum	some of you
	لِبَعْضِ	liba`din	for some
	بَعْضِكُمْ	ba`dakum	some of you
	بَعْضِ	ba`du	some
	بَعْضِ	ba`di	some
بَلْ	بَلْ	bal	Yes
	بَلِ	bali	yeah
حَمْ	حَمْ	ḥm	none
الْمِ	الْمِ	!m	none
الرِ	الرِ	!r	none
أَلْ	أَلْ	ālu	none
وَأَلْ	وَأَلْ	waāla	none
كَهَيْعَصْ	كَهَيْعَصْ	khy`ṣ	none
صْ	صْ	ṣ	none
قْ	قْ	q	none
نْ	نْ	n	none

Table 3: The Counts of the morpho-syntactic variations of the Arabic Qur'an stop-words

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
مِنْ	1673
فِي	1185
وَإِنَّهُ	812
الَّذِينَ	810
مِنْ	693
عَلَى	670
إِلَّا	662
لَا	658
وَمَا	646
إِلَيَّ	609
فَلَا	519
إِلَى	405
مَنْ	372
أَنْ	344
أَمْ	337
بِهِ	327
لَهُمْ	325
كُنْتُ	323
بِمَا	296
لَكُمْ	294
ذَلِكَ	280
لَهُ	275
الَّذِي	268
هُوَ	265
تُمْ	264
فِيهَا	241
وَمَنْ	234
كَانَ	229
وَدَا	195
هَذَا	190
كَانُوا	188
عَلَيْهِمْ	183
هُمْ	180
وَهُوَ	171
إِنْ	170
وَالَّذِينَ	164
وَإِنْ	163
إِنَّ	156
وَلَا	156
وَلَمَّا	156
عَنْ	153
وَيَا	150
إِنَّا	147

Morpho-Syntactic Forms *Counts*

عَلَيْكُمْ	146
مِنْهُمْ	145
عَلَيْهِ	145
حَتَّى	142
وَهُمْ	136
أَنَّهُ	131
إِذَا	130
فِيهِ	127
فَإِذَا	124
دُونَهَا	123
وَلِنَعْمَ	120
إِنَّهُ	113
فَلْيُبْسِئْ	113
مِمَّا	111
وَأَيَّاهُمْ	110
أَوْ	108
مِنْكُمْ	105
وَعِنْدَ	101
بِأَنَّ	99
فَعِنْدَ	98
فَلَنْ	89
مِنْهُ	88
مِنْهَا	86
فَمَا	86
كَذَلِكَ	83
لَنَا	83
دُونَ	83
فَقِيلَ لَهُمْ	82
إِنَّمَا	81
هُمْ	80
بِهَا	79
لِلَّذِينَ	79
أَلَمْ	78
فَلَمْ	77
إِلَيْهِ	76
إِلَيْكَ	74
وَإِنْ	71
تَحْتَهُ	70
وَبَيْنِكَ	70
لَكَ	69
إِذْ	69
عَنْهُمْ	67
فَمَنْ	67
هَلْ	66
تَحُنْ	65

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
وَمِنْ	64
كُنْتُمْ	63
كَيْفَ	62
فَإِنَّ	62
لِي	59
كَمَا	59
لَكِنَّا	59
لَعَلَّكَ	59
مَعَ	58
الَّتِي	57
أَمَّا	57
عَلَيْكَ	56
لِمَنْ	55
أَنْتَ	55
أَلَا	54
بَيْنَ	54
بَعْدَ	54
وَلَكِنْ	53
وَأَكْثَرَ	52
عَلَيْنَا	51
أَمِينِ	50
أَنَا	49
إِلَيْهِمْ	49
لَهُمْ	48
عَمَّا	47
وَأَنَا	47
لَمْ	47
كُنَّا	47
فَأَوْلَيْكَ	46
هَذِهِ	46
عَلَيْهَا	46
لَكُمْ	43
أَنَّ	43
أَكْبَرَ	43
لِهي	42
هُؤُلَاءِ	41
وَلَهُمْ	41
لَهَا	41
بِغَيْرِ	41
لَعَلَّكُمْ	41
وَأَنَّ	41
فَهُمْ	40
وَعَلَى	39
غَيْرِ	39
وَكَذَلِكَ	38

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
وَمِنْ	38
ذُوَيْهِ	38
لَوْ	38
عَنْهُ	37
مِثْلَهَا	37
بِهِمْ	36
قَبْلُ	36
أَوْلَمْ	35
فَبِأَيِّ	35
لِمَا	35
وَكَاثِرُوا	35
وَحَيْثُ	35
ذَلِكُمْ	34
مَعَهُ	34
مِنَّا	33
وَفِي	33
إِنَّكَ	33
هَا	33
فَلَمَّا	32
عَلَيْهِمْ	31
وَأِنَّا	31
لَمَّا	31
إِنَّكُمْ	30
إِخْدَاهُمَا	30
لَفِي	29
بِكُمْ	29
أَبَدًا	29
بَيْنَهُمْ	29
تِلْكَ	28
فَهُوَ	28
وَجِبِينَ	28
قَبْلِهِمْ	28
أَلْ	28
وَالِى	27
لَمِنْ	27
إِلَيْكُمْ	26
مَعَكُمْ	26
مَهُمَا	26
وَإِذَا	26
وَأَوْلَئِكَ	25
وَمِنْهُمْ	25
إِلَى	25
لَوْلَا	25
أَيُّهَا	25
قَبْلِكَ	25

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
فِيمَا	24
عَنْهَا	24
فَأَيْمًا	24
وَأَيْنَهُمَا	24
آل	24
فَهَلْ	23
أَيْنَ	23
وَلَهُ	23
كَأَنَّ	23
حَم	23
مَاذَا	22
وَالِيهِ	22
عَنْكُمْ	22
لَنْ	22
وَلَنْ	22
الْم	22
الر	22
آل	22
مِمَّنْ	21
مَنْ	21
عِنْدَ	21
وَأَنَّ	20
وَالِ	20
أَنَّ	19
بَعْدَ	19
عِنْدِ	19
عَلَيْكُمْ	18
وَبِذِي	18
قَبْلِ	18
وَلَكُمْ	17
لَهُوَ	17
أَتَيْتُمْ	17
بَعْدِهِ	17
فَكَيْفَ	16
فِيهِمْ	16
أَتَى	15
لِمَ	15
عَلَيَّ	15
فَمَنْ	15
وَالَّذِي	15
وَنَحْنُ	15
بِأَيْتِهِمْ	15
إِنِّهَا	15
فَرَفَعْتُمْ	15
قَبْلَكُمْ	15

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
قَبْلَهُ	15
بِهَذَا	14
وَأَنْتَ	14
كَيْ	14
أَنَا	14
فَأَيْتَهُ	14
بِأَنَّ	14
كَبِيعَصَ	14
ص	14
وَذَلِكَ	13
لَأَنْتُمْ	13
أَنْتُمْ	13
وَإِنِّي	13
هُنَالِكَ	13
بَيْنَهُمَا	13
ق	13
ن	13
كَمْ	12
مَعَهُمْ	12
أَمْ	12
وَلَمْ	12
وَإِنَّمَا	12
فَلَوْلَا	12
وَتِلْكَ	11
فَأَنَّى	11
عَلَيْهِمْ	11
بِي	11
عَنَّا	11
مَعَكَ	11
مَجِي	11
وَمِمَّا	11
أَمَّنْ	11
بِالَّذِي	11
إِنْ	11
لَمَّا	11
بَيْنِنَا	11
إِذَا	11
فَلَوْ	11
ذَلِكُمْ	10
مِنْهُمَا	10
بِكَ	10
فَلَهُ	10
فَالَّذِينَ	10
مِمَّنْ	10
غَيْرِ	10

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
وَلَهُنَّ	10
وَإِنَّهُمْ	10
فَأِنَّهُمْ	10
ذَا	10
وَكَمْ	9
مَتَى	9
فِيهِنَّ	9
كَالَّذِينَ	9
كَمَنْ	9
غَيْرُهُ	9
فَإِنْ	9
فَأَنَا	9
فَأَيْتَكَ	9
وَرَأَيْتَهُمْ	9
بَيْنِي	9
وَبَيْنَ	9
بَعْدَهُمْ	9
وَإِذْ	9
غَيْرُ	9
مِنْهُمْ	8
وَمِنْهَا	8
بِمَنْ	8
فِيهِمَا	8
عَلَيْهِمَا	8
عَنْهُمْ	8
عَنْكَ	8
الَّذِينَ	8
دُونَ	8
بِهِمَا	8
وَأَنَّهُ	8
أَنَا	8
بَيْنَ	8
بَلَى	8
فَأَمَّا	8
مِنْهُنَّ	7
لَكُمْ	7
بِالَّذِي	7
وَأَنْتُمْ	7
لَهُنَّ	7
وَلَوْ	7
أَيُّمَا	7
حَوْلَ	7
عِنْدَهُ	7
عِنْدَنَا	7
أَمَّا	7

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
فَسَوْفَ	7
سَوْفَ	7
لِمَنْ	6
وَمَنْ	6
أَيْهِمْ	6
أَيَّانَ	6
وَمِنْكُمْ	6
مِنْكَ	6
إِلَيْكُمْ	6
عَنِّي	6
بِالَّذِينَ	6
وَبِمَا	6
فَبِمَا	6
فَأَلَهُمْ	6
فَكَأَيُّنَ	6
وَلَكِنَّ	6
لَعَلَّهُمْ	6
فَإِنْ	6
فَأَيُّي	6
وَإِنَّكَ	6
ذَاتِ	6
فَوَرَى	6
وَبَيْنَكُمْ	6
عِنْدَنَا	6
وَأَمَّا	6
وَأَسَوْفَ	6
لَسَوْفَ	6
وَكَيْفَ	5
وَمِمَّنْ	5
كَالَّذِي	5
بِهِمْ	5
لَنَحْنُ	5
هِيَ	5
وَأَنَا	5
وَأَتَّهُمْ	5
إِنِّي	5
لَيْسَ	5
هَاهُنَا	5
خَلْفَهُمْ	5
فَالآنَ	5
حِينَ	5
عِنْدِي	5
كَأَلَّا	5
كُلُّ	5
كُلَّمَا	5

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
وَكُلٌّ	5
فَلَمْ	4
أَيُّكُمْ	4
أَنْتُمْ	4
مِنْكُمْ	4
فَعَلَيْهِ	4
لِي	4
لِغَيْرِ	4
إِيَّاهُ	4
لِيَتَّبِعَنَا	4
لَكِنْ	4
دَوَاتَا	4
لَدَيْهِمْ	4
لَدَيْ	4
تَحْتِهَا	4
عِنْدَكَ	4
عِنْدِهِ	4
أِذَا	4
كُلِّهَا	4
وَلِكُلِّ	4
كُلِّهِ	4
وَلِكُلِّ	4
وَكُلٌّ	4
كُلُّهُمْ	4
كُلِّ	4
بِذَلِكَ	3
فَمَاذَا	3
وَلِمَنْ	3
فَمِنْ	3
فَمِنْ	3
فَفِي	3
وَفِيهَا	3
فَعَلَيْهَا	3
إِلَيْهِمْ	3
فَلَهَا	3
وَعَنْ	3
الَّذِي	3
وَعَيْرِ	3
أَهُمْ	3
هُنَّ	3
لَهُمَا	3
وَأَيَّاكُمْ	3
أَمْ	3
لِكِي	3
وَلِكِنِّي	3

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
لَكُنَّا	3
لَعَلِّي	3
لَعَلَى	3
كَأَنَّ	3
كَأَنَّهُ	3
فَكَأَنَّمَا	3
كَأَنَّمَا	3
إِنَّنَا	3
وَلَيْنِ	3
وَإِنَّمَا	3
وَإِنْ	3
وَأَنْ	3
لَيْنِ	3
وَلَوْ	3
لَدُنْ	3
حَوْلَهُ	3
وَرَاءَ	3
تَحْتِ	3
وَبَيْنَهُمْ	3
بَيْنَكُمْ	3
قَبْلَهُمْ	3
بَعْدُ	3
عِنْدَكَ	3
عِنْدَهَا	3
إِذِ	3
وَإِذِ	3
لِكُلِّ	3
كُلُّهُ	3
أَوْكُلَّمَا	3
كَلَّمَا	3
لِهَذَا	2
فَتِلْكَ	2
أَوْلَاءِ	2
أَوْمَنْ	2
بِمَنْ	2
بِأَيِّ	2
أَيِّ	2
مِنْكُمْ	2
مُنَّا	2
وَمِنَّا	2
فِيهِ	2
أَفِي	2
لِأَيِّ	2
فَالِئِنَّا	2
وَبِهِ	2

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
فَلَكُمْ	2
لَكَ	2
وَلَهَا	2
وَعَنْ	2
اللَّائِي	2
وَاللَّائِي	2
إِلَّا	2
وَالْإِلَّا	2
أَأَنْتِ	2
لَأَنْتِ	2
فَأَنْتِ	2
بِهِنَّ	2
وَأِيَّايَ	2
إِيَّانَا	2
وَكَايُنَ	2
لَيْتَ	2
لَعَلَّ	2
فَلَعَلَّكَ	2
كَأَنَّهَا	2
وَأَنِّي	2
فَأَنَّهُ	2
فَأَنَّ	2
وَأِنَّكُمْ	2
وَأِنَّهَا	2
وَلَكِن	2
فَأِنَّكُمْ	2
أَأِنَّكَ	2
أَيُّهَا	2
تَارَةً	2
الآنَ	2
فَوْقَهُمْ	2
حَيْثُ	2
بَيْنَكُمْ	2
قَبْلِي	2
قَبْلِنَا	2
قَبْلِهَا	2
بَعْدَمَا	2
وَكُلُّهُمْ	2
كُلُّهُنَّ	2
كُلِّهَا	2
بِكُلِّ	2
بَعْضِ	2
بَعْضُهُمْ	2
بَعْضِ	2
بَعْضِكُمْ	2

