AN ANALYTICAL STUDY ON THE HOLY QURAN BASED ON THE ORDER OF WORDS IN ARABIC AND CONJUNCTION

Rahima Bentrcia[1], Samir Zidat[2], Farhi Marir[3]

[1, 2] Department of Computer Science, Lastic Laboratory of the Systems and the information and Communication Technologies, Chahid Mostefa Ben Boulaid, University of Batna 2, Batna 05078, Algeria

[3] College of Technological Innovation, Zayed University, P.O. Box 19282, Dubai, United Arab Emirates

Email: rahmabentrcia@yahoo.com[1], samir.zidat@gmail.com[2], Farhi.marir@zu.ac.ae[3]

DOI: https://doi.org/10.22452/mjcs.vol31no1.1

ABSTRACT

Some aspects of word relations are inspired from patterns of word co-occurrences. From this point, we conduct an analytical study on one type of these patterns, which is the AND conjunctive phrases, that exist in the holy Quran. First, we propose a set of AND conjunctive patterns in order to extract the conjunctive phrases from the Quranic Arabic Corpus, which we convert to Arabic script. Then, we analyze the order of the two words that form the conjunctive phrase. We report three different cases: words that have occurred in a specific order in the conjunctive phrase and repeated only once in the Quran, words that have occurred in a specific order in the conjunctive phrase and repeated many times in the Quran, and words that have occurred in two different orders in the conjunctive phrase and repeated one/many time(s) in the holy Quran. Finally, we show that different word orders in the conjunctive phrase yield different contextual meanings and association values between the combined words.

Keywords: word co-occurrence, text mining, pattern, association relations, Quranic Arabic Corpus, Quran

1.0 INTRODUCTION

Text mining concept can be defined as "the analysis of observational textual data sets to find un-suspected relationships and to summarize the text in novel ways that are both understandable and useful to the users" [1]. Word co-occurrences are considered as one of the most powerful text mining approaches that is used to extract statistical and associational relationships from textual documents [2]. Generally, two words co-occur if they are observed together in a given unit of text. However, the unit of text can be a window of a fixed number of words, or a sentence, or a group of sentences that may form a small paragraph or a document. Moreover, different text mining methods have been developed and applied in different fields such as information retrieval, which is widely used to answer queries like the case in search engines [3]. In addition, various ontology-based information extraction systems are also based on such methods either to extract keywords from a specific domain or to find the relationships among them [4].

Text mining and natural language processing methods are highly cooperated to extract information from text where such information is presented in an unstructured format that is not immediately suitable for automatic analysis by a computer. Applying text mining techniques, supported by machine learning methods, can play a significant role to extract useful information which provides potential benefits for a lot of applications such as text categorization, concept/entity extraction, and entity relation modeling.

On the other hand, researchers are also starting to exploit the text mining approaches to extract knowledge from sacred texts such as the holy Quran and the Bible in order to get better understanding of the Islamic and Christian religions [5]. Nevertheless, there is a lack in these approaches that deal with texts written in Arabic script due, for example, to the nature of Arabic writing, the semantic ambiguity of words, and the shortage in resources and tools that support Arabic [6]. For Quran mining, previous studies aim to achieve the purpose of understanding Quran as a source of knowledge and extracting useful information automatically. Therefore, Quran can be presented to the world and exploited very efficiently in many scientific, linguistic, and religious applications. Although few studies have been conducted in the literature on Arabic text mining [7, 8, 9], only very few mined Quranic Arabic text. Currently, the existing approaches to mine Quran are divided into computational and statistical methods, where statistics are used to extract information from Quran such as word co-occurrence, Quran concordance, and verses similarity [10, 11], and morphological and syntactical methods where Quran is analyzed to extract lexical and semantic information, or to construct a knowledge representation model such as ontologies and treebanks [12, 13].

In this paper, we conduct an analytical study that aims at mining the Arabic text of the holy Quran. To the best of our knowledge, there is no research study that analyzed this sacred text the way it is done in this paper. The main contribution of this paper is that, we combine statistical and grammatical methods to mine Quran. First, we exploit an efficient Arabic tool, which is AND conjunction, to extract the co-occurred words that are combined by AND conjunction and hence represent conjunctive phrases. Second, we propose a set of patterns that are used to extract the whole set of co-occurred words combined by AND. Moreover, we demonstrate various cases of the words that take different positions/orders in the conjunctive phrase. In particular, we show that different orders of one word yield different meanings and association measures. This study presents a novel approach since none of the existing methods illustrated the order concept of co-occurred words or even provided statistics about the different positions/ orders that co-occurred words had taken in Quran. Finally, we measure the value of the association relationship between the two co-occurred words in the conjunctive phrase using Pointwise Mutual Information method (PMI) and the Sketch Engine tool function (Word Sketch Difference). These basic analyses can be exploited very efficiently to build Quranic ontologies by extracting semantic relations from the holy Quran and assigning precise properties and restrictions to them [14].

The rest of the paper is organized as follows: Section 2 presents the previous work. Corpus preprocessing is introduced in Section 3 and conjunctive patterns extraction is illustrated in Section 4. Section 5 discusses the analysis of word order in the conjunctive phrase. Finally, a conclusion is presented in Section 6.

2.0 PREVIOUS WORK

The concept of text mining is becoming increasingly popular. Therefore, many studies are carried out to show the different methods used to analyze and extract knowledge from textual data. A study was conducted by Momtazi *et al.* where they proposed a term clustering algorithm to retrieve sentences from a corpus [15]. This algorithm is based on assigning similar terms to the same clusters based on their tendency to co-occur in similar contexts. Also, they compared four different methods for estimating word co-occurrence frequencies from two different corpora and discussed their effects on the system performance. In addition, Islam and Inkpen proposed a corpus-based method for calculating the semantic similarity of pair of words [16]. They used Point-wise Mutual Information (PMI) to measure the common words in the context of the two target words and exploit these PMI values to calculate the relative semantic similarity. The results were evaluated using four different corpora. Furthermore, Gomaa and Fahmy discussed three different methods of text similarity: String-based, Corpus-based, and Knowledge-based similarities [17]. A hybrid of these approaches was presented and useful similarity packages were mentioned.

For Arabic text, little has been written about text mining due to the nature of Arabic script [18]. One work was conducted by Alrabiah *et. al* where they performed two empirical studies by applying a number of probabilistic distributional semantic models to automatically identify lexical collocations [19]. They tested the performance of eight different association measures on the holy Quran in the first study, and they constructed a Classical Arabic corpus to be used in the second study. Their experiments showed that MI.log_freq association measure achieved the best results in extracting the collocations whereas mutual information association measure achieved the worst results. Another approach was presented by Attia *et. al* to design and implement an Arabic lexical semantics Language Resource (LR) that enables the retrieval of the possible senses of any given Arabic word at a high coverage [20]. Instead of tying full Arabic words to their possible senses, they related morphologically and POS-tags constrained Arabic lexical compounds to a predefined limited set of semantic fields across which the standard semantic relations are defined and hence the possible senses of the desired Arabic word are retrieved.

A different method was introduced by Thabtah *et. al* to classify Arabic documents, specifically the published Corpus of Contemporary Arabic (CCA), using four classification learning algorithms: Decision trees (C4.5), Hybrid (PART), Rule Induction (RIPPER), and Simple Rule (OneRule) [21]. They used WEKA, the open business intelligence tool, to evaluate the performance of these algorithms and they found that C4.5 is the most applicable algorithm to Arabic text classification in terms of error-rate, precision, and recall. Moreover, Al-Yahya *et al*. developed Badea system in order to enrich the ontological lexicon of Arabic language [22]. Badea was built semi-automatically to extract lexical relations specifically antonyms using a pattern-based approach. The method used an ontology of "seed" pairs of antonyms to facilitate the extraction of lexico-syntactic patterns in which the pairs occur. These patterns are then used to find new antonym pairs in a set of Arabic language corpora. The results showed important findings on the reliability of patterns in extracting antonyms for Arabic.

