

A Template-Based Approach for Tagging Non-Vocalized Arabic Nouns

By

Hashem Saadaldin Alghalib Alsharif

**A thesis submitted for the requirements of the degree
of Master of Science in Computer Science**

Supervised By

Dr. Khalid Omar Thabit

**FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY
KING ABDULAZIZ UNIVERSITY
JEDDAH – SAUDI ARABIA**



A Template-Based Approach for Tagging Non-Vocalized Arabic Nouns

By

Hashem Saadaldin Alghalib Alsharif

**A thesis submitted for the requirements of the degree
of Master of Science in Computer Science**

Supervised By

Dr. Khalid Omar Thabit

**FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY
KING ABDULAZIZ UNIVERSITY
JEDDAH – SAUDI ARABIA
Muharram 1436H – November 2014G**

A Template-Based Approach for Tagging Non-Vocalized Arabic Nouns

By

Hashem Saadaldin Alghalib Alsharif

**This thesis has been approved and accepted in partial
fulfillment of the requirements of the degree
of Master of Science in Computer Science**

EXAMINATION COMMITTEE

	Name	Rank	Field	Signature
Internal Examiner				
External Examiner				
Advisor				

KING ABDULAZIZ UNIVERSITY

Muharram 1436H – November 2014G (*Date of Defense*)

ABSTRACT

There exist no corpora of Arabic nouns. Furthermore, in any Arabic text, nouns can be found in different forms. In fact, by tagging nouns in an Arabic text, the beginning of each sentence can determine whether it starts with a noun or a verb.

Part of Speech Tagging (**POS**) is the task of labeling each word in a sentence with its appropriate category, which is called a **Tag** (*Noun, Verb and Article*). In this thesis, we attempt to tag non-vocalized Arabic text. The proposed POS Tagger for Arabic Text is based on searching for each word of the text in our lists of Verbs and Articles. Nouns are found by eliminating Verbs and Articles. Our hypothesis states that, if the word in the text is not found in our lists, then it is a Noun. These comparisons will be made for each of the words in the text until all of them have been tagged.

To apply our method, we have prepared a list of articles and verbs in the Arabic language with a total of 112 million verbs and articles combined, which are used in our comparisons to prove our hypothesis.

To evaluate our proposed method, we used pre-tagged words from "*The Quranic Arabic Corpus*", making a total of 78,245 words, with our method, the Template-based tagging approach compared with (AraMorph) a rule-based tagging approach and the Stanford Log-linear Part-Of-Speech Tagger.

Finally, AraMorph produced 40% correctly-tagged words and Stanford Log-linear Part-Of-Speech Tagger produced 68% correctly-tagged words, while our method produced 68,501 correctly-tagged words (88%).

TABLE OF CONTENTS

ABSTRACT.....	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	ix
Chapter I Introduction.....	1
1.1 What is a Word Template?.....	1
1.2 Natural Language Processing	2
1.3 Part of Speech Tagging	2
1.4 Objectives.....	2
1.5 Our Hypothesis.....	4
1.6 Arabic Verbs and Articles Lists	4
1.7 Thesis Layout	5
Chapter II Literature Review.....	6
2.1 Introduction	6
2.2 The Arabic Language	6
2.2.1 Arabic Morphology	7
2.2.2 Arabic Vowel Marks	8
2.2.3 Arabic as a Templatic Language	9
2.2.4 Four-Lettered Verb List.....	9

2.2.5	Articles List	11
2.3	POS Tagging Approaches used for Arabic	12
2.3.1	SVM.....	12
2.3.2	SVM + Morphological Analyzer.....	12
2.3.3	Statistical and Rule-based	12
2.3.4	Stanford PoS Tagger	13
2.3.5	Hidden Markov Model (HMM).....	13
2.3.6	Brill (Transformation) + Morphological Analyzer.....	13
2.3.7	Rules-based and Memory-based.....	14
2.3.8	Statistical	14
2.3.9	AraMorph.....	14
Chapter III	Methodology	16
3.1	Introduction	16
3.2	Verbs and Articles Lists	17
3.3	The Design	17
3.4	Implementation.....	20
3.4.1	Model Setup	20
3.4.2	Uploading the Articles and Verbs	21
3.4.3	Uploading the Raw Text File	21
3.4.4	Comparison Algorithm.....	22
3.4.5	Tagging Nouns	22
3.4.6	Python 3.4	23
3.5	Evaluation methods	23