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
وَذَلِكُمْ	1
فَهَذَا	1
هَاتَيْنِ	1
هَلِ	1
وَمَاذَا	1
أَمَّاذَا	1
لَمَنْ	1
وَبِمَنْ	1
فَأَيُّ	1
أَيِّ	1
فَأَيِّ	1
أَيْنَ	1
أَيْنَ	1
بِمَ	1
لِمَنْكُمْ	1
فَعَلَيْهِمْ	1
عَلَيْهِ	1
وَالِيَّ	1
مَعَكُمْ	1
مَعَهَا	1
لِلَّذِينَ	1
الَّذِينَ	1
وَبِالَّذِي	1
وَغَيْرُ	1
فَهُمْ	1
أَفَهُمْ	1
فَهِيَ	1
أَلَنْتُمْ	1
فَأَنْتُمْ	1
فَأَيَّايَ	1
إِيَّاكَ	1
وَأِيَّاكَ	1
فَأَيُّمَا	1
لَأَنْ	1
وَلَكِنَّا	1
وَلَكِنِّيهِمْ	1
بِأَنْتُمْ	1
وَأَنْتِي	1
كَأَنَّكَ	1
وَأَنْتُمْكُمْ	1
أَيْنَا	1
بِأَنْنَا	1
لِيُنَّ	1
وَلَوْلَا	1
دُو	1

<i>Morpho-Syntactic Forms</i>	<i>Counts</i>
لِذِي	1
ذَاتِ	1
وَذَاتِ	1
ذَاتُ	1
ذَوِي	1
وَلَيْسَ	1
لَسْتُ	1
فَقَدْ	1
وَقَدْ	1
فَقَدْ	1
قَدْ	1
وَلَقَدْ	1
لَقَدْ	1
وَلَقَدْ	1
لَقَدْ	1
الآنَ	1
حَوْلَهَا	1
فَرُوقَكُمْ	1
تَحْتِهِمْ	1
وَقِيلَ	1
وَإِذَا	1
لِيَبْغِضَ	1
بِعُضْنِكُمْ	1
بِعُضْرٍ	1
بِعُضْرٍ	1
بَلْ	1
بَلْ	1
قَدْ	1

Appendix B: The English List of Stop-Words in the Translations of the Qur'an

Table 4: The English list of Qur'an stop-words and their Arabic translation

<i>English stop word</i>	<i>Arabic translation [https://dictionary.cambridge.org/dictionary/english-arabic]</i>
a	أداة تُسْتَعْمَد لِلتَّكْبِير
about	حول
above	في الاعلى
according	علي حسب
after	بعد
again	مرة أخرى
against	ضد
ah	آه
all	الكل
also	أيضا
although	بالرغم من
am	فعل الكون المُسْتَعْمَد مَعَ الضَّمِير I
amid	وسط
among	من بين
amongst	وسط
an	أداة تُسْتَعْمَد لِلتَّكْبِير
and	و
another	آخر
any	أي
anybody	اي شخص
anyone	أي واحد
anything	اي شى
anywhere	في أى مكان
are	فعل الكون في صيغة الجَمْع للمُضَارِع
aren't	لا
as	مثل
at	في
be	فعل الكون
because	لان
been	التصريف الثالث من be
before	قبل
behind	خلف
being	يجرى
below	أدنى
beneath	تحت
besides	بالإضافة إلى
between	ما بين
betwixt	ما بين
both	على حد سواء
but	لكن
by	بواسطة

English stop word	Arabic translation [https://dictionary.cambridge.org/dictionary/english-arabic]
can	يستطيع
could	يُمكن
couldn't	لا يُمكن
couldst	يُمكن
did	الفعل الماضي من do
didn't	لم
didst	الفعل الماضي من do
do	فعل مُساعد يُستخدَم في أول الجُملة للسؤال بمعنى: هل/ فعل مُساعد يُستخدَم لمنع التكرار في الجُملة بمعنى: وأنا أيضا/ فعل مُساعد يُستخدَم في آخر الجُملة للسؤال بمعنى: أليس كذلك/ فعل مُساعد يُستخدَم للتأكيد
does	صيغة المضارع البسيط من do
doesn't	صيغة المضارع البسيط المنفي من do
don't	do not اختصار
dost	صيغة المضارع البسيط من do
doth	فعل مُساعد يُستخدَم في أول الجُملة للسؤال بمعنى: هل/ فعل مُساعد يُستخدَم لمنع التكرار في الجُملة بمعنى: وأنا أيضا/ فعل مُساعد يُستخدَم في آخر الجُملة للسؤال بمعنى: أليس كذلك/ فعل مُساعد يُستخدَم للتأكيد
down	أسفل
during	أثناء
each	كل
either	إما
ever	أبدا
every	كل
except	إلا
few	قليل
for	إلى عن على
forthwith	فورا
from	من عند
further	بالإضافة إلى ذلك
furthermore	علاوة على ذلك
had	الفعل الماضي والتصريف الثالث من have
hadn't	اختصار had not
hadst	الفعل الماضي والتصريف الثالث من have
has	فعل مُساعد يُستخدَم مع التصريف الثالث لفعل آخر للإشارة إلى أزمنة المضارع التام والماضي التام
hasn't	اختصار has not
hast	فعل مُساعد يُستخدَم مع التصريف الثالث لفعل آخر للإشارة إلى أزمنة المضارع التام والماضي التام
hath	الفعل الماضي والتصريف الثالث من have
have	فعل مُساعد يُستخدَم مع التصريف الثالث لفعل آخر للإشارة إلى أزمنة المضارع التام والماضي التام
haven't	اختصار have not
having	فعل مُساعد يُستخدَم مع التصريف الثالث لفعل آخر للإشارة إلى أزمنة المضارع التام والماضي التام
he	هو
henceforth	من الآن فصاعدا
her	لها
here	هنا
hers	لها
herself	نفسها
him	له
himself	نفسه
his	له
how	ماذا
however	ومع ذلك

<i>English stop word</i>	<i>Arabic translation [https://dictionary.cambridge.org/dictionary/english-arabic]</i>
i	أنا
if	إذا
in	في
into	إلى
is	هو
isn't	ليس
it	هذا
its	انها
its	انها
itself	بحد ذاتها
just	مجرد
manysoever	كثير إلى أي حد ممكن أو معروف
may	قد
mayest	الأكثر قابلية
me	أنا
mere	مجرد
midst	وسط
might	ربما
mightn't	قد لا
more	أكثر من
moreover	علاوة على ذلك
most	عظم
must	يجب
mustn't	ممنوع
my	لي
myself	نفسي
nay	لا بل
near	قريب
neither	لا هذا ولا ذاك
never	أبدا
no	لا
nor	ولا
not	ليس
now	الآن
o'er	على
of	من
off	إيقاف
oft	كثيرا
often	غالبا
on	على
only	فقط
or	أو
other	آخر
otherwise	غير ذلك
Our	لنا
ours	لنا
ourselves	أنفسنا
out	خارج
over	على
own	خاصة

<i>English stop word</i>	<i>Arabic translation [https://dictionary.cambridge.org/dictionary/english-arabic]</i>
perchance	بالمصادفة
perhaps	ربما
rather	بدلاً
same	نفسه
shall	سوف
shalt	سوف
shan't	لا يجوز
she	هي
she's	انها
should	من المُتَوَقَّع أن ...
should've	يجب أن قمت
shouldn't	لا ينبغي
shouldst	من المُتَوَقَّع أن ...
so	وبالتالي
some	بعض
specially	خصيصاً
such	هذه
than	من
that	أن
that'll	التي سوف
the	ال
thee	الميك
their	هم
theirs	لهم
them	معهم
themselves	أنفسهم
then	ثم
thenceforth	من ذلك الحين فصاعداً
there	هناك
thereat	في ذلك الزمان
thereby	وبالتالي
therefore	وبالتالي
therefrom	من ذلك
therein	في ذلك
thereof	من ذلك
thereon	في هذا الشأن
therewith	مع ذلك
these	هؤلاء
they	هم
thine	لك
this	هذه
thither	هناك
those	أولئك
thou	أنت
though	على أية حال
through	عبر
throughout	على مدار
thus	وهكذا
thy	خاصتك
thyself	نفسك

<i>English stop word</i>	<i>Arabic translation [https://dictionary.cambridge.org/dictionary/english-arabic]</i>
till	حتى
to	إلى
too	جدا
towards	تجاه
under	تحت
unless	ما لم
until	حتى
unto	حتى
up	فوق
upon	بناء على
us	لنا
very	للمغاية
was	كان
wasn't	لم يكن
wast	جملت
we	نحن
were	كانت
weren't	لم تكن
what	ماذا
whatever	ايا كان
when	متى
whenever	كلما كان
where	أين
whereby	بواسطة
wherein	حيث
whereof	م
whereon	ينفع
wheresoever	حيثما
wherever	أينما
wherewith	بماذا
whether	سواء
which	التي
while	في حين
whilst	في حين
who	من الذى
whoever	من
whom	من
whomsoever	أيا كان
whose	ملك من
whosoever	أيا كان
why	لمماذا
will	سوف
with	مع
within	في غضون
without	بدون
won't	متعود
would	فعل مُساعدِ بمعنى سيكون
wouldn't	لن
wouldst	فعل مُساعدِ بمعنى سيكون
ye	أيها

<i>English stop word</i>	<i>Arabic translation [https://dictionary.cambridge.org/dictionary/english-arabic]</i>
yea	نعم
yet	بعد
you	أنت
you'd	كنت
you'll	عليك
you've	كنت قد
your	ك
you're	أنت على
yours	خاصة بك
yourself	نفسك
yourselves	أنفسكم

Appendix C: The Lexico-grammatical Patterns of Nature in the Qur'an

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
people people	people praise	sent people	many people	day subservient		earth verily
signs people	people say	made people	grateful people	earth six		
truth people	people disbelieve	say earth	six day	heavens six		
sign people	day know	gather day	merciful heavens	heavens all		
people god	day established	created day	seven heavens	soul good		
people nation	day judge	knows day	wrong soul	soul astray		
bounty people	earth see	merge day	lost soul	soul one		
food people	earth assemble	made day	punished soul	wrongdoers		
guidance people	earth appointed	created earth	merciful man	life good		
Noah people	sky made	know earth	alike fruit	heart sealed		
Moses people	heavens raise	made earth	alike fruit	rivers eternal		
people lut	soul earned	belong earth	righteous garden	verily heart		
people pharaoh	soul know	glorifying of God	active angel			
day rejecters	soul paid	praise earth	greatest angel			
woe day	resurrection wronged	swallow earth	merciful night			
day faces	resurrection know	created heavens	one death			
day faces	resurrection promise	belong heavens				
earth day	resurrection waste	know earth				
dominion day	hell abode	send heavens				
days throne	hell pray	know heavens				
day things	hell guide	sent heavens				
night days	life die	glorifying of God				
day sun	life reply	praise heavens				

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
day moon	man remind	say heavens				
day signs	man say	kill soul				
days people	man-made	doeth soul				
heavens earth	garden flow	create soul				
mischief earth	garden abode	guide resurrection				
dominion earth	heart understand	reject resurrection				
earth mountains	hand earn	fill resurrection				
earth mountains	hand say	made hell				
water earth	devil said	enter hell				
water earth	devil work	taste hell				
earth Almighty	angel gather	invite hell				
god earth	night subjected	reject hell				
earth god	night prostrate	gave man				
earth death	night praise	made man				
earth people	death say	created man				
creation earth	death reject	gather man				
earth thing	face submit	work garden				
earth AllWise	water raise	believe garden				
insolence earth	water emerge	enter garden				
animal earth	water made	doeth garden				
sustenance earth	village wronged	made heart				
earth heavens	village made	believers				
heavens earth	village said					
heavens water	rivers abode	believe hand				
dominion heavens	light guide	spread hand				
creation heavens	light send					
heavens All-Mighty	light disbelieve	believe hand				

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
lord heavens	light belong	follow devil				
day heavens	ship drown	prostrate angel				
day heavens	wind blast	praise angel				
descend heavens	spring drink	command angel				
sustenance heavens	spring drink	say angel				
unknown heavens	fruit said	knows angel				
truth heavens		merge night				
heavens clouds		appointed night				
soul Allah		created night				
effort soul		turn face				
burden soul		sent water				
soul witness		send water				
soul spouse		create water				
soul thing		said worlds				
book soul		made mountains				
earth soul		flow rivers				
soul spacious		flow rivers				
guidance soul		enter rivers				
day resurrection		made river				
Noah resurrection		made river				
prophet resurrection		made light				
resurrection injustices		emerge light				
Moses resurrection		conceal chest				
warning resurrection		know chest				
lord resurrection		relieve chest				
Resurrection worshiper		subjected sun				
triumph resurrection		subjected moon				

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
resurrection fire						
resurrection ignorance						
brain resurrection						
knowledge resurrection						
disbelief hell						
companions hell						
hell doom						
destination hell						
guardian hell						
day hell						
mankind hell						
injustice hell						
disbelief hell						
people hell						
hell sign						
food hell						
life death						
people life						
glory life						
life earth						
water earth						
water earth						
life garnish						
sustenance life						
life fate						
life people						
punishment life						