On the other side, Quran mining occupies a large area in text mining although very few approaches have been developed for Quranic Arabic due to the depth of knowledge needed in this field and the challenges related to Arabic script. A research study was conducted by Safeena and Kammani to review Qur'anic computation methods in term of research and application [23]. The work surveyed the development of Quranic computation using a literature review and classification of journal articles, conference proceedings and dissertations from 1997 to 2011. This study also covered general Arabic besides Quranic Arabic and helped to facilitate the understanding of Quranic text. Hamam *et al.* created an illustrative graphic-based tool which helps Quran experts to easily mine Quran [24]. This platform not only links one chapter to another chapter, or one verse to another verse through words, but also connects chapters and verses together through concepts and dependencies. Also, it provides expert users the ability to add new aspects and their dependencies to a shared database.

A different approach was presented by Al-Kabi *et al.* to classify the verses of Al-Fatiha and Yaseen chapters automatically [11]. The classifier normalizes the verses in the first step then applies the score function to categorize each verse to the class for which it has the highest score value. The accuracy rate reached 91% although it can be improved using a full corpus of the holy Quran and a better stemmer. Furthermore, Siddiqui *et al*. proposed a Probabilistic Topic Model method to discover the thematic structure of the holy Quran [25]. First, they applied a number of preprocessing steps to the Arabic Quranic chapters (Surahs) in order to obtain the final set of features from the raw text in those documents. Then, they used the Latent Dirichlet Allocation (LDA) algorithm which was run with different values of the input parameters to identify topics at different levels of granularity. Finally, the topics contained in each surah along with the most important terms that defined those topics were extracted.

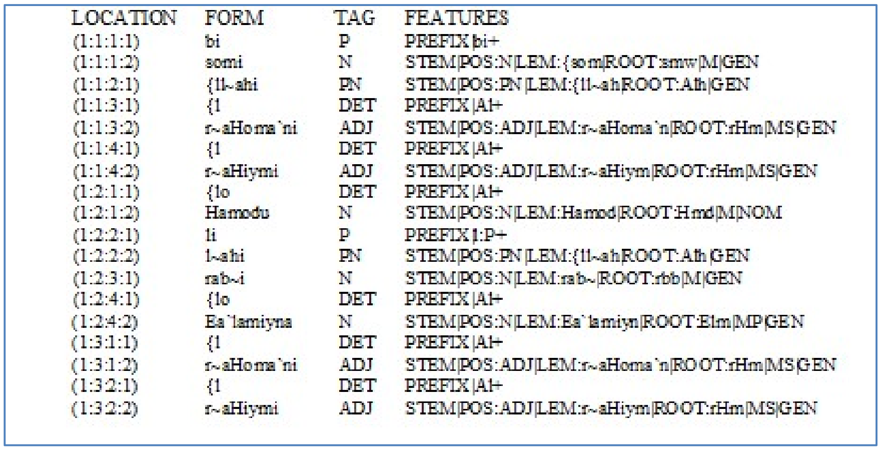
In addition, Sharaf and Atwell presented QurAna, a large corpus created from the holy Quran, and more than 24000 pronouns were tagged with their antecedence information [26]. These antecedents were maintained as an ontological list of concepts which improves information systems performance. Finally, some useful applications that can exploit this corpus were mentioned. The same authors proposed a different corpus QurSim where semantically similar or related verses of the Arabic Quran were linked together [27]. A total set of 7600 pairs of related verses were included in the corpus with different relatedness degree. Moreover, the authors provided an online query page to demonstrate, for a given verse, a network of all direct and indirect related verses. Some useful applications of this corpus were also mentioned. Abbas exploited an existing index of Quranic topics from a scholarly source: Tafsir of Ibn Kathir, to develop Qurani which is a search tool that looks for concepts in the holy Quran and provides English translations for the verses containing these concepts [28].

An efficient framework for modelling and retrieving knowledge from different sources primarily related to the holy Quran and scholarly texts was developed by Ul Ain and Basharat [29]. The documents were annotated using the domain ontology and the system employed semantic web, information extraction, and natural language processing techniques so users can query that filtered and concise knowledge using a semantic based intelligent search engine. The Quranic Arabic Corpus (QAC) is another linguistic and religious resource which was initiated by Dukes *et al*. in order to enable further analysis of the Quran [30]. The authors relied on the Arabic traditional grammar to provide multiple layers of Quran annotation including part-of-speech tagging, morphological segmentation, and syntactic analysis. Besides that, they presented a new online supervised collaboration approach to linguistic annotation of Quranic Arabic which passes through automatic rule-based tagging, initial manual verification, and online supervised collaborative proofreading to ensure a high quality resource.

In the proposed work, the framework for conducting the analytical study on Quran was built in three distinct phases that include corpus preprocessing, conjunctive patterns extraction, and analyzing the order of words in the conjunctive phrase. The following sections discuss these phases in detail.

3.0 CORPUS PREPROCESSING

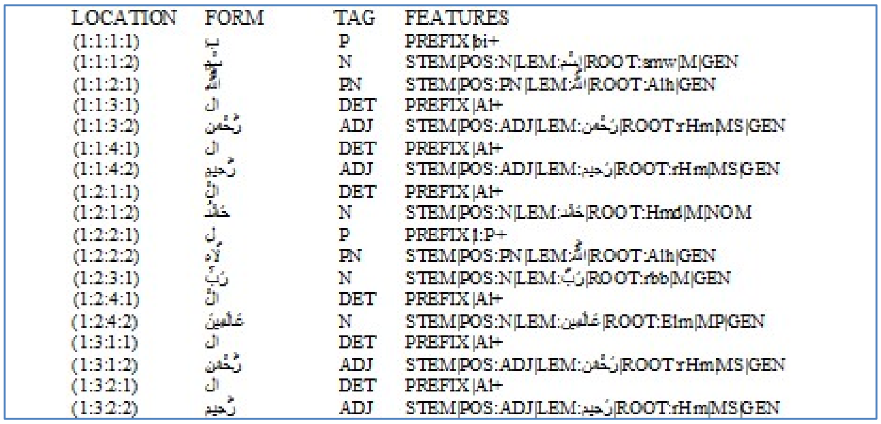
Quranic Arabic Corpus (Quranic Arabic Corpus) is an integrated and reliable linguistic resource developed by Kais Dukes in Leeds University. The corpus provides three levels of analysis: morphological annotation, a syntactic treebank, and a semantic ontology. This annotated linguistic resource consists of 77430 words of Quranic Arabic, divided into 114 documents. Each word is tagged with its part-of-speech as well as multiple morphological features that are based on the traditional Arabic grammar. Also, it is stored as a text file and is available for free. The corpus is written in Buckwalter Arabic transliteration (Buckwalter code) as shown in Fig. 1, which displays the first three verses of Al-Fatihah (The Opener) chapter.