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
life reward						
life doom						
life guidance						
man woman						
jinn man						
jinns man						
day man						
devil man						
disbelief man						
gardens river						
garden Eden						
companions gardens						
bliss gardens						
garden believers						
heaven people						
heart disease						
sight heart						
heart resurrection						
feet hands						
mercy hand						
prophet hand						
hand worshipper						
mankind hand						
hand injustice						
book hand						
book hand						
hand thing						

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
penalty hand						
mankind Allah						
step devil						
devil enemy						
party devil						
devil fire						
devil punishment						
angel spirit						
angel prophet						
earth angels						
night day						
night moon						
night sun						
dawn night						
signs night						
night sight						
life death						
earth death						
soul death						
day faces						
day faces						
face direction						
sky water						
water life						
lord worlds						
prophet worlds						
blessings worlds						

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
people village						
village messenger						
sky mountains						
day mountains						
day mountains						
rivers reward						
Eden rivers						
son women						
man women						
women women						
women soul						
darkness light						
light darkness						
Allah light						
sky light						
things cattle						
jinn mankind						
jinn mankind						
jinn mankind						
jinn bounties						
ship sea						
ship sea						
land sea						
pomegranate						
date-palm fruit						
date-palm fruit						
grape fruit						

N+N	N+V	V+N	ADJ+N	N+ADJ	AV+N	N+AV
grape fruit						
river fruit						
river fruit						
sun moon						
day sun						
night sun						
sun stars						
sun stars						
ship bounties						
night ship						
ship people						
earth tree						
sky moon						
garden date-palm						
date-palm grapevine						

Appendix D: The Lexico-grammatical Patterns of Nature and Their Discourse SPs in English Translations

The lexico-grammatical patterns of nature and their meanings in Pickthall's Translation

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
people of scripture	disbelievers	people believe	believers	said people	disbelievers	last day	the horror of judgment day	people clear	emphasis on message	indeed people	emphasis on message	earth verily	emphasis on message
moses	emphasis on	people serve	believers	sent people	emphasis on	awful day	horror of judgment day	people wrong	emphasis on message	verily heart	emphasis on message	life verily	emphasis on message
people revelations	emphasis on	people make	glorifying of God	created earth	glorifying of God	right soul	believers	people good	wrong doers	heart	message	night verily	emphasis on message
people portents	emphasis on	people follow	emphasis on message	belong earth	glorifying of God	good soul	believers	people earth wise	believers	glorifying of God			
people sense	emphasis on	people save	message	know earth	God	guilty hell	wrongdoers	earth six	glorifying of God	heavens six	glorifying of God		
brother	emphasis on	day raised	the horror of judgment day	praise earth	glorifying of God	able life	glorifying of God	evil doings	glorifying of God	of Satan			
people city chieftains	wrongdoers	day blown	judgment day	made earth	God	garden	reward	soul evil heart	glorifying of God	sealed man	hypocrites	Islamic teaching	
people	wrong doers	day gather	message	heavens	God	garden	reward	death	glorifying of God	one	disbelievers		
Noah people	message	day come	message	heavens	God	right hand believing	reward	women	glorifying of God	good	believers		
people knowledge	emphasis on	day promised	message	hurt soul	wrong doers	women	believers	one	glorifying of God	women	believers		
doom day	Afterlife punishment	earth say	message	soul	God	vile women	wrongdoers	good	glorifying of God	women	believers		
woe day	Afterlife punishment	heavens measure	glorifying of God	believe soul	believers	hypocritical women	hypocrites	humble	glorifying of God	cattle	believers		
day soul	emphasis on	heavens glorify	glorifying of God	judge resurrection	emphasis on	virtuous women	believers	forbidde n	glorifying of God	Islamic teaching			
day trumpet	the horror of judgment day	heavens give	glorifying of God	made resurrection	glorifying of God	aware chest	emphasis on	message	glorifying of God				
day decision	emphasis on	soul wronged	wrong doers	bring resurrection	emphasis on	hidden chest	glorifying of God		glorifying of God				

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
footsteps	evil	doings	of										
devil	Satan	glorifying	of										
lord angels	God	horror	of										
angels day	judgment day	glorifying	of										
angels spirit	God	emphasis	on										
angels witness	message	glorifying	of										
earth angels	God	horror	of										
heavens angels	judgment day	glorifying	of										
heavens angels	God	glorifying	of										
night day	God	glorifying	of										
night sun	God	glorifying	of										
night moon	God												

The lexico-grammatical patterns of nature and their meanings in Ali's Translation

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
people book	emphasis	on	people	said			emphasis	on	glorifying	of	verily	emphasis	on
	message		rejected	people	disbelievers	las day	message	people	God	day	message	firmly	God
people noah	emphasis	on	horrors	of	sent	emphasis	on	glorifying	of	people	emphasis	on	death
people pharaoh	message	people fear	judgment day	people	message	one soul	God	clear	message	heavens	message	verily	message
	disbelievers	people	worship	believers	message	lost soul	wrong doers	earth	glorifying	of	verily	emphasis	on
	glorifying	of		emphasis	on	believe	wrong	spacious	life	message	verily	emphasis	on
lord people	God	day say	message	day	believers	soul	wrong doers	garden	reward	mankind	message	emphasis	on
thamud	emphasis	on	emphasis	on	explain	message	good life	heart	disbelievers	indeed	message	emphasis	on
people	message	day see	message	day	message	wicked	believers	sealed	miracle	verily	emphasis	on	message
people lut	wrongdoers	day raised	judgment day	earth	glorifying	of	God	wrongdoers	hand white	story	worlds	message	
	glorifying	of	emphasis	on	made	glorifying	of	impure	mankind	glorifying	of		
night day	God	earth say	message	earth	God	men	wrongdoers	one	God				

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
earth signs	glorifying of	God	glorifying of	work	God	reward	glorifying of						
earth power	glorifying of	God	glorifying of	garden	God	knows	glorifying of						
earth glory	glorifying of	God	glorifying of	heart	God	reward	glorifying of						
earth rain	glorifying of	God	glorifying of	sent heart	God	learn	emphasis on						
earth end	emphasis on	God	emphasis on	heart	God	message	glorifying of						
water heavens	message of	God	message of	know	God	glorifying of	emphasis on						
rain heavens	glorifying of	God	glorifying of	mankind	God	sent	emphasis on						
things heavens	glorifying of	God	glorifying of	mankind	God	angels	emphasis on						
dominion heavens	glorifying of	God	glorifying of	sent	God	angels	emphasis on						
power heavens	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
heavens glory	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
creation heavens	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
praise heavens	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
creatures heavens	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
knowledge heavens	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
heavens affairs	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
day soul	emphasis on	God	emphasis on	angels	God	angels	emphasis on						
judgment soul	emphasis on	God	emphasis on	angels	God	angels	emphasis on						
soul death	emphasis on	God	emphasis on	angels	God	angels	emphasis on						
soul man	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
angels soul	glorifying of	God	glorifying of	angels	God	angels	emphasis on						
things souls	glorifying of	God	glorifying of	angels	God	angels	emphasis on						

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
light earth	glorifying	of											
Allah chest	God												
pairs cattle	glorifying	of											
food cattle	God												
day years	glorifying	of											
lord sea	God												
works fruit	reward												
fruit deeds	reward												
kind fruit	glorifying	of											
kind fruit	God												
kind fruit	reward												
night sun	glorifying	of											
sun star	God												
sun star	horror	of											
sun star	judgment day												
sun law	glorifying	of											
day moon	God												
moon law	glorifying	of											
gardens springs	God												
gardens springs	glorifying	of											
gardens springs	God												
gardens springs	reward												

The lexico-grammatical patterns of nature and their meanings in Arberry's Translation

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
signs	emphasis	on	people	Present-life	sent	emphasis	on	day	horror	of	surely	emphasis	
people	message	cried		punishment	people	message	last day	dreadful	judgment	people	on message	earth surely	on message
noah people	emphasis	on	people		delivered	emphasis	on	sevens	glorifying		emphasis	heavens	emphasis
	message	believe		believers	people	message	heavens	of God	of God	indeed	on message	surely	on message

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
Moses people	emphasis on message	day say	emphasis on message	guides people	emphasis on message	living soul	glorifying of God	earth firm	glorifying of God	surely man	emphasis on message	hand surely	emphasis on message
people book	emphasis on message	day raised	the horror of judgment day	enter day	Afterlife punishment	one soul	glorifying of God	heavens six	glorifying of God	surely chest	emphasis on message	death surely	emphasis on message
people lot	wrong doers	day cry	day	enter day	reward	lost soul	wrong doers	soul good	believers	surely sea	emphasis on message	river forever	reward an emphasis on the message
people evildoers	wrong doers	day see	emphasis on message	appointed earth	glorifying of God	present life	emphasis on message	man pure	believers			flood surely	
people lies pharaoh	disbelievers	day blown	the horror of judgment day	know earth	glorifying of God	right hand	believers	man equal	Islamic teaching				wrong doers
people	disbelievers	day come	emphasis on message	frustrate earth	wrong doers	holy house	emphasis on message	man evil	wrongdoers				wrongdoers
people city brother	wrongdoers emphasis on message	earth created	glorifying of God	belong heavens	glorifying of God	determined death	emphasis on message	man unthankful	miracle story				
people	glorifying of God	earth say	glorifying of God	heavens know heavens	glorifying of God	one death	disbelievers	hand white	miracle story				
night day			emphasis on message	heavens	glorifying of God	good faces	reward	angels good	believers				Afterlife punishment
Chastisement day	wrongdoers Afterlife punishment	heavens say	emphasis on message	heavens praise	glorifying of God	boiling water	Afterlife punishment	faces blackened	Afterlife punishment				Afterlife punishment
day doom		heavens magnify	glorifying of God	created soul	glorifying of God	firm mountain	glorifying of God	river well pleased	reward				
woe day	Afterlife punishment	soul paid	emphasis on message	resurrecti on	emphasis on message	believing women	believers						
day decision	emphasis on message	soul know	glorifying of God	created man	glorifying of God	corrupt women	glorifying of God	wrongdoers					
day lies	disbelievers	wronged	wrong doers	say man	emphasis on message	humble women	emphasis on message	believers					
day trumpet	the horror of judgment day	soul know	glorifying of God	swallow heaven	Present-life punishment	truthful women	Present-life punishment	believers					
day faces	Afterlife punishment	soul desires	wrongdoers	heart know	believers	hypocrites women	believers	hypocrites					
day faces	reward	life gives	glorifying of God	heart know	glorifying of God	evil women	glorifying of God	wrongdoers					
day hour	emphasis on message	life makes	glorifying of God	heart make	glorifying of God	thousand years	glorifying of God	wrongdoers					

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
day sun	glorifying of God												
ship bounty	emphasis on the message												
moon start	glorifying of God												
garden spring	glorifying of God												
garden spring	reward												
drowned signs	emphasis on the message												
shadows light	emphasis on message												
sun star	glorifying of God												
sun star	horror of judgment day												
signs shadow	emphasis on message												

The lexico-grammatical patterns of nature and their meanings in Saheeh's Translation

N+N	Pragmatic Function	N+V	Pragmatic Function	V+N	Pragmatic Function	ADJ+N	Pragmatic Function	N+ADJ	Pragmatic Function	AV+N	Pragmatic Function	N+AV	Pragmatic Function
signs people	emphasis on message	people believe	believers	sent people	emphasis on message	wrongdoings people	wrongdoers	people clear	emphasis on message	indeed people	emphasis on message	soul indeed	emphasis on message
people pharaoh	wrongdoers	people worship	message	guide people	emphasis on message	praiseworthy earth	glorifying of God	earth spread	glorifying of God	surely hell	emphasis on message	garden forever	reward
Noah people	emphasis on message	day say	message	assemble day	emphasis on message	seven heavens	glorifying of God	earth six	glorifying of God	certainly man	emphasis on message	rivers forever	reward
Moses people	emphasis on message	day come earth	message	believe day	emphasis on message	righteous garden	reward	heavens six	glorifying of God	indeed death	emphasis on message	heart indeed	emphasis on message
people city	wrongdoers glorifying of God	established soul earned	glorifying of God	belong earth created	glorifying of God	two gardens	reward	man good hand	believers				
night day punishment	Afterlife punishment	soul killed	wrongdoers	made earth	glorifying of God	right hand home	believers	white face black women	emphasis on message	miracle story Afterlife punishment			
day horn	the horror of judgment day	soul wronged	wrongdoers	know earth	glorifying of God	first death	disbelievers	beautiful	reward				

Appendix E: The Evaluative SP of Nature in the Qur'an Translations

Pickthall Nature Terms	Frequency	Pos	Neg	Neu	Pos %	Neg %	Neu %	SP
people	238	65	45	1	27.31092	18.90756	0.420168	pos
day	519	106	129	2	20.42389	24.85549	0.385356	neg
earth/s	393	226	14	0	57.50636	3.562341	0	pos
sky/s/heavens	247	173	14	0	70.04049	5.668016	0	pos
soul/s	116	26	11	0	22.41379	9.482759	0	pos
resurrection	77	21	18	0	27.27273	23.37662	0	pos
hell/fire	94	0	51	0	0	54.25532	0	neg
life	140	34	37	1	24.28571	26.42857	0.714286	neg
man/human/men	354	25	13	0	7.062147	3.672316	0	pos
heaven/the garden	139	28	1	0	20.14388	0.719424	0	pos
heart/s	130	25	29	0	19.23077	22.30769	0	neg
hand/s	135	26	7	3	19.25926	5.185185	2.222222	pos
mankind	209	51	6	0	24.40191	2.870813	0	pos
devil/s	54	0	2	0	0	3.703704	0	neg
angel/s	98	28	9	0	28.57143	9.183673	0	pos
night/s	106	44	4	1	41.50943	3.773585	0.943396	pos
homes/houses	112	8	8	0	7.142857	7.142857	0	pos/neg
death	76	23	4	0	30.26316	5.263158	0	pos
face/s	58	3	5	1	5.172414	8.62069	1.724138	neg
water/rain	76	36	11	0	47.36842	14.47368	0	pos
worlds	44	41	2	0	93.18182	4.545455	0	pos
village/town	6	0	0	0	0	0	0	neu
mountain	24	8	5	0	33.33333	20.83333	0	pos
river/s	58	41	0	0	70.68966	0	0	pos
women	109	14	7	5	12.84404	6.422018	4.587156	pos
light	65	27	9	0	41.53846	13.84615	0	pos
chest/s/breast/s	24	10	1	0	41.66667	4.166667	0	pos
sea /s	44	18	5	0	40.90909	11.36364	0	pos
fruit /s	37	3	0	0	8.108108	0	0	pos
sun	35	30	5	0	85.71429	14.28571	0	pos
cattle	33	10	0	0	30.30303	0	0	pos
ship /s	33	4	0	0	12.12121	0	0	pos
tree /s	28	3	3	0	10.71429	10.71429	0	pos/neg
jinn	31	7	8	0	22.58065	25.80645	0	neg
moon	27	17	1	0	62.96296	3.703704	0	pos
year/s	36	5	0	0	13.88889	0	0	pos
springs/water springs	33	6	1	0	18.18182	3.030303	0	pos
wind/s	29	6	1	0	20.68966	3.448276	0	pos
drowning /flood	26	0	5	0	0	19.23077	0	neg

darkness	25	14	0	0	56	0	0	pos
date-palm /s	11	0	0	0	0	0	0	neu
star/s	15	6	0	0	40	0	0	pos
Ali Nature Term	Frequency	Pos	Neg	Neu	Pos %	Neg %	Neu %	SP
people	511	90	75	6	17.61252	14.6771	1.174168	pos
day	520	49	53	5	9.423077	10.19231	0.961538	neg
earth/s	419	250	24	1	59.66587	5.727924	0.238663	pos
sky/s/heavens	247	217	20	0	87.85425	8.097166	0	pos
soul/s	176	43	30	4	24.43182	17.04545	2.272727	pos
resurrection	20	13	26	0	65	130	0	neg
hell/fire	81	0	26	0	0	32.09877	0	neg
life	213	65	38	0	30.51643	17.84038	0	pos
man/human/men	574	42	7	2	7.317073	1.219512	0.348432	pos
heaven/the garden	129	100	4	0	77.51938	3.100775	0	pos
heart/s	161	32	33	0	19.87578	20.49689	0	neg
hand/s	138	16	7	4	11.5942	5.072464	2.898551	pos
mankind	72	9	4	1	12.5	5.555556	1.388889	pos
devil/s	2	0	2	0	0	100	0	neg
angel/s	100	27	10	0	27	10	0	pos
night/s	109	34	4	1	31.19266	3.669725	0.917431	pos
homes/houses	130	14	7	0	10.76923	5.384615	0	pos
death	80	28	4	0	35	5	0	pos
face/s	79	5	15	0	6.329114	18.98734	0	neg
water/rain	69	20	14	2	28.98551	20.28986	2.898551	pos
worlds	49	40	0	0	81.63265	0	0	pos
village/town	17	0	0	0	0	0	0	neu
mountain/hills	45	15	7	0	33.33333	15.55556	0	pos
river/s	55	38	0	0	69.09091	0	0	pos
women	98	7	6	12	7.142857	6.122449	12.2449	neu
light	78	27	8	0	34.61538	10.25641	0	pos
chest/s/breast/s	14	4	1	0	28.57143	7.142857	0	pos
sea /s/ocean /s	42	2	1	0	4.761905	2.380952	0	pos
fruit /s	60	9	2	0	15	3.333333	0	pos
sun	37	21	4	0	56.75676	10.81081	0	pos
cattle	34	10	2	3	29.41176	5.882353	8.823529	pos
ship /s/ark	35	4	2	0	11.42857	5.714286	0	pos
tree /s	30	0	0	0	0	0	0	neu
jinn	32	5	9	0	15.625	28.125	0	neg
moon	28	17	1	0	60.71429	3.571429	0	pos
year/s	36	11	3	0	30.55556	8.333333	0	pos
spring /s/fountain/s	32	5	2	0	15.625	6.25	0	pos
wind/s	30	5	0	0	16.66667	0	0	pos
drowning /flood	30	0	5	0	0	16.66667	0	neg
darkness	30	11	2	1	36.66667	6.666667	3.333333	pos

date-palm /s	8	0	0	0	0	0	0	neu
star/s	13	4	0	0	30.76923	0	0	pos
Arberry Nature Term	Frequency	Pos	Neg	Neu	Pos %	Neg %	Neu %	SP
people	521	114	114	5	21.881	21.881	0.959693	pos/neg
day	462	76	123	0	16.45022	26.62338	0	neg
earth/s	388	206	21	0	53.09278	5.412371	0	pos
sky/s/heavens	191	202	13	0	105.7592	6.806283	0	pos
soul/s	122	15	2	0	12.29508	1.639344	0	pos
resurrection	70	32	39	0	45.71429	55.71429	0	neg
hell/fire	24	0	26	0	0	108.3333	0	neg
life	122	36	27	0	29.5082	22.13115	0	pos
man/human/men	497	22	5	0	4.426559	1.006036	0	pos
heaven/the garden	123	58	0	0	47.15447	0	0	pos
heart/s	132	18	32	0	13.63636	24.24242	0	neg
hand/s	126	9	2	0	7.142857	1.587302	0	pos
mankind	37	9	5	0	24.32432	13.51351	0	pos
devil/s	70	0	3	0	0	4.285714	0	neg
angel/s	92	14	2	0	15.21739	2.173913	0	pos
night/s	96	26	0	0	27.08333	0	0	pos
homes/houses	71	4	1		5.633803	1.408451	0	pos
death	43	4	3	0	9.302326	6.976744	0	pos
face/s	79	11	13	1	13.92405	16.4557	1.265823	neg
water/rain	76	30	16	0	39.47368	21.05263	0	pos
worlds/all beings	27	10	0	0	37.03704	0	0	pos
village/townships	3	0	0	0	0	0	0	neu
mountain/hills	47	17	6	0	36.17021	12.76596	0	pos
river/s	56	37	0	0	66.07143	0	0	pos
women	101	9	6	3	8.910891	5.940594	2.970297	pos
light	57	25	2	0	43.85965	3.508772	0	pos
chest/s/breast/s	33	21	3	0	63.63636	9.090909	0	pos
sea /s	46	17	2	0	36.95652	4.347826	0	pos
fruit /s	38	11	0	0	28.94737	0	0	pos
sun	36	16	1	0	44.44444	2.777778	0	pos
cattle	32	9	1	1	28.125	3.125	3.125	pos
ship /s/ark	35	12	0	0	34.28571	0	0	pos
tree /s	19	0	0	0	0	0	0	neu
jinn	27	8	41	0	29.62963	151.8519	0	pos
moon	27	17	1	0	62.96296	3.703704	0	pos
year/s	26	9	4	1	34.61538	15.38462	3.846154	pos
spring/s/fountain/s	25	7	1	0	28	4	0	pos
wind/s	27	7	1	0	25.92593	3.703704	0	pos
drowning /flood	26	0	6	0	0	23.07692	0	neg
darkness /shadows	24	0	8	0	0	33.33333	0	neg
date-palm /s/palm trees	11	0	0	0	0	0	0	neu

star/s	14	2	0	0	14.28571	0	0	pos
Saheeh Nature Term	Frequency	Pos	Neg	Neu	Pos %	Neg %	Neu %	SP
people/mankind	732	81	133	2	11.06557	18.1694	0.273224	neg
day	497	81	98	2	16.29779	19.71831	0.402414	neg
earth/s	403	215	25	0	53.34988	6.203474	0	pos
sky/s/heavens	261	216	15	0	82.75862	5.747126	0	pos
soul/s	105	18	5	0	17.14286	4.761905	0	pos
resurrection	84	36	40	0	42.85714	47.61905	0	neg
hell/fire	82	0	46	0	0	56.09756	0	neg
life	161	66	15	2	40.99379	9.31677	1.242236	pos
man/human/men	206	8	2	1	3.883495	0.970874	0.485437	pos
heaven/the garden	140	51	0	0	36.42857	0	0	pos
heart/s	124	0	18	0	0	14.51613	0	neg
hand/s	107	10	6	1	9.345794	5.607477	0.934579	pos
mankind	62	9	6	0	14.51613	9.677419	0	pos
devil/s	27	0	1	0	0	3.703704	0	neg
angel/s	126	13	9	0	10.31746	7.142857	0	pos
night/s	105	33	0	1	31.42857	0	0.952381	pos
homes/houses	120	7	0	1	5.833333	0	0.833333	pos
death	88	26	6	0	29.54545	6.818182	0	pos
face/s	65	4	4	0	6.153846	6.153846	0	pos/neg
water/rain	54	7	16	0	12.96296	29.62963	0	neg
worlds	71	53	1	0	74.64789	1.408451	0	pos
village/town	3	0	0	0	0	0	0	neu
mountain/hills	55	13	3	0	23.63636	5.454545	0	pos
river/s	59	51	0	0	86.44068	0	0	pos
women	103	13	7	5	12.62136	6.796117	4.854369	pos
light	52	28	8	0	53.84615	15.38462	0	pos
chest/s/breast/s	44	19	1		43.18182	2.272727	0	pos
sea /s	46	13	1	0	28.26087	2.173913	0	pos
fruit /s	52	18	1	0	34.61538	1.923077	0	pos
sun	35	19	1	1	54.28571	2.857143	2.857143	pos
cattle	7	0	0	0	0	0	0	neu
ship /s	36	13	0	0	36.11111	0	0	pos
tree /s	49	17	4	1	34.69388	8.163265	2.040816	pos
jinn	33	5	11	0	15.15152	33.33333	0	neg
moon	28	17	1	0	60.71429	3.571429	0	pos
year/s	32	8	2	0	25	6.25	0	pos
spring/s/fountain/s	26	7	3	0	26.92308	11.53846	0	pos
wind/s	31	9	0	0	29.03226	0	0	pos
drowning	25	0	3	0	0	12	0	neg
darkness	32	0	12	0	0	37.5	0	neg
date-palm /s	21	1	0	0	0	0	0	pos
star/s	19	6	0	0	31.57895	0	0	pos

Haleem Term	Nature	Frequency	Pos	Neg	Neu	Pos %	Neg %	Neu %	SP
people		785	93	55	2	11.84713	7.006369	0.254777	pos
day		440	112	103	3	25.45455	23.40909	0.681818	pos
earth/s		367	190	20	0	51.77112	5.449591	0	pos
sky/s/heavens		276	186	6	0	67.3913	2.173913	0	pos
soul/s		114	22	7	1	19.29825	6.140351	0.877193	pos
resurrection		74	33	35	0	44.59459	47.2973	0	neg
hell		91	0	27	0	0	29.67033	0	neg
life		203	62	30	0	30.54187	14.77833	0	pos
man/human/men		244	18	12	0	7.377049	4.918033	0	pos
heaven/the garden		136	72	1	0	52.94118	0.735294	0	pos
heart/s		131	28	20	2	21.37405	15.26718	1.526718	pos
hand/s		97	10	5		10.30928	5.154639	0	pos
mankind		23	4	2	1	17.3913	8.695652	4.347826	pos
devil/s		7	0	0	0	0	0	0	neg
angel/s		96	14	4	0	14.58333	4.166667	0	pos
night/s		108	46	1	1	42.59259	0.925926	0.925926	pos
homes/houses		144	14	22	0	9.722222	15.27778	0	neg
death		72	19	0	0	26.38889	0	0	pos
face/s		83	9	14	0	10.84337	16.86747	0	neg
water/rain		83	45	14	0	54.21687	16.86747	0	pos
worlds		31	33	1	0	106.4516	3.225806	0	pos
village/town		47	0	1	0	0	2.12766	0	neg
mountain		59	13	3	0	22.0339	5.084746	0	pos
river/s		18	38	0	0	211.1111	0	0	pos
women		91	12	6	9	13.18681	6.593407	9.89011	neu
light		52	30	7	0	57.69231	13.46154	0	pos
chest/s/breast/s		5	0	0	0	0	0	0	neu
sea /s		43	19	3	0	44.18605	6.976744	0	pos
fruit /s		34	3	0	0	8.823529	0	0	pos
sun		34	22	0	1	64.70588	0	2.941176	pos
cattle		9	0	0	0	0	0	0	neu
ship /s/ark		32	9	0	0	28.125	0	0	pos
tree /s		40	4	2	0	10	5	0	pos
jinn		37	3	8	0	8.108108	21.62162	0	neg
moon		26	18	0	0	69.23077	0	0	pos
year/s		26	3	1	0	11.53846	3.846154	0	pos
springs /fountain		26	3	1		11.53846	3.846154	0	pos
wind /s		31	8	0	0	25.80645	0	0	pos
drowning		22	4	0	0	18.18182	0	0	pos
darkness		28	4	6		14.28571	21.42857	0	neg
date-palm/s		14	9	0	0	64.28571	0	0	pos
star/s		16	4	0	0	25	0	0	pos

Appendix F: A Diachronic Analysis of English Stop-words in the English translations of the Qur'an (1930-2004)²⁵⁴

This section presents the results of an experiment which explores the historical change in the use of English function words in the translations of the Qur'an in the period between 1930 and 2014. It was conducted following Rayson's and Garside's (2000) approach of comparing corpora using frequency profiling; and which relies on the use of the Log-Likelihood association measure to rank the compared frequencies between corpora. This comparison reveals the decline in the use of some words, the increase in others, as well as *lock words*, which are steady in their use over time. Pickthall's translation was treated as a subcorpora that was compared to the second corpus, which is comprised of the other translations in this research. Results of this study were useful in determining the archaic English functions words that were used by the earliest translations of the Quran. The researcher appended them to the list of stop-words (i.e., functions words) available in the NLTK package in Python during the text pre-processing stage to explore the collocations of natural phenomena terms in the Quran. The following figure shows a sample of the decline of the use of the words 'unto' as a preposition and 'thee'²⁵⁵ as a pronoun. It also shows the decline of 'hath' and 'doth' in contrast to 'has' and 'does', which demonstrate an increase in use over time.