LOCATION	FORM	TAG	FEATURES
(1:1:1:1)	bi	P	PREFIX|bi+
(1:1:1:2)	somi	N	STEM|POS:N|LEM:{som|ROOT:smw|M|GEN
(1:1:2:1)	{ll~ahi	PN	STEM|POS:PN|LEM:{ll~ah|ROOT:Alh|GEN
(1:1:3:1)	{l	DET	PREFIX|Al+
(1:1:3:2)	r~aHoma`ni	ADJ	STEM|POS:ADJ|LEM:r~aHoma`n|ROOT:rHm|MS|GEN
(1:1:4:1)	{l	DET	PREFIX|Al+
(1:1:4:2)	r~aHiymi	ADJ	STEM|POS:ADJ|LEM:r~aHiym|ROOT:rHm|MS|GEN
(1:2:1:1)	{lo	DET	PREFIX|Al+
(1:2:1:2)	Hamodu	N	STEM|POS:N|LEM:Hamod|ROOT:Hmd|M|NOM
(1:2:2:1)	li	P	PREFIX|l:P+
(1:2:2:2)	l~ahi	PN	STEM|POS:PN|LEM:{ll~ah|ROOT:Alh|GEN
(1:2:3:1)	rab~i	N	STEM|POS:N|LEM:rab~|ROOT:rbb|M|GEN
(1:2:4:1)	{lo	DET	PREFIX|Al+
(1:2:4:2)	Ea`lamiyna	N	STEM|POS:N|LEM:Ea`lamiyn|ROOT:Elm|MP|GEN
(1:3:1:1)	{l	DET	PREFIX|Al+
(1:3:1:2)	r~aHoma`ni	ADJ	STEM|POS:ADJ|LEM:r~aHoma`n|ROOT:rHm|MS|GEN
(1:3:2:1)	{l	DET	PREFIX|Al+
(1:3:2:2)	r~aHiymi	ADJ	STEM|POS:ADJ|LEM:r~aHiym|ROOT:rHm|MS|GEN

Fig. 1: Sample of the Quranic Arabic Corpus in Buckwalter transliteration

There are four columns in the corpus; the LOCATION column consists of four numbers which illustrate the chapter number in Quran, the verse number in this chapter, the word number in this verse, and the part number in this word. The FORM column divides each word into its main parts, whereas the TAG column assigns for each part in the previous column its part of speech (POS) tag such as noun, determinant, verb, etc. Finally, the FEATURE column describes the morphological structure of each part in the word such as prefix, stem, suffix, etc.



LOCATION	FORM	TAG	FEATURES
(1:1:1:1)	بِ	P	PREFIX|bi+
(1:1:1:2)	سْمِ	N	STEM|POS:N|LEM:ٱسْم|ROOT:smw|M|GEN
(1:1:2:1)	ٱللَّهِ	PN	STEM|POS:PN|LEM:ٱللَّه|ROOT:Alh|GEN
(1:1:3:1)	ٱل	DET	PREFIX|Al+
(1:1:3:2)	رَّحْمَٰنِ	ADJ	STEM|POS:ADJ|LEM:رَحْمَٰن|ROOT rHm|MS|GEN
(1:1:4:1)	ٱل	DET	PREFIX|Al+
(1:1:4:2)	رَّحِيمِ	ADJ	STEM|POS:ADJ|LEM:رَحِيم|ROOT:rHm|MS|GEN
(1:2:1:1)	ٱلْ	DET	PREFIX|Al+
(1:2:1:2)	حَمْدُ	N	STEM|POS:N|LEM:حَمْد|ROOT:Hmd|M|NOM
(1:2:2:1)	لِ	P	PREFIX|l:P+
(1:2:2:2)	لَّهِ	PN	STEM|POS:PN|LEM:ٱللَّه|ROOT:Alh|GEN
(1:2:3:1)	رَبِّ	N	STEM|POS:N|LEM:رَبّ|ROOT:rbb|M|GEN
(1:2:4:1)	ٱلْ	DET	PREFIX|Al+
(1:2:4:2)	عَٰلَمِينَ	N	STEM|POS:N|LEM:عَٰلَمِين|ROOT:Elm|MP|GEN
(1:3:1:1)	ٱل	DET	PREFIX|Al+
(1:3:1:2)	رَّحْمَٰنِ	ADJ	STEM|POS:ADJ|LEM:رَحْمَٰن|ROOT rHm|MS|GEN
(1:3:2:1)	ٱل	DET	PREFIX|Al+
(1:3:2:2)	رَّحِيمِ	ADJ	STEM|POS:ADJ|LEM:رَحِيم|ROOT:rHm|MS|GEN

Fig. 2: Sample of the Quranic Arabic Corpus converted from Buckwalter transliteration to Arabic.

The main objective of the preprocessing step is to facilitate the understanding and hence the use of the corpus. This was accomplished by converting the available Quranic corpus to Arabic version. We develop a conversion method to transfer back each character from Buckwalter code to its equivalent Arabic character. These include

Arabic alphabet and diacritics. Fig. 2 demonstrates a sample of the Quranic Arabic corpus converted to Arabic script.

4.0 CONJUNCTIVE PATTERNS EXTRACTION

Arabic is a very challenging language due to its morphological structure and the richness of its grammatical rules. In this work, we exploit one common grammatical rule, which is AND conjunction, to define a set of patterns that helps in extracting co-occurred words. The patterns consist of two words, which could be nouns, adjectives, or proper nouns, enclosing AND conjunction in between. The AND conjunctive rule states that the two combined words must share a kind of association between each other.

There are nine conjunction tools in Arabic. However, only six of them have a conjunctive role in the holy Quran, and have been repeated for several times [31], as shown in Table 1.

Table 1: The Arabic conjunction tools mentioned in the holy Quran

Conjunctive	ثم THEN	أو OR	و AND	بل BUT	الفاء THEN	أم OR
Frequency	330	328	234	144	41	29

Based on a deep study of Arabic grammar [32, 33], POS tagging, and morphology features found in the Quranic Arabic corpus, we treat only the cases where the two combined words are nouns, proper nouns, and adjectives. Other complex cases are beyond the scope of this work because they need specific knowledge resources such as exegesis of the holy Quran. This set is as follows:

1. Noun + "AND" Conjunction + Noun: this pattern is for extracting any two nouns with AND conjunction in between, such as: "رَعْد و بَرْق", which means thunder and lightning. Different cases of this pattern are explained bellow:
   a. Noun + "AND" Conjunction + Noun + Determinant 'ال' + Noun: the two combined nouns are followed by a third noun which starts with a determinant, like "الرَّسُولَ و أُولِي الأَمْرِ", which means messenger and those in authority among you.
   b. Noun + "AND" Conjunction + Noun + Noun: the two combined nouns are followed by a third noun, like "أَعْمَى و أَضَلُّ سَبِيلًا", which means blind and will be further astray from the path.
   c. Noun + "AND" Conjunction + Noun + Determinant 'ال' + Adjective: the two combined nouns are followed by an adjective which starts with a determinant, like "الآيَاتِ و الذِّكْرِ الْحَكِيم", which means the verses and the wise remembrance.
   d. Noun + "AND" Conjunction + Noun + Adjective: the two combined nouns are followed by an adjective like "غُصَّةٍ و عَذَاباً أَلِيمًا", which means choking food and a painful punishment.
2. Adjective + "AND" Conjunction + Adjective: this rule is for extracting any two adjectives with AND conjunction in between, such as "شَقِيٌّ و سَعِيد", which means the wretched and the prosperous.
3. Proper Noun + "AND" Conjunction + Determinant 'ال' + Proper Noun: this rule is for extracting any two Proper nouns with AND conjunction in between, and the second one starts with a determinant, such as "اسْمَاعِيلَ و الْيَسَعَ", which means Ishmael and Elisha.
4. Proper Noun + "AND" Conjunction + Proper Noun: this rule is for extracting any two Proper nouns with AND conjunction in between, for example: "يَأْجُوج و مَأْجُوج", which means Gog and Magog.
5. Proper Noun + "AND" Conjunction + Determinant 'ال' + Noun: this rule extracts any Proper noun followed by a noun which starts with a determinant, for example: "نُوحٍ و النَّبِيِّينَ", which means Noah and the prophets.
6. Noun + Pronoun + "AND" Conjunction + Noun + Pronoun: this rule extracts any two nouns combined with AND, and the first noun ends with a Pronoun, for example: "سِرَّهُم و نَجْوَاهُم", which means their secrets and their private conversations.

Moreover, we define a set of negation conjunctive patterns, where the negation letter ‘لا, NOT’ is used with the "AND" conjunction, as clarified next.