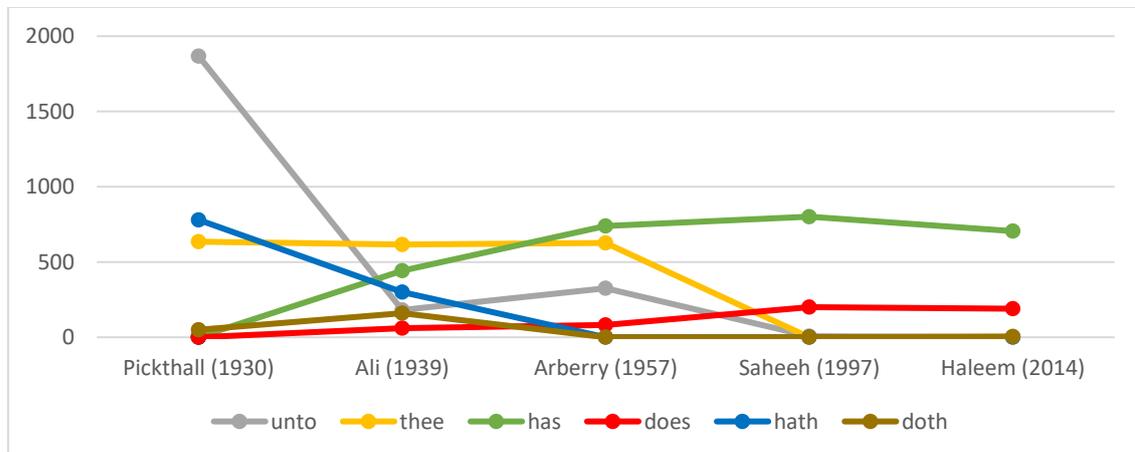


Figure 1: The decline and increase of six English stop-words in the translation of the Qur'an (1930-2004)

Finally, the following table shows the results of this experiment with the Log-Likelihood score used to rank the difference between the corpora from the highest to lowest. It also demonstrates the

²⁵⁴ These were calculated by using the *Significance and Effect Calculator*, [<http://ucrel.lancs.ac.uk/llwizard.html>]

²⁵⁵ *Thee* is an old-fashioned, poetic, or religious word for the pronoun 'you' when talking to only one person. Available from: [<https://www.collinsdictionary.com/dictionary/english/thee>], [Accessed 1 December 2019]

relative frequencies of the English stop-words, which were obtained both from the Python list and the close reading of the Pickthall translation.

Key for table:256

O1 is observed frequency in Corpus 1 [Pickthall, 1930].

O2 is observed frequency in Corpus 2 [the other four translations].

% 1 and %2 values show relative frequencies in the texts.

+ **indicates overuse** in O1 relative to O2.

- **indicates underuse** in O1 relative to O2.

Table 1: The comparison between Pickthall's stop-word frequencies and the four other translations.

<i>Word</i>	<i>O1</i>	<i>1%</i>	<i>O2</i>	<i>2%</i>	<i>Log Likelihood</i>
unto	1868	1.2	514	0.08 +	3774.36
ye	1861	1.2	2019	0.32 +	1531.98
hath	780	0.5	300	0.05 +	1375.13
has	1	0	2688	0.43 -	1178.65
you	1521	0.98	13563	2.17 -	1068.77
to	2074	1.34	15598	2.50 -	838.22
which	955	0.62	1797	0.29 +	327.57
thereof	166	0.11	81	0.01 +	259.32
does	0	0	532	0.09 -	236.16
do	471	0.3	3652	0.59 -	212.06
except	33	0.02	861	0.14 -	206.23
thou	760	0.49	1592	0.26 +	200.52
thee	635	0.41	1243	0.20 +	198.8
whoever	6	0	480	0.08 -	167.79
about	32	0.02	692	0.11 -	148.31
what	501	0.32	3471	0.56 -	147.71
that	2552	1.65	7767	1.24 +	142.67
whom	409	0.26	870	0.14 +	103.52
thy	440	0.28	965	0.15 +	102.15
till	104	0.07	106	0.02 +	91.71
off	110	0.07	120	0.02 +	89.98
wherewith	58	0.04	29	0.00 +	89.37
most	102	0.07	974	0.16 -	87.02
of	4857	3.13	16752	2.69 +	86.96
shall	364	0.23	2387	0.38 -	84.76
and	7871	5.08	28148	4.51 +	83.72
theirs	87	0.06	88	0.01 +	77.34

256 The key is borrowed from [<http://ucrel.lancs.ac.uk/llwizard.html>], my brackets.

<i>Word</i>	<i>O1</i>	<i>1%</i>	<i>O2</i>	<i>2%</i>	<i>Log Likelihood</i>
nay	148	0.1	225	0.04 +	76.61
whereof	55	0.04	33	0.01 +	75.78
wherein	90	0.06	98	0.02 +	73.78
over	93	0.06	843	0.14 -	68.54
so	531	0.34	3073	0.49 -	65.04
besides	3	0	194	0.03 -	64.74
anything	5	0	211	0.03 -	62.26
towards	1	0	160	0.03 -	62.1
out	118	0.08	917	0.15 -	53.54
down	160	0.1	1140	0.18 -	52.77
whatever	24	0.02	341	0.05 -	51.82
any	156	0.1	1097	0.18 -	48.85
may	405	0.26	1084	0.17 +	45.89
anyone	12	0.01	231	0.04 -	45.69
him	1324	0.85	4343	0.70 +	40.83
perhaps	1	0	99	0.02 -	35.97
hast	100	0.06	187	0.03 +	34.78
within	16	0.01	226	0.04 -	34.14
thus	135	0.09	288	0.05 +	33.9
therein	220	0.14	551	0.09 +	32.84
when	970	0.63	3151	0.51 +	32.7
have	1116	0.72	5399	0.87 -	32.66
amongst	0	0	72	0.01 -	31.96
just	13	0.01	195	0.03 -	31.27
such	222	0.14	575	0.09 +	28.99
your	773	0.5	3790	0.61 -	26.17
is	3185	2.05	11571	1.85 +	25.59
will	1679	1.08	7725	1.24 -	25.48
through	26	0.02	260	0.04 -	25.1
who	2028	1.31	7193	1.15 +	24.61
into	104	0.07	685	0.11 -	24.38
up	92	0.06	619	0.10 -	23.99
thine	25	0.02	25	0.00 +	22.49
ever	147	0.09	366	0.06 +	22.44
because	154	0.1	395	0.06 +	20.93
until	65	0.04	460	0.07 -	20.87
this	422	0.27	2151	0.34 -	20.78
therewith	37	0.02	54	0.01 +	20.46
but	988	0.64	4639	0.74 -	20.04
yourself	0	0	44	0.01 -	19.53
henceforth	6	0	0	0.00 +	19.37
its	121	0.08	733	0.12 -	19.14
then	832	0.54	2811	0.45 +	18.99

<i>Word</i>	<i>O1</i>	<i>1%</i>	<i>O2</i>	<i>2%</i>	<i>Log Likelihood</i>
during	1	0	58	0.01 -	18.84
he	2259	1.46	8204	1.32 +	18.27
being	34	0.02	276	0.04 -	17.86
nor	295	0.19	899	0.14 +	17.61
shouldst	18	0.01	18	0.00 +	16.2
rather	10	0.01	123	0.02 -	15.9
mere	24	0.02	31	0.00 +	15.89
mayest	0	0	32	0.01 -	14.2
throughout	0	0	32	0.01 -	14.2
other	133	0.09	750	0.12 -	13.9
by	387	0.25	1908	0.31 -	13.88
all	512	0.33	2460	0.39 -	13.85
themselves	100	0.06	589	0.09 -	13.56
didst	33	0.02	57	0.01 +	13.55
or	464	0.3	2237	0.36 -	13.05
very	16	0.01	150	0.02 -	12.97
according	12	0.01	125	0.02 -	12.87
with	955	0.62	4345	0.70 -	12.12
would	266	0.17	1344	0.22 -	12.1
why	48	0.03	319	0.05 -	11.86
the	7638	4.93	29405	4.71 +	11.77
ourselves	6	0	80	0.01 -	11.36
in	2162	1.39	9413	1.51 -	11.1
wherever	5	0	72	0.01 -	11.09
were	398	0.26	1915	0.31 -	10.94
hadst	13	0.01	14	0.00 +	10.79
without	42	0.03	281	0.05 -	10.69
thereon	11	0.01	11	0.00 +	9.9
be	1212	0.78	5369	0.86 -	9.32
whereby	11	0.01	12	0.00 +	9
could	63	0.04	374	0.06 -	8.93
amid	8	0.01	7	0.00 +	8.21
midst	3	0	48	0.01 -	8.17
among	227	0.15	1114	0.18 -	7.76
too	25	0.02	175	0.03 -	7.69
for	1985	1.28	7451	1.19 +	7.48
if	725	0.47	3259	0.52 -	7.44
though	97	0.06	282	0.05 +	7.22
also	34	0.02	218	0.03 -	7.14
under	8	0.01	77	0.01 -	6.97
upon	284	0.18	1351	0.22 -	6.81
wont	17	0.01	31	0.00 +	6.25
them	2322	1.5	8825	1.41 +	5.93

<i>Word</i>	<i>O1</i>	<i>1%</i>	<i>O2</i>	<i>2%</i>	<i>Log Likelihood</i>
perchance	1	0	25	0.00 -	5.85
on	589	0.38	2642	0.42 -	5.81
must	33	0.02	80	0.01 +	5.55
wast	18	0.01	36	0.01 +	5.35
however	10	0.01	15	0.00 +	5.29
been	241	0.16	1137	0.18 -	5.21
shalt	8	0.01	69	0.01 -	5.09
couldst	14	0.01	26	0.00 +	4.94
more	83	0.05	431	0.07 -	4.79
no	496	0.32	2222	0.36 -	4.78
each	78	0.05	407	0.07 -	4.68
thither	5	0	5	0.00 +	4.5
ah	29	0.02	72	0.01 +	4.47
either	9	0.01	71	0.01 -	4.3
are	1354	0.87	5118	0.82 +	4.14
moreover	1	0	20	0.00 -	4.07
thereby	34	0.02	90	0.01 +	4.05
beneath	13	0.01	91	0.01 -	4
am	141	0.09	468	0.08 +	3.88
at	175	0.11	823	0.13 -	3.63
anywhere	0	0	8	0.00 -	3.55
an	284	0.18	1007	0.16 +	3.46
i	665	0.43	2466	0.40 +	3.44
anybody	1	0	0	0.00 +	3.23
should	172	0.11	802	0.13 -	3.17
again	55	0.04	167	0.03 +	3.12
forthwith	0	0	7	0.00 -	3.11
having	15	0.01	95	0.02 -	2.97
against	177	0.11	818	0.13 -	2.87
did	214	0.14	975	0.16 -	2.78
ours	8	0.01	15	0.00 +	2.77
therefore	25	0.02	139	0.02 -	2.38
from	1074	0.69	4549	0.73 -	2.31
dost	8	0.01	55	0.01 -	2.28
myself	8	0.01	55	0.01 -	2.28
their	1252	0.81	5260	0.84 -	1.9
hers	2	0	2	0.00 +	1.8
my	487	0.31	1831	0.29 +	1.75
further	8	0.01	51	0.01 -	1.63
every	131	0.08	595	0.10 -	1.62
we	1868	1.2	7280	1.17 +	1.51
before	295	0.19	1282	0.21 -	1.44
wouldst	13	0.01	35	0.01 +	1.43

<i>Word</i>	<i>O1</i>	<i>1%</i>	<i>O2</i>	<i>2%</i>	<i>Log Likelihood</i>
they	2882	1.86	11313	1.81 +	1.41
those	1196	0.77	4999	0.80 -	1.4
whereon	11	0.01	29	0.00 +	1.33
these	100	0.06	455	0.07 -	1.27
oft	1	0	11	0.00 -	1.23
himself	56	0.04	264	0.04 -	1.2
doth	49	0.03	165	0.03 +	1.16
here	22	0.01	112	0.31	1.07
although	8	0.01	47	0.01 -	1.07
behind	62	0.04	215	0.03 +	1.04
might	98	0.06	441	0.07 -	1.03
whenever	13	0.01	70	0.01 -	0.99
own	103	0.07	461	0.07 -	0.97
had	349	0.23	1324	0.21 +	0.95
thereat	3	0	6	0.00 +	0.89
oer	0	0	2	0.00 -	0.89
me	428	0.28	1638	0.26 +	0.85
her	101	0.07	367	0.06 +	0.81
his	998	0.64	3890	0.62 +	0.8
a	2280	1.47	8986	1.44 +	0.78
yourselves	62	0.04	278	0.04 -	0.61
than	192	0.12	819	0.13 -	0.54
yet	105	0.07	457	0.07 -	0.53
unless	34	0.02	119	0.02 +	0.5
our	480	0.31	1864	0.30 +	0.48
was	380	0.25	1590	0.25 -	0.47
thysself	11	0.01	35	0.01 +	0.44
manysoever	0	0	1	0.00 -	0.44
thenceforth	0	0	1	0.00 -	0.44
how	193	0.12	818	0.13 -	0.43
both	48	0.03	213	0.03 -	0.38
herself	3	0	8	0.00 +	0.35
she	80	0.05	299	0.05 +	0.34
otherwise	3	0	17	0.00 -	0.32
yours	19	0.01	66	0.01 +	0.31
itself	7	0	22	0.00 +	0.31
whilst	1	0	7	0.00 -	0.31
betwixt	1	0	2	0.00 +	0.3
furthermore	1	0	2	0.00 +	0.3
while	126	0.08	481	0.08 +	0.27
us	455	0.29	1783	0.29 +	0.25
where	29	0.02	129	0.02 -	0.24
yea	13	0.01	45	0.01 +	0.22