7. Negation ‘NOT لا’ + Adjective + "AND" Conjunction + Negation ‘NOT لا’ + Adjective: this pattern finds out any two negative adjectives combined with AND conjunction, such as “لاَ مَقْطُوعَة وَ لاَ مَمْنُوعَة”, which means neither limited nor forbidden.
8. Negation ‘NOT لا’ + Determinant ’ال’ + Noun + "AND" Conjunction + Negation ‘NOT لا’ + Determinant ’ال’ + Noun: this pattern finds out any two negative nouns combined with AND conjunction, such as “ لاَ الظُّلُمَات وَ لاَ النُّور”, which means neither the darknesses nor the light.
9. Adjective + "AND" Conjunction + Negation ‘NOT لا’ + Adjective: this pattern finds out any two adjectives combined with AND, where the second one is directly preceded by a negation, such as “ كَاهِن و لاَ مَجْنُون”, which means not a soothsayer or a madman.
10. Noun + "AND" Conjunction + Negation ‘NOT لا’ + Noun: this pattern finds out any two nouns combined with AND, where the second one is directly preceded by a negation, such as “صَاحِبَة وَ لاَ وَلَدًا”, which means not a wife or a son.

5.0 ANALYZING THE ORDER OF WORDS IN THE CONJUNCTIVE PHRASE

The holy Quran is the last heavenly books that God revealed to the Prophet Muhammad, peace be upon him. It is divided into 114 chapters called Surah, of different size, and each chapter consists of several verses named Ayah, which in total, make 6243 verses and 77430 words [12].

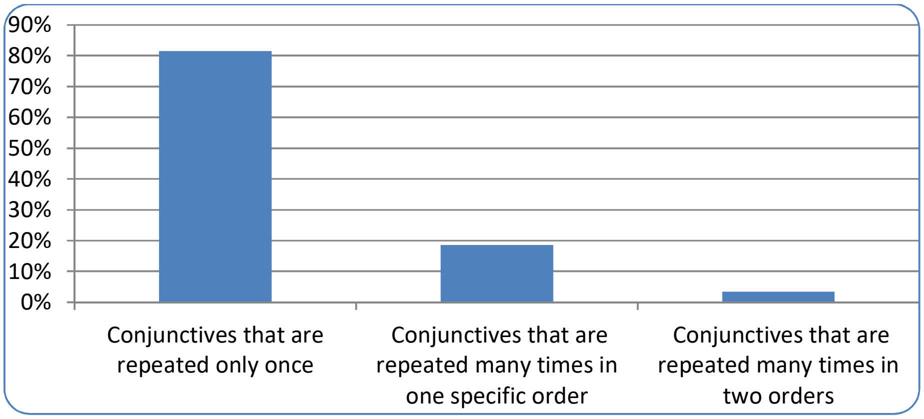


Fig. 3: The three categories of word orders and their percentages

The Quranic text is very challenging to be studied because it is the word of God. Therefore, every word in the Quran counts a great deal and needs a solid knowledge of Arabic in general and the language of the holy Quran in particular. We have tested this fact during the conducting of this work where we found that each word in Quran reserves a specific position in the verse because of important reasons related to the interpretation of that verse [34]. More accurately, a word may precede an adjacent word because of a special care, the more care you pay for a word in Quran, the more precedence among words it has in the verse. For this reason, we face some words which precede adjacent words in many verses whereas they follow them in others. In the case of conjunctive phrases, we can divide the two combined words based on their position/order in the conjunctive phrase into three main categories:

- Words that occurred in a specific order in the conjunctive phrase and repeated only one time in Quran. It occupies a high percentage of 81.47% of the total number of AND conjunctive phrases.
- Words that occurred in a specific order in the conjunctive phrase and repeated many times in Quran. It occupies a reasonable percentage of 18.62% of the total number of AND conjunctive phrases.

- Words that occurred in two different orders in the conjunctive phrase and repeated one/ many time(s) in the holy Quran. It occupies a small percentage of 3.43% of the total number of AND conjunctive phrases.

The three categories and their percentages are shown in Fig. 3.

5.1 Words that have occurred in one specific order in the conjunctive phrase and repeated only once in the Quran

This set includes words that are combined together with AND conjunction and occurred together in that order only once in the holy Quran even if they are repeated many times separately. As elements of this set, we can find conjunctive phrases of proper nouns and nouns, as shown in Table 2.

Table 2: Sample of conjunctive phrases that occurred once in one specific order

TERM2		AND	TERM1		TYPE
Solomon	سُلَيْمَان	وَ	هَارُون	Aaron	Proper Nouns
Uzza	ٱلْعُزَّى	وَ	ٱللَّات	Lat	Proper Nouns
Cucumbers	قِثَّائ	وَ	بَقْل	Green Herbs	Proper Nouns
Summer	صَيْف	وَ	شِتَاء	Winter	Proper Nouns
al-Marwah	مَرْوَة	وَ	صَفَا	as-Safa	Proper Nouns
Sheep	غَنَم	وَ	بَقَر	Cow	Proper Nouns
Morning Star	طَارِق	وَ	سَمَاء	Heaven	Nouns
No disputing	لَا جِدَالَ	وَ	لَا فُسُوقَ	No disobedience	Nouns
Diver	غَوَّاص	وَ	بَنَّاء	Builder	Nouns
Weaning	فِصَال	وَ	حَمْل	Gestation	Nouns
Bowls	جِفَان	وَ	تَمَاثِيل	Statues	Nouns
Nor sleep	لَا نَوْمٌ	وَ	سِنَةٌ	Drowsiness	Nouns

5.2 Words that have occurred in one specific order in the conjunctive phrase and repeated many times in Quran

There are many Arabic and Islamic studies that talk about order in Quranic co-occurred words, and explain the reasons that make a word precedes or follows an adjacent word in the verse [35, 36]. In the case of conjunctive phrases, we find a set of words that follow the same order many times in the Quran. This repetition could be considered as a sign for the existence of a relationship between these words. Table 3 illustrates some elements of this set.

Islamic scholars indicate many reasons for words precedence. One of them is the word preference. We find this, for example, in the phrase "**الذَّكَرَ وَالأُنْثَى**", "**Male AND Female**", in the verse 45 of An-Najm (The Star) chapter:

وَأَنَّهُ خَلَقَ الزَّوْجَيْنِ الذَّكَرَ وَالأُنْثَى

(*And that He creates the two mates - the male and female* -) [53:45]

The word male always precedes the word female because male exhibits some distinct features that female do not i.e. physical capabilities that make him stronger and more capable of performing some tasks that female cannot.

Another reason is word precedence in the sense of existence such as in the phrase "**إِسْحَاقَ وَيَعْقُوبَ**", "**Isaak AND Jacob**" where the prophet Isaak was born before his brother the prophet Jacob and the prophet Ishmael was born before his brother Isaak, as shown in the verse 84 of Al-An'am (The Cattle) chapter:

وَوَهَبْنَا لَهُ إِسْحَاقَ وَيَعْقُوبَ كُلًّا هَدَيْنَا وَنُوحًا هَدَيْنَا مِن قَبْلُ وَمِن ذُرِّيَّتِهِ دَاوُودَ وَسُلَيْمَانَ وَأَيُّوبَ وَيُوسُفَ وَمُوسَىٰ
وَهَارُونَ وَكَذَٰلِكَ نَجْزِي الْمُحْسِنِينَ

(*And We gave to Abraham, Isaac and Jacob - all [of them] We guided. And Noah, We guided before; and among his descendants, David and Solomon and Job and Joseph and Moses and Aaron. Thus do We reward the doers of good*) [6:84]

Table 3: Sample of conjunctive phrases that occurred many times in one specific order

The Conjunctive Phrase		**Frequency in Quran**
'Judgment AND Knowledge'	'حُكْم 'و ' عِلْم '	4
' East AND west'	'مَشْرِق 'و ' مَغْرِب'	6
'Unseen AND the Witnessed'	'غَيْب 'و ' شَهَادَة '	10
'Guidance AND Mercy'	'هُدًى 'و ' رَحْمَة '	13
' Isaac AND Jacob'	'إِسْحَاق 'و ' يَعْقُوب'	10
'World AND Hereafter'	'دُّنْيَا 'و ' اخِر '	16
' Male AND Female '	'ذَكَر 'و ' أُنْثَى'	4
' Night AND Day '	' لَيْل 'و' نَهَار'	21
'Protector NOR Helper'	' وَلِي ' و لَا 'نَصِير'	12
' Ishmael AND Isaac' '	إِسْمَاعِيل' و ' إِسْحَاق '	6
'Forgiveness AND Reward'	' مَّغْفِرَة ' و 'أَجْر '	6