<i>Word</i>	<i>O1</i>	<i>1%</i>	<i>O2</i>	<i>2%</i>	<i>Log Likelihood</i>
only	271	0.17	1058	0.17 +	0.2
below	4	0	20	0.00 -	0.16
it	1546	1	6286	1.01 -	0.13
can	205	0.13	804	0.13 +	0.11
therefrom	14	0.01	61	0.01 -	0.07
between	143	0.09	589	0.09 -	0.06
as	654	0.42	2607	0.42 +	0.05
not	1849	1.19	7474	1.20 -	0.03
there	472	0.3	1916	0.31 -	0.03
whose	85	0.05	335	0.05 +	0.03
after	196	0.13	779	0.12 +	0.02
neither	72	0.05	294	0.05 -	0.01
another	112	0.07	449	0.07 +	0
above	62	0.04	251	0.04 -	0
often	6	0	24	0.00 +	0

Appendix G: Examples of the Five Most Frequent Discourse Prosodies in the Qur'an

This appended section discusses examples of the five most frequent discourse prosodies of natural phenomena in the Qur'an. It is not part of the qualitative analysis of this research; it is an attempt on the researcher's part to reflect on the five most frequent uncovered discourse prosodies of nature in the Qur'an and comment on them with examples.²⁵⁷

a. Glorifying God through nature

The concept of praise is defined as the acknowledgement of the infinite bounties of Allah. This signifies obedience on the part of the Muslim who acknowledges that Allah alone should be praised for everything. It is an invitation to believe in God through contemplation and praise. This echoes description of the role of natural phenomena in the Qur'an by Fazlur Rahman (2009) who writes:

Often, they portray God's unlimited power and majesty to invite man to believe in Him, or depict His infinite mercy and require man to be grateful to Him. In both cases, nature's magnitude and utility for man, as well as the stability and regularity of natural phenomena, are stressed (p.66).

- 1- God, with his power and might, created natural phenomena with perfection and knows all that is in the universe.

الَّذِي خَلَقَ سَبْعَ سَمَاوَاتٍ طِبَاقًا مَا تَرَى فِي خَلْقِ الرَّحْمَنِ مِنْ تَفَوتٍ فَأَرْجِعِ الْبَصَرَ هَلْ تَرَى
مِنْ فُتُورٍ ﴿٣﴾

ثُمَّ ارْجِعِ الْبَصَرَ كَرَّتَيْنِ يَنْقَلِبْ إِلَيْكَ الْبَصَرُ خَاسِئًا وَهُوَ حَسِيرٌ ﴿٤﴾

Saheeh International: (And) who created seven **heavens** in layers. You do not see in the creation of the Most Merciful any inconsistency. So return (your) vision (to the **sky**); do you see any breaks? Look again and again- your sight will return to you frustrated (in the attempt to find any discontinuity or irregularity) and fatigued.

Verses (67:3-4)

وَرَأَى الْجِبَالَ تَحْسَبُهَا جَمَادًا وَهِيَ نُورٌ مُرْتَمٍ السَّحَابِ صُغُرٌ اللَّهُ الَّذِي أَنْفَقَ كُلَّ شَيْءٍ وَإِنَّهُ خَبِيرٌ
بِمَا تَعْمَلُونَ ﴿٨٨﴾

²⁵⁷ When reading the examples, the words printed in bold in the English translations of the verses in the examples below were identified as natural phenomena in this research.

Saheeh International: And you see the **mountains**, thinking them rigid, while they will pass as the passing of clouds. (It is) the work of Allah, who perfected all things. Indeed, He is Acquainted with that which you do.

Verse (27:88)

2- The perfection of the creation of nature is that it works with regularity.

يُولِجُ اللَّيْلَ فِي النَّهَارِ وَيُؤَلِّجُ النَّهَارَ فِي اللَّيْلِ وَسَخَّرَ الشَّمْسَ وَالْقَمَرَ كُلًّا
يَجْرِي لِأَجَلٍ مُّسَمًّى ذَلِكُمْ اللَّهُ رَبُّكُمْ لَهُ الْمُلْكُ وَالَّذِينَ تَدْعُونَ مِنْ دُونِهِ
مَا يَمْلِكُونَ مِنْ فِطْمِيرٍ ﴿١٣﴾

Saheeh International: He causes the night to enter the **day**, and He causes the **day** to enter the **night** and has subjected the **sun** and the **moon** - each running (its course) for a specified term. That is Allah, your Lord; to Him belongs sovereignty. And those whom you invoke other than Him do not possess (as much as) the membrane of a date seed.

Verse (35:13)

This again asserts Fazlur Rahman's (2009) observation of the concepts of nature in the Qur'an when he writes:

Nature is so well-knit and works with such regularity that it is the prime miracle of God... References to phenomena like the regularity of the day following the night and the night following the day, the rainy season when the earth is quickened following a dry season when it had been parched and dead, are strewn through the pages of the Qur'an (p.68).

3- God subjected natural phenomena to man and made them subservient to him, and he should be grateful.

إِنَّ رَبَّكُمْ اللَّهُ الَّذِي خَلَقَ السَّمَوَاتِ وَالْأَرْضَ فِي سِتَّةِ أَيَّامٍ ثُمَّ اسْتَوَىٰ عَلَى الْعَرْشِ
يُعْشَىٰ اللَّيْلَ النَّهَارَ يَطْلُبُهُ حَثِيثًا وَالشَّمْسَ وَالْقَمَرَ وَالنُّجُومَ مُسْحَرَاتٍ يَا أَيُّهَا آلَ آدَمَ
الْحَافِظُ وَالْآتَمُّ تَبَارَكَ اللَّهُ رَبُّ الْعَالَمِينَ ﴿٥٤﴾

Saheeh International: Indeed, your Lord is Allah, who created the **heavens** and **earth** in six **days** and then established Himself above the Throne. He covers the **night** with the **day**, (another night) chasing it rapidly; and (He created) the **sun**, the **moon**, and the **stars**, subjected by His command. Unquestionably, His is the creation and the command; blessed is Allah, Lord of the worlds. (Verse 7:54)

4- God, Who should be glorified and praised, has the ultimate power and dominion over the universe and all natural causes work for Him.

فَلِلَّهِ الْحَمْدُ رَبِّ السَّمَوَاتِ وَرَبِّ الْأَرْضِ رَبِّ الْعَالَمِينَ ﴿٣٦﴾

Saheeh International: Then, to Allah belongs (all) praise - Lord of the **heavens** and Lord of the **earth**, Lord of the **worlds**. (Verse 45:36)

b. The emphasis on the message of the Qur'an

In a tradition related by Ali Ibn Abi Talib²⁵⁸, we are told that the Prophet said, 'There shall be great sedition after me.' 'Ali asked how such calamity could be averted. The Prophet answered:

By means of the Book of God! In it is the report concerning those who were before you, the narrative of what is the report concerning those who were before you, the narrative of what is to come after you, and the criterion of judgment among you... Whoever seeks guidance in anything other than it, God shall cause him to go astray. It is the rope of God; it is the 'wise remembrance' and 'the straight way.' With it, hearts shall not swerve nor tongues utter confusion. The learned shall never be sated of it. It shall not wear out from constant use, nor will its marvels ever be exhausted. ...Whoever utters it speaks the truth, and whoever abides by it shall have his rich reward. Whoever judges by it shall judge justly, and whoever call others to it shall be guided to the straightway (Ibn Kathīr in Abdul-Rahman, 2009, p.434).

The tradition above provides an intact summary of the role of the Qur'an in man's life; a motif which also co-occurs with natural phenomena throughout the Qur'an. The following statements can be made with examples of the discourse prosody of the emphasis on the message in the Qur'an (natural phenomena terms in boldface).

- 1- The Qur'an is described as a light to guide man in life.

يَأَيُّهَا النَّاسُ قَدْ جَاءَكُمْ بُرْهَانٌ مِنْ رَبِّكُمْ وَأَنْزَلْنَا إِلَيْكُمْ نُورًا مُبِينًا ﴿١٧٤﴾

Saheeh International: O **mankind**, there has come to you a conclusive proof from your Lord, and We have sent down to you a clear **light**. (Verse 4:174)

فَتَأْمُرُوا بِالْبِرِّ سَوَاءً وَالنُّورِ الَّذِي أَنْزَلْنَا وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ ﴿٨﴾

Yusuf Ali: Believe, therefore, in Allah and His Messenger, and in the **Light** which we have sent down. And Allah is well acquainted with all that ye do. (Verse 64:8)

²⁵⁸ Ali Ibn Abi Talib is the cousin and son-in-law of Muhammad (peace be upon him), fourth caliph of Muslims, and the first male convert to Islam and second convert after Muhammad's wife Khadijah. He married Muhammad's daughter Fatimah, and he is the father of sons Hasan and Husayn. He participated in most expeditions during Muhammad's lifetime. Shiis believe that he was appointed by Muhammad as successor and place him next to God and Muhammad as the centre of religious belief. Sunnis acknowledge him only as the fourth caliph, a distinguished judge, pious believer, and brave warrior. His rule as caliph was marked by political crisis and civil strife, and he was assassinated while praying in a mosque in 660. [http://www.oxfordislamicstudies.com], [Accessed 18 May 2019].

يُرِيدُونَ أَنْ يُطْفِئُوا نُورَ اللَّهِ بِأَفْوَاهِهِمْ وَيَأْبَى اللَّهُ لِأَنَّ يَسْمُرُوهُ، وَلَوْ كَرِهَ
الْكَافِرُونَ ﴿٣٢﴾

Saheeh International: They want to extinguish the **light** of Allah with their mouths, but Allah refuses except to perfect His **light**, although the disbelievers dislike it. (Verse 9:32)

Tafsir Aljalalyin: They desire to extinguish God's light, His Law and His proofs, with their tongues, with what they say about Him; and God refuses but to perfect, to make manifest, His light, even though the disbelievers be averse, to this.

- 2- All of the creations of nature in the Qur'an are mere signs of the truth of the message of the Qur'an and the unity of God. It is as Fazlur Rahman remarks, 'the Qur'an does not 'prove' God but 'points to' Him from the existing universe' (p.10).

أَوَلَمْ يَنْظُرُوا فِي مَلَكُوتِ السَّمَوَاتِ وَالْأَرْضِ وَمَا خَلَقَ اللَّهُ مِنْ شَيْءٍ وَأَنْ عَسَى أَنْ
يَكُونَ قَدِ اقْتَرَبَ أَجَلُهُمْ فَبِأَيِّ حَدِيثٍ بَعْدَهُ يُؤْمِنُونَ ﴿١٨٥﴾

Saheeh International: Do they not look into the realm of the **heavens** and the **earth** and everything that Allah has created and (think) that perhaps their appointed time has come near? So in what statement hereafter will they believe? (Verse 7:185)

أَفَلَمْ يَرَوْا إِلَى مَا بَيْنَ أَيْدِيهِمْ وَمَا خَلْفَهُمْ مِنْ السَّمَاءِ وَالْأَرْضِ إِنَّ نَسْأًا نَحْصِفُ بِهِمْ
الْأَرْضَ أَوْ نَسْقِطُ عَلَيْهَا حُمْلاً مِنَ السَّمَاءِ إِنَّ فِي ذَلِكَ لَآيَةً لِكُلِّ عَبْدٍ مُنِيبٍ ﴿٧﴾

Yusuf Ali: See they not what is before them and behind them, of the **sky** and the **earth**? If We wished, We could cause the **earth** to swallow them up, or cause a piece of the **sky** to fall upon them. Verily in this is a Sign for every devotee that turns to Allah (in repentance). (Verse 34:9)

أَوَلَمْ يَرَوْا إِلَى الْأَرْضِ كَرَأَيْنَا فِيهَا مِنْ كُلِّ زَوْجٍ كَرِيمٍ ﴿٧﴾

Pickthall: Have they not seen the **earth**, how much of every fruitful kind We make to grow therein? (Verse 26:7)

- 3- The documentation and recurrent references of natural phenomena in the bounties of the promised life in the reward of those who believe and likewise the promised penalty for those who disbelieve or transgress are mentioned in the Qur'an as proof of its credible message.