Also, it appears clearly in the verse 39 of Ibrahim (Abraham) chapter:

الْحَمْدُ لِلَّهِ الَّذِي وَهَبَ لِي عَلَى الْكِبَرِ إِسْمَاعِيلَ وَإِسْحَاقَ إِنَّ رَبِّي لَسَمِيعُ الدُّعَاءِ

(*Praise to Allah, who has granted to me in old age Ishmael and Isaac. Indeed, my Lord is the Hearer of supplication*) [14:39]

A different reason is word precedence in the sense of time such as in the phrase "**الْمَشْرِقُ وَالْمَغْرِبُ**", "**East AND West**", where the day starts by the sunrise from east to west, as mentioned bellow in the verse 115 of Al-Baqarah (The Cow) chapter:

وَلِلّهِ الْمَشْرِقُ وَالْمَغْرِبُ فَأَيْنَمَا تُوَلُّواْ فَثَمَّ وَجْهُ اللّهِ إِنَّ اللّهَ وَاسِعٌ عَلِيمٌ

(*And to Allah belongs the east and the west. So wherever you [might] turn, there is the Face of Allah. Indeed, Allah is all-Encompassing and knowing*) [2:115]

In addition, we find word precedence according to the development situation such as "**السَّمْعَ وَالْبَصَرَ**", "**Hearing AND Vision**" in the fetus where the evolution of hearing is completed before the evolution of vision which is delayed after the birth of the fetus. The verse 78 of An-Nahl (The Bees) chapter states this clearly:

وَاللَّهُ أَخْرَجَكُمْ مِنْ بُطُونِ أُمَّهَاتِكُمْ لَا تَعْلَمُونَ شَيْئًا وَجَعَلَ لَكُمُ السَّمْعَ وَالْأَبْصَارَ وَالْأَفْئِدَةَ لَعَلَّكُمْ تَشْكُرُونَ

(*And Allah has extracted you from the wombs of your mothers not knowing a thing, and He made for you hearing and vision and intellect that perhaps you would be grateful*) [16:78]

5.3 Words that have occurred in two different orders in the conjunctive phrase and repeated one/many time(s) in the holy Quran

One main application of word co-occurrences is to extract semantic relations that may exist between them [2]. In Arabic grammar, the association relation between two words in AND conjunctive phrase word_1 and word_2 is the same as the relation between word_2 and word_1, which is not the case in Quranic conjunctive phrases. Our contribution in this study is to reveal and discuss the differences between the two types of association that may exist between word_1 and word_2, and word_2 and word_1 in the Quranic conjunctive phrases from the contextual meaning side and the association magnitude side.

There are no extra or meaningless words in the Quran; on the contrary, there exist words which have more than one meaning based on their positions in the verse. Moreover, the order which a word follows in a verse may also influence its interpretation. In the case of conjunctive phrases, we find a set of words that follow two different orders one/ many time(s) in the Quran such as the examples of Table 4.

Whether a specific word precedes or follows its adjacent word is based on the context of the verse where they occur [37]. For example, in the phrase "**الْأَرْضَ وَالسَّمَاوَاتِ**", "**Earth AND Heavens**", the word 'Earth' precedes the word 'Heavens' because earth is created before heavens, as illustrated in the verse 4 of Taha (Ta-Ha) chapter:

تَنزِيلًا مِّمَّنْ خَلَقَ الْأَرْضَ وَالسَّمَاوَاتِ الْعُلَى

(*A revelation from He who created the earth and highest heavens*) [20:4]

However, in more than 100 verses, we find the word 'heavens' comes before 'earth' because of its huge space and great creation. An example is the verse 77 of An-Nahl (The Bees) chapter:

وَلِلَّهِ غَيْبُ السَّمَاوَاتِ وَالْأَرْضِ وَمَا أَمْرُ السَّاعَةِ إِلَّا كَلَمْحِ الْبَصَرِ أَوْ هُوَ أَقْرَبُ إِنَّ اللَّهَ عَلَىٰ كُلِّ شَيْءٍ قَدِيرٌ

(*And to Allah belongs the unseen [aspects] of the heavens and the earth. And the command for the Hour is not but as a glance of the eye or even nearer. Indeed, Allah is over all things competent*) [16:77]

Another example of words which have occurred in two different orders is the phrase "**الْجِنَّ وَالْإِنسَ**" "**Jinn AND Mankind**" in the verse 56 of Adh-Dhariyat (The Winnowing Winds) chapter, we found that the word 'Jinn' precedes 'Mankind' because Jinn are created before Mankind.

وَمَا خَلَقْتُ الْجِنَّ وَالْإِنسَ إِلَّا لِيَعْبُدُونِ

(*And I did not create the jinn and mankind except to worship Me*) [51:56]

Moreover, in the verses where there is a kind of challenging in movement and speed, we also find Jinn before Men because of their supernatural ability, as presented in the verse 33 of Ar-Rahman (The Beneficent) chapter:

يَا مَعْشَرَ الْجِنِّ وَالْإِنسِ إِنِ اسْتَطَعْتُمْ أَن تَنفُذُوا مِنْ أَقْطَارِ السَّمَاوَاتِ وَالْأَرْضِ فَانفُذُوا لَا تَنفُذُونَ إِلَّا بِسُلْطَانٍ

(*O company of jinn and mankind, if you are able to pass beyond the regions of the heavens and the earth, then pass. You will not pass except by authority [from Allah*]) [55:33]

Table 4: Sample of conjunctive phrases that occurred one/ many time(s) in Quran in two orders

The Conjunctive Phrase		**Frequency**	**(PMI) Method**	**Word Sketch Difference Function**
'Heavens AND Earth'	'سَمَاء ' و'أَرْض '	2	1.1827	5.4
' Earth AND Heavens '	'أَرْض ' و 'سَمَاء '	148	5.0673	10.2
' Thamud AND ‘Ad '	'ثَمُود ' و 'عَاد '	5	7.8071	9.0
' ‘Ad AND Thamud'	'عَاد ' و 'ثَمُود '	1	5.9997	6.7
' Warner AND 'Bearer of glad tidings '	'نَذِير ' و 'بَشِير '	5	8.1641	10.0
'Bearer of glad tidings AND Warner'	'بَشِير ' و 'نَذِير '	2	7.1641	9.8
'Jinn AND Mankind'	'جِنّ ' و 'إِنس '	3	7.5626	9.1
'Mankind AND Jinn'	'إِنس ' و 'جِنّ '	9	9.2996	10.8
'Harm NOR Benefit'	'ضَر' وَلَا 'نَفْع '	3	9.4106	9.9
'Benefit NOR Harm'	' نَفْع' وَلَا ' ضَر '	4	10.1476	10.4

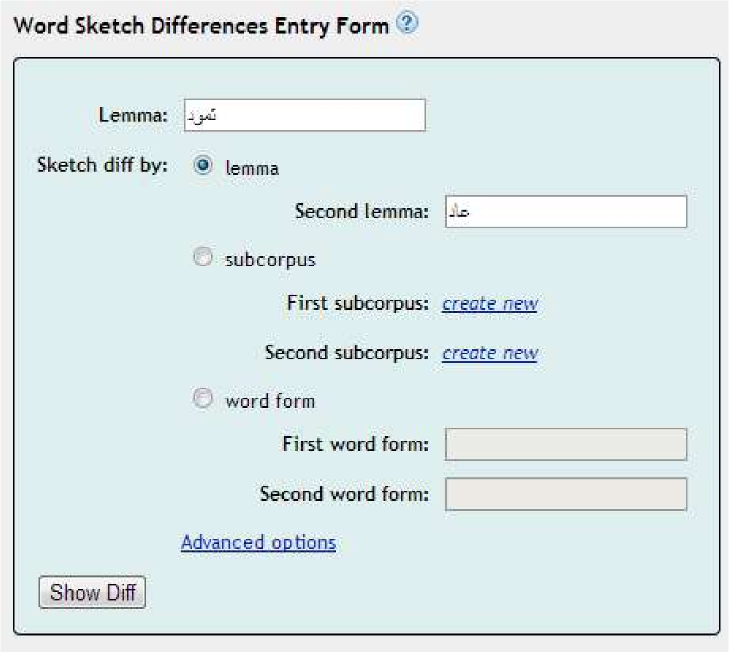
$$PMI(x,y) = \log \frac{p(x,y)}{p(x)p(y)} \quad (1)$$

However, in some verses, such as the verse 88 of Al-Israa (The Night Journey) chapter, God asked Men before Jinn to create Quran because it is a challenge for them first and foremost:

قُل لَّئِنِ اجْتَمَعَتِ الْإِنسُ وَالْجِنُّ عَلَىٰ أَن يَأْتُوا بِمِثْلِ هَٰذَا الْقُرْآنِ لَا يَأْتُونَ بِمِثْلِهِ وَلَوْ كَانَ بَعْضُهُمْ لِبَعْضٍ ظَهِيرًا

(*Say, "If mankind and the jinn gathered in order to produce the like of this Qur'an, they could not produce the like of it, even if they were to each other assistants.*") [17:88]

On the other side, in order to find the difference in the association values between the two words in the conjunctive phrase, we apply Pointwise Mutual Information method (PMI) to measure how much information one word can give about the other one which occurs with it [38]. This method is derived from information theory and widely proposed to find semantic relations between either adjacent words that occur together frequently or trigger pairs, which are long distance word pairs.



Word Sketch Differences Entry Form

Lemma: ثمود

Sketch diff by: ● lemma

Second lemma: عاد

○ subcorpus

First subcorpus: create new

Second subcorpus: create new

○ word form

First word form:

Second word form:

Advanced options

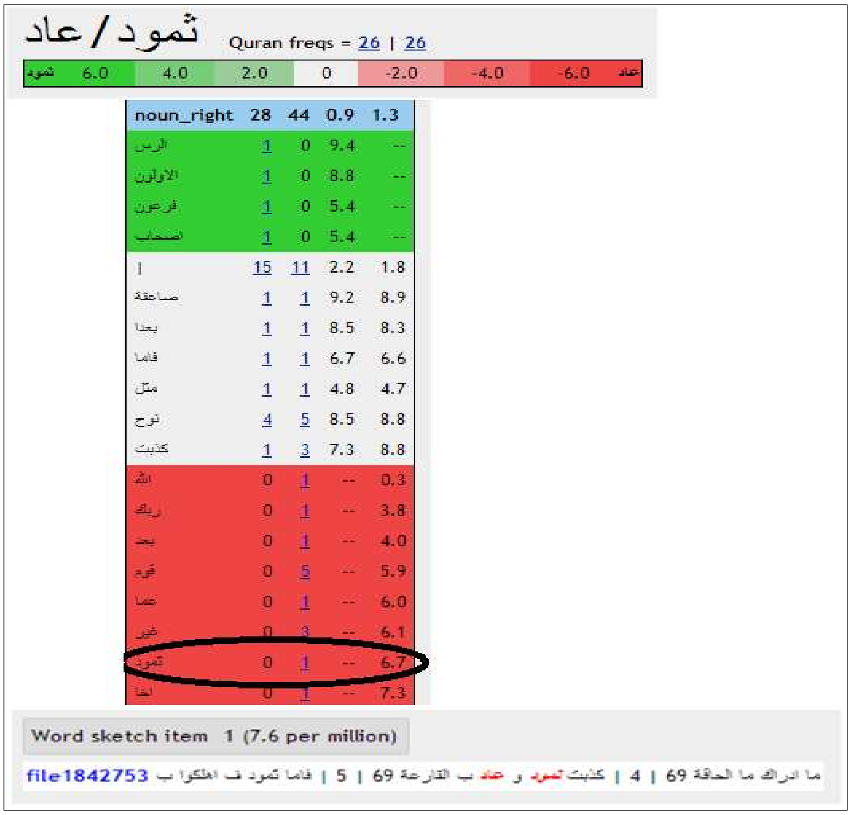
Show Diff

Fig. 4: Word Sketch differences entry form for the phrase “Thamud AND ‘Ad”

Where *p(x, y)* is the probability that the two words *x* and *y* occur together in the same verse, *p(x)* is the probability that word x occurs alone in that verse, and the same for *p(y).*

From Table 4, we can notice the difference in the association values between the two combined words with a different order in the conjunctive phrase. The phrase “بَشِير و نَذِير” “Bearer of glad tidings AND Warner” has an association value of 7.1641 whereas the phrase “نَذِير و بَشِير” “Warner AND Bearer of glad tidings” has 8.1641. High PMI value indicates a high degree of association relationship between the words and vice versa. Moreover, high frequent pairs of words have high association values compared to those with low frequency.

In addition, to validate the first approach, we use another method which is the word sketch difference function available in the Sketch Engine tool [39]. This function is used to compare any two words in their lemma form by displaying those patterns and combinations that the two words have shared in common or differentiated by. Besides that, there are four numbers next to each pattern; the first two show the frequency of co-occurrence with the first and the second word, whereas the last two show the salience scores for the pattern with both words (the Word Sketch Difference help).



ثمود / عاد Quran freqs = 26 | 26

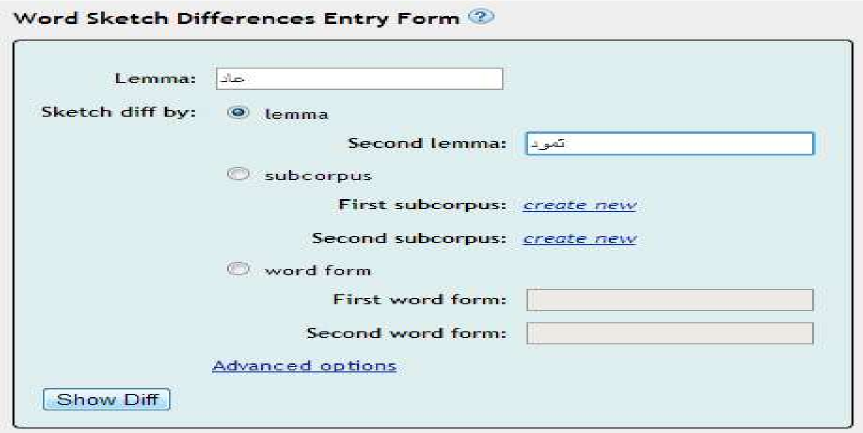
ثمود 6.0	4.0	2.0	0	-2.0	-4.0	-6.0 عاد

noun_right	28	44	0.9	1.3
الرس	1	0	9.4	--
الاولين	1	0	8.8	--
فرعون	1	0	5.4	--
اصحاب	1	0	5.4	--
ا	15	11	2.2	1.8
صاعقة	1	1	9.2	8.9
بعدا	1	1	8.5	8.3
فاما	1	1	6.7	6.6
مثل	1	1	4.8	4.7
نوح	4	5	8.5	8.8
كذبت	1	3	7.3	8.8
الله	0	1	--	0.3
ربك	0	1	--	3.8
بعد	0	1	--	4.0
قوم	0	5	--	5.9
عصا	0	1	--	6.0
غير	0	3	--	6.1
ثمود	0	1	--	6.7
اخا	0	1	--	7.3

Word sketch item 1 (7.6 per million)

file1842753 ما ادراك ما الحاقة 69 | 4 | كذبت ثمود و عاد ب القارعة 69 | 5 | فاما ثمود ف اهلكوا ب