يَوْمَ نَجْمَعُكُمْ لِيَوْمِ الْجَمْعِ ذَلِكَ يَوْمُ النَّعَابِ وَمَنْ يُؤْمِنْ بِاللَّهِ وَعَمِلْ صَالِحًا نَكْفِرْ عَنْهُ
سَيِّئَاتِهِ، وَيُدْخِلْهُ جَنَّاتٍ تَجْرِي مِنْ تَحْتِهَا الْأَنْهَارُ خَالِدِينَ فِيهَا أَبَدًا ذَلِكَ الْفَوْزُ
الْعَظِيمُ ﴿١٩﴾

Saheeh International: The **Day** He will assemble you for the **Day** of Assembly - that is the Day of Deprivation. And whoever believes in Allah and does righteousness - He will remove from him

his misdeeds and admit him to **gardens** beneath which **rivers** flow, wherein they will abide forever. That is the great attainment. (Verse 64:9)

كُلُّ نَفْسٍ ذَائِقَةُ الْمَوْتِ وَإِنَّمَا تُوَفَّقُ أُجُورَكُمْ يَوْمَ الْفَيْكَمَةِ فَمَنْ زُحِرَ عَنْ
النَّارِ وَأُدْخِلَ الْجَنَّةَ فَقَدْ فَازَ وَمَا الْحَيَاةُ الدُّنْيَا إِلَّا مَتَاعُ الْعُرُورِ ﴿١٨٥﴾

Pickthall: Every **soul** will taste of **death**. And ye will be paid on the **Day** of Resurrection only that which ye have fairly earned. Whoso is removed from the **Fire** and is made to enter paradise; he indeed is triumphant. The **life** of this **world** is but the comfort of illusion. (Verse 3:185)

- 4- The parables of the previous prophets and the miracles they performed are also signs of the truth of the message and a warning for those who disbelieve. In this regard, Robinson (1999) writes:

The Qur'an mentions several miracles or 'signs' as it calls them, which previous prophets performed by Allah's permission. For instance, it narrates how Moses' staff turned into a serpent and how Jesus created birds from clay, healed the blind and raised the dead (3.49) (p.28).

The following example illustrates an instance of the Qur'an describing the miracles granted to Solomon and David:

وَوَرِثَ سُلَيْمَانُ دَاوُدَ وَقَالَ يَا أَيُّهَا النَّاسُ عَلِمْنَا مَنطِقَ الطَّيْرِ وَأَوْتَيْنَا مِمَّنْ كُلِّ شَيْءٍ وَإِن هَذَا
هُوَ الْفَضْلُ الْمُبِينُ ﴿١٦﴾

Saheeh International: And Solomon inherited David. He said, 'O **people**, we have been taught the language of **birds**, and we have been given from all things. Indeed, this is evident bounty.' (Verse 27:19)

Tafsir Aljalalyin: And Solomon inherited from David, prophethood, and knowledge exclusively from among all his other children; and he said, 'O people, we have been taught the speech of the birds, that is, the ability to understand their sounds, and we have been given of all things, given to prophets and kings. Indeed this, gift, is the manifest, the evident, favour'.

- 5- The Qur'an invites man to use his cognitive skills in thinking, understanding and remembering. It also tells him to witness the truth of its message through engaging his senses to see the wonders of nature and the consequences of those who were deniers of God all to reach the higher state of believing in the unity of God and worship Him. Hence, the recurring Qur'anic invitations and exhortations (Fazular Fazlur Rahman 2009, p.11), 'Do you not reflect?' 'Do you not think?' 'Do you not take heed' usually occur with natural phenomena terms.

هَذَا بَلَّغٌ لِلنَّاسِ وَلِيُنذَرُوا بِهِ، وَلِيَعْلَمُوا أَنَّمَا هُوَ إِلَهٌ وَجِدٌ وَيَلْبِذُوا إِلَى الْآبَتِ ﴿٥٤﴾

Yusuf Ali: Here is a Message for **mankind**: Let them take warning therefrom and let them know that He is (no other than) One Allah: let **men** of understanding take heed. Verse (14:52)

وَلَقَدْ بَعَثْنَا فِي كُلِّ أُمَّةٍ رَسُولًا أَنِ اعْبُدُوا اللَّهَ وَاجْتَنِبُوا الطَّاغُوتَ فَمِنْهُمْ
مَنْ هَدَى اللَّهُ وَمِنْهُمْ مَنْ حَقَّتْ عَلَيْهِ الضَّلَالَةُ فَمَسِيرُوا فِي الْأَرْضِ فَانظُرُوا كَيْفَ
كَانَ عَاقِبَةُ الْمُكَذِّبِينَ ﴿٣٦﴾

Pickthall: And verily We have raised in every nation a messenger, (proclaiming): Serve Allah and shun false gods. Then some of them (there were) whom Allah guided, and some of them (there were) upon whom error had just held. Do but travel in the **land** and see the nature of the consequence for the deniers! (Verse 16:36)

- 6- The recurrent contrast between darkness (disbelief) and light (faith) in the Qur'an is a motif of the emphasis on the truth of the message.

هُوَ الَّذِي يُزِيلُ عَلَى عَبْدِهِ مَا يَشَاءُ وَيُنَزِّلُ الْإِنشَارَ مِنَ السَّمَاءِ وَإِلَى الْأَرْضِ وَإِنَّ اللَّهَ لَكَلِيمٌ
لَرءُوفٌ رَحِيمٌ ﴿٩﴾

Pickthall: He it is Who sendeth down clear revelations unto His slave, that He may bring you forth from **darkness** unto **light**; and lo! for you, Allah is Full of Pity, Merciful. (Verse 57:9)

Tafsir Aljalalyin: It is He Who sends down upon His servant clear signs, the signs of the Qur'an that He may bring you forth from the darkness (from) disbelief to the light, (to) faith. For truly God is Kind, Merciful to you, in bringing you forth from disbelief to faith.

وَمَا يَسْتَوِي الْأَعْمَى وَالْبَصِيرُ ﴿١١﴾ وَلَا الظُّلُمَاتُ وَلَا النُّورُ ﴿٢٠﴾

Saheeh International: Not equal are the blind and the seeing; Nor are the darknesses and the light, (Verses 35:19-20)

Tafsir Aljalalyin: Not equal are the blind and the seeing man, nor darkness — disbelief — and light — faith;

أَوْ مَنْ كَانَ مَيِّتًا فَأَحْيَيْنَاهُ وَجَعَلْنَا لَهُ نُورًا يَمْشِي بِهِ فِي النَّاسِ كَمَنْ مَثَلَهُ فِي
الظُّلُمَاتِ لَيْسَ بِخَارِجٍ مِنْهَا كَذَلِكَ زُيِّنَ لِلْكَافِرِينَ مَا كَانُوا يَعْمَلُونَ ﴿١٢٢﴾

Saheeh International: And is one who was dead and We gave him **life** and made for him **light** by which to walk among the **people** like one who is in **darkness**, never to emerge therefrom? Thus it has been made pleasing to the disbelievers that which they were doing. (Verse 6:122)

Asbab al-Nuzul (i.e., the contexts and occasions of the Revelation of the Qur'an) al-Wahidi: The following was revealed regarding Abū Jahl²⁵⁹ and others: Why, is he who was dead, through unbelief, and We gave him life, through guidance, and appointed for him a light by which to walk among people, distinguishing thereby the truth from falsehood — this (light) being faith — as him whose likeness (*ka-man mathaluhu*: *mathal*, 'likeness', is extra; in other words, (read)*ka-*

259Abū Jahl is a member of the ruling Quraysh of Makkah and one of early Islam's staunchest enemies. He attempted to do physical harm to the Prophet. His name means 'father of ignorance/savagery'; his real name was Amr ibn Hisham. He was killed in the Battle of Badr. Available from: [http://www.oxfordislamicstudies.com], [Accessed 25 May 2019].

man huwa, ‘as him who’) is in darkness whence he cannot emerge? — and this is the disbeliever — No! So, in the same way, that faith has been adorned for believers, what the disbelievers have been doing, in the way of disbelief and acts of disobedience, has been adorned for them.

- 7- The Qur’an makes an exhaustive reference to the resurrection and witnessing the promised day of resurrection to emphasise the truth of the message.

فَكَيْفَ إِذَا جُمِعْتَهُمْ لِيَوْمٍ لَا رَيْبَ فِيهِ وَوُفِّيَتْ كُلُّ نَفْسٍ مَّا كَسَبَتْ وَهُمْ لَا يُظْلَمُونَ ﴿٢٥﴾

Saheeh International: So how will it be when We assemble them for a **Day** about which there is no doubt? And each **soul** will be compensated (in full for) what it earned, and they will not be wronged. (3:25)

ثُمَّ إِنَّكُمْ يَوْمَ الْقِيَامَةِ تُبْعَثُونَ ﴿١٦﴾

Saheeh International: Then indeed, you, on the **Day** of Resurrection, will be resurrected. (Verse 23:16)

وَالْيَوْمِ الْمَوْعُودِ ﴿٢﴾

Saheeh International: And (by) the promised **Day**. (Verse 85:2)

c. The afterlife punishment

Punishment and Reward is the name of one of the four pivotal Qur’anic concepts in Abdul-Raof (2003, p.105). Abdul-Raof claims that monotheism التوحيد, prophethood النبوة, resurrection البعث, and reward-punishment الثواب والعقاب occur consistently. He says that they usually are found occurring consecutively and sometimes in a linear order. An example he gives of this is found below (pp. 105-107):

الَّذِي جَعَلَ لَكُمُ الْأَرْضَ فِرَاشًا وَالسَّمَاءَ بِنَاءً وَأَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجَ بِهِ مِنَ الثَّمَرَاتِ رِزْقًا لَكُمْ فَلا تَجْعَلُوا لِلَّهِ أَنْدَادًا وَأَنْتُمْ تَعْلَمُونَ ﴿٢٢﴾
 وَإِن كُنْتُمْ فِي رَيْبٍ مِّمَّا نَزَّلْنَا عَلَيَّ عَبْدَانَا فَأْتُوا بِسُورَةٍ مِّثْلِهِ ۚ وَادْعُوا شُهَدَاءَكُمْ
 فَمَنْ لَمْ تَفْعَلُوا وَلَنْ تَفْعَلُوا فَاتَّقُوا النَّارَ الَّتِي أُهِيَ وَفُودُهَا النَّاسُ وَالْحِجَارَةُ أُعِدَّتْ لِلْكَافِرِينَ
 مِنْ دُونِ اللَّهِ ۗ إِنَّ كُنْتُمْ صَادِقِينَ ﴿٢٣﴾

وَيَسِّرِ الْآيَاتِ ۗ آمَنُوا وَعَسَلُوا الصَّالِحِينَ أَن لَّهُمْ جَنَّاتٌ تَجْرِي مِنْ تَحْتِهَا الْأَنْهَارُ
 كَمَا رَفُؤُوا مِنْهَا مِنْ شَمْرٍ وَرِزْقًا قَالُوا هَذَا الَّذِي رَزَقْنَا مِنْ قَبْلُ وَأَنْتُمْ بِهِ
 مُتَسَاهِبُونَ ۗ وَهُمْ فِيهَا خَالِدُونَ ﴿٢٥﴾
 كَيْفَ نَكْفُرُونَ بِاللَّهِ وَكُنْتُمْ أَمْوَنًا فَأَحْسِبْكُمْ ثُمَّ يُمَسِّكُكُمْ ثُمَّ يُغْسِقُكُمْ
 ثُمَّ إِلَيْهِ تُرْجَعُونَ ﴿٢٤﴾

Saheeh International: (2:21) O mankind, worship your Lord, who created you and those before you, that you may become righteous –(2:22)(He) who made for you the earth a bed (spread out) and the sky a ceiling and sent down from the sky, rain and brought forth thereby fruits as provision

for you. So do not attribute to Allah equals while you know (that there is nothing similar to Him). (2:23) And if you are in doubt about what We have sent down upon Our Servant (Muhammad), then produce a Surah the like thereof and call upon your witnesses other than Allah, if you should be truthful. (2:24) But if you do not - and you will never be able to - then fear the Fire, whose fuel is men and stones, prepared for the disbelievers. (2:25) And give good tidings to those who believe and do righteous deeds that they will have gardens (in Paradise) beneath which rivers flow. Whenever they are provided with a provision of fruit therefrom, they will say, 'This is what we were provided with before.' And it is given to them in likeness. And they will have therein purified spouses, and they will abide therein eternally... (2:28) How can you disbelieve in Allah when you were lifeless and He brought you to life; then He will cause you to die, then He will bring you (back) to life, and then to Him you will be returned.

In the above verses of the Qur'an, Abdul-Raof illustrates that monotheism is presented in Verses (2:21-22), prophethood in Verse (2:23), reward and punishment in Verses (2:24-25), and resurrection in Verse (2:28). Similar to the findings of this study, Abdul-Raof states that the presentation of the pivotal Qur'anic concepts is encountered throughout the Qur'anic discourse and is a major text-linguistic feature of the Qur'an. The following are general statements which describe this discourse prosody with examples from the Qur'an (natural phenomena terms in boldface):

- 1- Afterlife punishment is linked to the resurrection and Day of Judgment.

يَوْمَ تَبْيَضُّ وُجُوهٌ وَتَسْوَدُّ وُجُوهٌ فَأَمَّا الَّذِينَ اسْوَدَّتْ وُجُوهُهُمْ أَكْفَرْتُمْ بَعْدَ إِيمَانِكُمْ فَذُوقُوا الْعَذَابَ بِمَا كُنْتُمْ تَكْفُرُونَ ﴿١٠٦﴾

Saheeh International: On the **Day**(some) faces will turn white and (some)**faces** will turn black. As for those whose faces turn black, (to them it will be said), 'Did you disbelieve after your belief? Then taste the punishment for what you used to reject.' (Verse 3:106)

أَفَمَنْ يَنْقَى وُجُوهُهُ، سُوءَ الْعَذَابِ يَوْمَ الْقِيَامَةِ وَقِيلَ لِلظَّالِمِينَ ذُوقُوا مَا كُنْتُمْ تَكْسِبُونَ ﴿١٤﴾

Saheeh International: Then is he who will shield with his face the worst of the punishment on the Day of Resurrection (like one secure from it)? And it will be said to the wrongdoers, 'Taste what you used to earn.' (Verse 39:24)

- 2- The Qur'an describes the companions of the Fire.