Fig. 5: The association score and frequency of the phrase “Thamud AND ‘Ad”



Word Sketch Differences Entry Form

Lemma: عاد

Sketch diff by: ● lemma

Second lemma: ثمود

○ subcorpus

First subcorpus: create new

Second subcorpus: create new

○ word form

First word form:

Second word form:

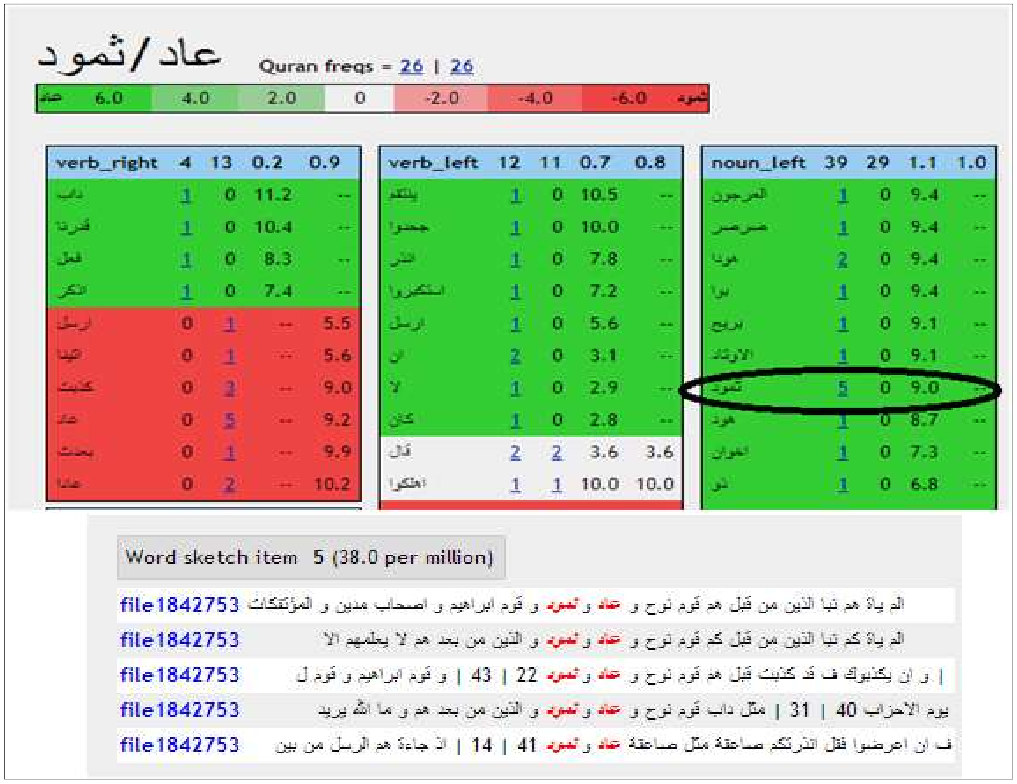
Advanced options

Show Diff

Fig. 6: Word Sketch differences entry form for the phrase “‘Ad AND Thamud”

As an example, we compare the two phrases “ثَمُود و عَاد” “Thamud AND ‘Ad” and “عَاد و ثَمُود” “‘Ad AND Thamud” using Word Sketch Differences as shown in Fig. 4 and Fig. 6.

Fig. 5 illustrates the impact of the word order in the conjunctive phrase on the association score. It is clear that the word 'ثمود' comes to the right of the word 'عاد' in this form only once in the holy Quran with an association value of 6.7. However, when the same word 'ثمود' comes to the left of the same word 'عاد', the association value increases to 9.0 with a frequency of 5, as depicted in Fig. 7.


عاد/ثمود Quran freqs = 26 | 26

عاد	6.0	4.0	2.0	0	-2.0	-4.0	-6.0	ثمود

verb_right	4	13	0.2	0.9
داب	1	0	11.2	--
قدرنا	1	0	10.4	--
فعل	1	0	8.3	--
الذكر	1	0	7.4	--
ارسل	0	1	--	5.5
اتينا	0	1	--	5.6
كذبت	0	3	--	9.0
عاد	0	5	--	9.2
بعدت	0	1	--	9.9
عادا	0	2	--	10.2

verb_left	12	11	0.7	0.8
ياتكم	1	0	10.5	--
جحدوا	1	0	10.0	--
انذر	1	0	7.8	--
استكبروا	1	0	7.2	--
ارسل	1	0	5.6	--
ان	2	0	3.1	--
لا	1	0	2.9	--
كان	1	0	2.8	--
قال	2	2	3.6	3.6
اهلكوا	1	1	10.0	10.0

noun_left	39	29	1.1	1.0
المرجون	1	0	9.4	--
صرصر	1	0	9.4	--
هودا	2	0	9.4	--
نبا	1	0	9.4	--
بريح	1	0	9.1	--
الاوتاد	1	0	9.1	--
ثمود	5	0	9.0	--
هود	1	0	8.7	--
اخوان	1	0	7.3	--
ذو	1	0	6.8	--

Word sketch item 5 (38.0 per million)

file1842753	الم ياة هم نبا الذين من قبل هم قوم نوح و عاد و ثمود و قوم ابراهيم و اصحاب مدين و المؤتفكات
file1842753	الم ياة كم نبا الذين من قبل كم قوم نوح و عاد و ثمود و الذين من بعد هم لا يعلمهم الا
file1842753	| و ان يكذبوك ف قد كذبت قبل هم قوم نوح و عاد و ثمود 22 | 43 | و قوم ابراهيم و قوم ل
file1842753	يوم الاحزاب 40 | 31 | مثل داب قوم نوح و عاد و ثمود و الذين من بعد هم و ما الله يريد
file1842753	ف ان اعرضوا فقل انذرتكم صاعقة مثل صاعقة عاد و ثمود 41 | 14 | اذ جاءة هم الرسل من بين

Fig. 7: The association score and frequency of the phrase "'Ad AND Thamud"

6.0 CONCLUSION

In this work, we have performed an analytical study on the Arabic conjunctive phrases, namely AND conjunction, extracted from the Quranic Arabic corpus. This research is very useful for religious scholars, scientist, and linguists because it shows the linguistic miracle of the holy Quran based on scientific evidence. We have analyzed the order of the two words that form the conjunctive phrase and its effect on the contextual meaning of the Quranic verse, where they have occurred and the association relationship between them. We have reported three different cases: words that have occurred in a specific order in the conjunctive phrase and repeated only once in the Quran, words that have occurred in a specific order in the conjunctive phrase and repeated many times in the Quran, and words that have occurred in two different orders in the conjunctive phrase and repeated one/many time(s) in the holy Quran. In the future, we plan to explore a wider range of Quranic co-occurred words and test different association measurements.

REFERENCES

[1] D. J. Hand, H. Mannila, and P. Smyth, *Principles of Data Mining (Adaptive Computation and Machine Learning)*, Massachusetts, USA, MIT Press, 2001.

[2] R.G. Raj and S. Abdul-Kareem, “Information Dissemination And Storage For Tele-Text Based Conversational Systems' Learning”, *Malaysian Journal of Computer Science*, Vol. 22 No. 2, 2009. pp. 138-159.

[3] J. Xu and W. B. Croft, “Improving the Effectiveness of Information Retrieval”, *ACM Transactions on Information Systems (TOIS)*, Vol. 18, No.1, 2000, pp. 79-112.

[4] A. Qazi, R. G. Raj, M. Tahir, M. Waheed, S. U. R. Khan, and A. Abraham, “A Preliminary Investigation of User Perception and Behavioral Intention for Different Review Types: Customers and Designers Perspective,” *The Scientific World Journal*, Vol. 2014, Article ID 872929, 8 pages, 2014. doi:10.1155/2014/872929.

[5] R. E. Banchs, *Text Mining with MATLAB*, New York, USA, Springer, 2013.