وَالَّذِينَ كَسَبُوا السَّيِّئَاتِ جَزَاءُ سِنِينَ يَبْتَلِهَا وَتَرْهَقُهُمْ ذِلَّةٌ مَّا لَهُمْ مِنَ اللَّهِ مِنْ عَاصِرٍ كَانَمَا أَغْشَيْتْ وُجُوهُهُمْ قِطْعًا مِنَ اللَّيْلِ مُظْلِمًا أُولَئِكَ أَصْحَابُ النَّارِ هُمْ فِيهَا خَالِدُونَ ﴿١٧﴾

Saheeh International: But they who have earned (blame for) evil doings - the recompense of an evil deed is its equivalent, and humiliation will cover them. They will have from Allah no protector. It will be as if their **faces** are covered with pieces of the night - so dark (are they). Those are the companions of the **Fire**; they will abide therein eternally (Verse 27:14)

3- The Qur'an describes the food and drink in the Fire.

وَلَا طَعَامٌ إِلَّا مِنْ غَسِيلِينَ ﴿٣٦﴾

Pickthall: Nor any **food** save filth (Verse 69:36)

إِنَّ شَجَرَتَ الزَّقُّومِ ﴿٤٣﴾ طَعَامٌ لِلْآثِمِينَ ﴿٤٤﴾

Saheeh International: Indeed, the **tree** of Zaqquq. Is **food** for the sinful. (Verses 44:43-44)

Tafsīr Ibn Kathīr : (Verily, the tree of Zaqquq will be the food of the sinners. Like boiling oil, it will boil in the bellies, like the boiling of scalding water. (It will be said:) ‘Seize him and drag him into the midst of blazing Fire, then pour over his head the torment of boiling water. Taste you (this)! Verily, you were (pretending to be) the mighty, the generous! Verily, this is that whereof you used to doubt!’) (44:43-50)

وَنَادَى أَصْحَابُ النَّارِ أَصْحَابَ الْجَنَّةِ أَنْ أَفِيضُوا عَلَيْنَا مِنَ الْمَاءِ أَوْ مِمَّا رَزَقَكُمُ اللَّهُ
قَالُوا إِنَّ اللَّهَ حَرَّمَهُمَا عَلَى الْكَافِرِينَ ﴿٥٠﴾

Pickthall: And the dwellers of the **Fire** cry out unto the dwellers of the Garden: Pour on us some **water** or some wherewith Allah hath provided you. They say: Lo! Allah hath forbidden both to disbelievers (in His guidance) (Verse 7:50)

d. The shunning of disbelievers

The shunning of disbelievers is one of the motifs that run throughout the Qur'an. The following are two general statements which describe this discourse prosody with examples from the Qur'an (natural phenomena terms in boldface):

- 1- A description of the state of mind and the hearts of disbelievers is a recurrent meaning in the Qur'an.

وَقَالُوا إِن هِيَ إِلَّا حَيَاتُنَا الدُّنْيَا وَمَا نَحْنُ بِمَبْعُوثِينَ ﴿٢٩﴾

Yusuf Ali: And they (sometimes) say: ‘There is nothing except our **life** on this **earth**, and never shall we be raised up again.’ (Verse 6:29)

إِلَهُكُمْ إِلَهٌ وَاحِدٌ فَالَّذِينَ لَا يُؤْمِنُونَ بِالْآخِرَةِ قُلُوبُهُمْ مُنْكَرَةٌ وَهُمْ مُسْتَكْبِرُونَ ﴿٢٢﴾

Saheeh International: Your god is one God. But those who do not believe in the Hereafter - their **hearts** are disapproving, and they are arrogant. (Verse 16:22)

وَجَعَلْنَا عَلَى قُلُوبِهِمْ أَكِنَّةً أَنْ يَفْقَهُوهُ وَفِي آذَانِهِمْ وَقْرًا وَإِذَا ذُكِرْتُمْ فِي الْقُرْآنِ إِحْدَاهُ
 وَلَوْ عَلَىٰ أَذْبُرِهِمْ نُفُورًا ﴿٤٦﴾

Yusuf Ali: And We put coverings over their **hearts** (and minds) lest they should understand the Qur'an, and deafness into their ears: when thou dost commemorate thy Lord and Him alone in the Qur'an, they turn on their backs, fleeing (from the Truth). (Verse 17:46)

وَإِذَا ذُكِرَ اللَّهُ وَحْدَهُ اشْمَأَزَّتْ قُلُوبُ الَّذِينَ لَا يُؤْمِنُونَ بِالْآخِرَةِ وَإِذَا ذُكِرَ
 الَّذِينَ مِنْ دُونِهِ إِذَا هُمْ يَسْتَبْشِرُونَ ﴿٤٥﴾

Saheeh International: And when Allah is mentioned alone, the **hearts** of those who do not believe in the Hereafter shrink with aversion, but when those (worshipped) other than Him are mentioned, immediately they rejoice. (Verse 39:45)

- 2- People of the previous times who denied the messages of their prophets are mentioned as negative examples.

وَقَالَ الْمَلَأُ مِنْ قَوْمِهِ الَّذِينَ كَفَرُوا وَكَذَّبُوا بِآيَاتِنَا فِي الْحَيَاةِ الدُّنْيَا مَا هَذَا
 إِلَّا بَشَرٌ مِثْلُكُمْ أَيُّهَا كُلُّ مَمَاتًا كُلُّ مَمَاتًا مِنْهُ وَيَشْرَبُ مِمَّا تَشْرَبُونَ ﴿٣٣﴾

Saheeh International: And the eminent among his **people** who disbelieved and denied the meeting of the Hereafter while We had given them luxury in the worldly **life** said, 'This is not but a man like yourselves. He eats of that from which you eat and drinks of what you drink.

كَذَّبَتْ قَبْلَهُمْ قَوْمُ نُوحٍ وَالْأَحْزَابُ مِنْ بَعْدِهِمْ وَهَمَّتْ كُلُّ أُمَّةٍ بِرَسُولِهِمْ
 لِيَأْخُذُوهُ وَجَدَلُوا بِالْبَاطِلِ لِيُدْحِضُوا بِهِ الْحَقَّ فَأَخَذْتَهُمْ فَكَيْفَ كَانَ عِقَابِ ﴿٥﴾

Saheeh International: The **people** of Noah denied before them and the (disbelieving) factions after them, and every nation intended (a plot) for their messenger to seize him, and they disputed by (using) falsehood to (attempt to) invalidate thereby the truth. So I seized them, and how (terrible) was My penalty (Verse 40:5)

e. The afterlife reward

In the interpretation of the description of the reward in the Garden of the Afterlife (Paradise), Ibn Kathir writes:

When Allah, the Exalted, mentioned the condition of the wretched, He also commended the people of delight (the believers). They are those who believe and work righteous deeds. Thus, their hearts believed, and their limbs worked righteous deeds, both in statements and actions. This includes their performance of deeds of obedience and their abandonment of evils. In this way, they are the inheritors of Gardens (of Paradise), which contain lofty rooms and seats arranged in rows. Therein they will find bunches of fruit near to them, elevated couches, fair and beautiful wives, various types of fruit, desired kinds of food and delicious drinks. They also will be allowed to see the Creator of the heavens and the earth and they will be in this state of pleasure forever. They will not die, nor will they grow old. (Ibn Kathīr in Abdul-Rahman, 2009, p.30)

The following are general statements which describe this discourse prosody with examples from the Qur'an (natural phenomena terms in boldface):

- 1- The Qur'an speaks of the promised reward in the Gardens, in an overlap with the meaning of emphasis on the message in the previous section.

﴿ وَسَارِعُوا إِلَىٰ مَغْفِرَةٍ مِّن رَّبِّكُمْ وَجَنَّةٍ عَرْضُهَا السَّمَاوَاتُ وَالْأَرْضُ أُعِدَّتْ لِلْمُتَّقِينَ ﴾

Saheeh International: And hasten to forgiveness from your Lord and a **garden** as wide as the **heavens** and **earth**, prepared for the righteous. (Verse 3:133)

- 2- The Qur'an describes in detail the bounties given to those who enter the Garden.

﴿ مَثَلُ الْجَنَّةِ الَّتِي وَعَدَ الْمُتَّقُونَ فِيهَا أَنْهَارٌ مِّن مَّاءٍ غَيْرِ آسِنٍ وَأَنْهَارٌ مِّن لَّبَنٍ لَّمْ يَتَغَيَّرَ طَعْمُهُ وَأَنْهَارٌ مِّنْ حَمْرٍ لَّدَّةٍ لِلشَّرَابِ بَيِّنٍ وَأَنْهَارٌ مِّنْ عَسَلٍ مُّصَفًّى وَلَهُمْ فِيهَا مِن كُلِّ الثَّمَرَاتِ وَمَغْفِرَةٌ مِّن رَّبِّهِمْ كَمَنْ هُوَ خَالِدٌ فِي النَّارِ وَسُقُوا مَاءً حَمِيمًا فَقَطَّعَ أَمْعَاءَهُمْ ﴾

Arberry: This is the similitude of Paradise which the godfearing have been promised: therein are **rivers** of **water** unstaling, rivers of milk unchanging in flavour, and **rivers** of wine -- a delight to the drinkers, rivers, too, of honey purified; and therein for them is every **fruit**, and forgiveness from their Lord -- Are they as he who dwells forever in the Fire, such as are given to drink boiling **water**, that tears their bowels asunder?

﴿ أُولَٰئِكَ جَزَاءُهم مَّغْفِرَةٌ مِّن رَّبِّهِمْ وَجَنَّاتٌ تَجْرِي مِن تَحْتِهَا الْأَنْهَارُ خَالِدِينَ فِيهَا وَيَعْمَلُونَ فِيهَا الْعَمَلِينَ ﴾

Pickthall: The reward of such will be forgiveness from their Lord, and Gardens underneath which rivers flow, wherein they will abide forever - a bountiful reward for workers! (Verse 3:136)

إِنَّ الَّذِينَ ءَامَنُوا وَعَمِلُوا الصَّالِحَاتِ وَأَخْبَتُوا إِلَىٰ رَبِّهِمْ أُولَٰئِكَ أَصْحَابُ الْجَنَّةِ ۖ هُمْ فِيهَا خَالِدُونَ ﴿١٣﴾

Yusuf Ali: But those who believe and work righteousness, and humble themselves before their Lord, - They will be companions of the **gardens**, to dwell therein for aye! (Verse 11:23)

﴿ مَثَلُ الْجَنَّةِ الَّتِي وَعِدَ الْمُتَّقُونَ تَجْرِي مِنْ تَحْتِهَا الْأَنْهَارُ أُكُلُهَا دَائِمٌ وَظِلُّهَا ۚ تِلْكَ عُقْبَى الَّذِينَ اتَّقَوْا وَعُقْبَى الْكَافِرِينَ النَّارُ ﴾ ﴿١٣﴾

Pickthall: A of the Garden which is promised unto those who keep their duty (to Allah): Underneath it **rivers** flow; its **food** is everlasting, and its shade; this is the reward of those who keep their duty, while the reward of disbelievers is the Fire. (Verse 13:35)

3- The Qur'an describes the companions of the Garden and their good deeds in the present life, which make them earn this reward in the Afterlife.

وَالسَّابِقُونَ السَّابِقُونَ أُولَٰئِكَ مِنْ السَّابِقِينَ وَالَّذِينَ آمَنُوا بِالْحَسَنِاتِ وَالَّذِينَ سَبَقُوا بِالْحَسَنَاتِ وَالَّذِينَ سَبَقُوا بِالْحَسَنَاتِ وَالَّذِينَ سَبَقُوا بِالْحَسَنَاتِ وَالَّذِينَ سَبَقُوا بِالْحَسَنَاتِ ۚ تِلْكَ عُقْبَى الَّذِينَ اتَّقَوْا وَعُقْبَى الْكَافِرِينَ النَّارُ ﴿١٣﴾

Saheeh International: And the first forerunners (in the faith) among the Muhajireen and the Ansar and those who followed them with good conduct - Allah is pleased with them, and they are pleased with Him, and He has prepared for them **gardens** beneath which **rivers** flow, wherein they will abide forever. That is the great triumph. (Verse 9:100)

مَنْ عَمِلَ صَالِحًا مِمَّنْ ذَكَرْنَا يُرَىٰ فِي سَمَوَاتِنَا مِنْهُ حَسَنَاتٌ ۖ أُولَٰئِكَ سَيَرْجُوهُمْ رَبُّهُم بِرَحْمَتِهِمْ ۖ وَسَيُدْخِلُهُمْ الْجَنَّاتِ الَّتِي وَعَدَ لَهُمْ ۚ وَهُمْ فِيهَا خَالِدُونَ ﴿١٧﴾

Yusuf Ali: Whoever works righteousness, **man** or **woman**, and has Faith, verily, to him will We give a new Life, a life that is good and pure and We will bestow on such their reward according to the best of their actions. (Verse 16:97)

To conclude, this commentary is evidence that SP, especially its discourse prosodies, is, in fact, a corpus linguistic feature in the language of the Qur'an; that is, the corpus-based analysis yielded up these meanings. As the reader's eyes examine the lines of the Qur'an, he/she will sense that these meanings recur and emerge, particularly in verses similar to the examples in this section. They can be considered a part of the skeleton of meanings on which the major themes of the Qur'an (e.g., nature) stand and an awareness of them can help enrich and consolidate his/her understanding of the Qur'an.