[6] A. Farghaly and K. Shaalan, “Arabic Natural Language Processing: Challenges and Solutions”, *ACM Transactions on Asian Language Information Processing (TALIP),* Vol. 8, No. 4, 2009, pp. 1-22.

[7] A. M. Saif and M. J. Aziz, “An Automatic Collocation Extraction from Arabic Corpus”, *Journal of Computer Science,* Vol. 7, No. 1, 2011, pp. 6-11.

[8] M. Alrabiah, A. Al-salman, and E. Atwell, “A New Distributional Semantic Model for Classical Arabic”, *2nd International Conference on Islamic Applications in Computer Science and Technology (IMAN 2014),* Amman, Jordan, 12-13 October 2014.

[9] M. Al-Kabi, H. Al-Belaili, B. Abul-Huda, and A. H. Wahbeh, “Keyword Extraction Based on Word Co-Occurrence Statistical Information for Arabic Text”, *ABHATH AL-YARMOUK: "Basic Sci. & Eng.",* Vol. 22, No. 1, 2013, pp. 75- 95.

[10] M. H. Panju, *Statistical extraction and visualization of topics in the qur’an corpus*, Master’s thesis, University of Waterloo, Ontario, Canada, 2014.

[11] M. Al-Kabi, G. Kanaan, R. Al-Shalabi, K. Nahar, and B. Bani-Ismail, “Statistical Classifier of the Holy Quran Verses (Fatiha and YaSeen Chapters)”, *Journal of Applied Sciences*, Vol. 5, No. 3, 2005, pp. 580-583.

[12] K. Dukes and N. Habash, “Morphological annotation of Quranic Arabic”, *The 7th International Conference on Language Resources and Evaluation (LREC)*, Valletta, Malta, May 2010, pp. 2530-2536.

[13] K. Dukes and T. Buckwalter, “A dependency treebank of the Quran using traditional Arabic grammar”, *The 7th International Conference on Informatics and Systems (INFOS),* Cairo, Egypt, March 2010, pp. 1-7.

[14] F. J. Alvarez, A. Vaquero, F. Sáenz, and M. Buenaga, “Semantic Relation Modeling and Representation for Problem-Solving Ontology-Based Linguistic Resources: Issues and Proposals”, *The 9th International Conference on Enterprise Information Systems (ICEIS),* Madeira, Portugal, June 2007, pp. 59-70.

[15] S. Momtazi, S. Khudanpur, and D. Klakow, “A comparative study of word co-occurrence for term clustering in language model-based sentence retrieval”, *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT 10),* California, USA, June 2010, pp. 325-328.

[16] A. Islam and D. Inkpen, “Second Order Co-occurrence PMI for Determining the Semantic Similarity of Words”, *The International Conference on Language Resources and Evaluation (LREC),* Genoa, Italy, May 2006, pp. 1033–1038.

[17] W. H. Gomaa and A. A. Fahmy, “A survey of text similarity approaches”, *International Journal of Computer Applications*, Vol. 68, No. 13, 2013, pp. 13-18.

[18] N. Habash, *"Introduction to Arabic Natural Language Processing*", California, USA, Morgan & Claypool Publisher, 2010.

[19] M. Alrabiah, N. Alhelewh, A. Al-Salman, and E. Atwell, "An Empirical Study On The Holy Quran Based On A Large Classical Arabic Corpus", *International Journal of Computational Linguistics (IJCL),* Vol. 5, No. 1, 2014, pp. 1-13.

[20] M. Attia, M. Rashwan, A. Ragheb, M. Al-Badrashiny, H. Al-Basoumy, and S. Abdou, "A compact Arabic lexical semantics language resource based on the theory of semantic fields", *The 6th International Conference on Natural Language Processing (GoTAL 2008),* Gothenburg, Sweden, August 2008, pp. 65-76.

[21] F. Thabtah, O. Gharaibeh, and R. Al-Zubaidy, "Arabic text mining using rule based classification", *Journal of Information & Knowledge Management*, Vol. 11, No. 01, 2012, pp. 1-10.

[22] M. Al-Yahya, S. Al-Malak, and L. Aldhubayi, "Ontological Lexicon Enrichment: The Badea System for Semi-automated Extraction of Antonymy Relations from Arabic Language Corpora*", Malaysian Journal of Computer Science*, Vol. 29, No. 1, 2016, pp 56-73.

[23] R. Safeena and A. Kammani, "Quranic Computation: A Review of Research and Application", *Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences (NOORIC 2013),* Madinah, Saudi Arabia, December 2013, pp. 203-208.

[24] H. Hamam, M. T. Ben Othman, A. Kilani, M. Ben Ammar, and F. Ncibi, "Data Mining in the Quran Using Aspects and Dependencies", *The 3rd International Conference on Islamic Applications in Computer Science and Technologies (IMAN 2015*), Konya, Turkey, 1st - 3rd October 2015.

[25] M. A. Siddiqui, S. M. Faraz, and S. A. Sattar, "Discovering the Thematic Structure of the Quran using Probabilistic Topic Model", *Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences (NOORIC 2013),* Madinah, Saudi Arabia, December 2013, pp. 234-239.

[26] A. B. M. Sharaf and E. Atwell, "QurAna: Corpus of the Quran annotated with Pronominal Anaphora", *The 8th International Conference on Language Resources and Evaluation (LREC 2012),* Istanbul, Turkey, May 2012, pp. 130-137.

[27] A. B. M. Sharaf and E. Atwell, "QurSim: A corpus for evaluation of relatedness in short texts", *The 8th International Conference on Language Resources and Evaluation (LREC 2012),* Istanbul, Turkey, May 2012, pp. 2295-2302.

[28] N. Abbas, *Qurany: A Tool to Search for Concepts in the Quran*, MSc Research Thesis, School of Computing, University of Leeds, Leeds, UK, 2009.

[29] Q. Ul Ain and A. Basharat, "Ontology driven Information Extraction from the Holy Quran related Documents", *The 26th Institution of Electrical and Electronics Engineers Pakistan Students' Seminar (IEEEP 2011),* Karachi, Pakistan, 16th - 17th March 2011.

[30] K. Dukes, E. Atwell, and N. Habash, "Supervised Collaboration for Syntactic Annotation of Quranic Arabic", *Language Resources and Evaluation*, Vol. 47, No. 1, 2013, pp. 33-62.

[31] M. Adhima, *Derasat li Osloob Al-Quraan Al-Karim* (In Arabic), Cairo, Egypt, Dar Al-Hadith, 1972.

[32] M. Al-Ghalayini, *Jamea Al-Dorous Al-Arabiya* (In Arabic), Cairo, Egypt, Dar Al-ghad Al-Jadid, 2007.

[33] A. Al-Zujaji, *Horouf Al-Maani* (In Arabic), Beirut, Lebanon, Muassasat Al-Ressala, 1984.

[34] J. Al- Soyouti, *Al-Itkan fi Oloum Al-Quran* (In Arabic), Beirut, Lebanon, Almaktaba Althakafiya, 1973.

[35] A. Abderrahman, *Al-Iajaz Albayani li Al-Quran* (In Arabic), Cairo, Egypt, Dar Al-Maaref, 1987.

[36] F. Al-Samiraii, *Al-Taabir Al-Qurani* (In Arabic), Amman, Jordan, Dar Ammar, 2006.

[37] M. Al- Masiri, *Dalalat Al-Takdim wa Al-Taakhir fi Al-Quran Al- Karim, Dirassa Tahliliya* (In Arabic), Cairo, Egypt, Maktabat Wahba, 2005.

[38] L. B. Huang, V. Balakrishnan, R.G. Raj, "Improving the relevancy of document search using the multi-term adjacency keyword-order model." Malaysian Journal of Computer Science, Vol. 25, No. 1, 2012, pp. 1-10.

[39] A. Kilgarriff, V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, and V. Suchomel, “The Sketch Engine: ten years on”, *Lexicography*, Vol. 1, No. 1, 2014, pp. 7–36.