Chapter IV Results and Conclusions	24
4.1 Introduction	24
4.2 AraMorph	25
4.3 The Stanford POS Tagger	26
4.4 The Quranic Arabic Corpus	26
4.5 Results	27
4.6 Conclusions	28
LIST OF REFERENCES	29
Appendix A	32
Appendix B	48

LIST OF FIGURES

Figure	Page
Figure 1: Transistor, clock, and power in Intel CPUs over time.	3
Figure 2: This table specifies the limits on physical memory for Windows 7	3
Figure 3: Arabic Verbs	9
Figure 4: Morphology rule attributes for Arabic verbs.....	10
Figure 5: An example of a four-lettered Arabic verb derivation.	10
Figure 6: The function used to add prefixes to Verbs and Articles.	11
Figure 7: The function employed to delete duplication.	11
Figure 8: An example of Arabic articles derivation.	12
Figure 9: An example of tagging an Arabic sentence.....	17
Figure 10: The arrays used in the buffer and the total number of elements.	18
Figure 11: Our proposed approach: the Template-Based Approach to POS Tagging.....	19
Figure 12: Reading the Verbs and Articles files and then removing any duplications, before sorting and storing them according to length.	21
Figure 13: Uploading the raw text file.....	21
Figure 14: Comparison Algorithm.....	22
Figure 15: Tagging nouns by default.	22
Figure 16: Part of the result.	23
Figure 17: Part of the results.....	24
Figure 18: Screenshot of the AraMorph system	25

Figure 19: Screenshot of the Stanford system	26
Figure 20: Screenshot of the Quranic Arabic Corpus website.....	26
Figure 21: Proof of the validity of the methodology	27
Figure 22: The Comparison Table	27

Chapter I

Introduction

Arabic texts can be either: **Vocalized** (*The Holy Quran*) or **Non-Vocalized** (*newspapers, books, and the media*). Handling non-vocalized texts is confusing; due to the ambiguity problem, words may have more than one meaning. "كتب" (*this can be a noun "books" or a verb "to write"*).

There exist no corpora of Arabic nouns. Furthermore, in any Arabic text, nouns can be found in different forms and can refer to non-Arabic items or things. In fact, by tagging nouns in an Arabic text, the beginning of each sentence can determine whether it starts with a noun or a verb.

In this chapter, we will discuss Word Templates, Natural Language Processing, and Part of Speech Tagging (**POS**), then we will state out objectives and the hypothesis used in this thesis, followed by an introduction to the Arabic words lists used in our approach. Finally, the structure of our thesis will be outlined.

1.1 What is a Word Template?

The term 'Word Template' is defined as any processing element that can be combined with a data model and processed by a template engine to produce a result. [1]

1.2 Natural Language Processing

Any language that is naturally used by humans (*Arabic and English*) is called Natural Language. In other words, a natural language is not an artificial or a man-made language, such as programming languages. Therefore, Natural language processing (NLP) is an area of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things. [2]

1.3 Part of Speech Tagging

Part of Speech Tagging (**POS**) is the task of labeling each word in a sentence with its appropriate category, which is called a **Tag** (*Noun, Verb and Article*) [3]. In this thesis, we attempt to tag non-vocalized Arabic text.

We intended to use a simplified tags set in our approach (*Noun, Verb and Article*) because our objective is to tag non-vocalized Arabic text, by applying a Template-based approach without explaining the construction of the Arabic sentence and also because most of the Arabic language is rooted to a verb, noun or article.

POS tagging is useful for a large number of applications. It is the first analysis step in many syntactic parsers; it is required for the correct lemmatization of words; and it is used in information extraction, vocalization, translation, speech synthesis, lexicographic research, term extraction, and many other applications.

1.4 Objectives

There are many methods of POS tagging which can be classified into three categories: the Statistical Approach, the Rule-Based Approach and the Hybrid Approach. All of these approaches are CPU consuming while our approach is memory consuming.

Since the 1980's, processor speed has increased by 1,000 times. Meanwhile, memory capacity has increased by 1,000,000 times.

It has been found that, since 2005, CPU speed has not followed the estimation of Moore's law. In fact, due to the thermal wall – *as he described it* – CPU speed has reached a certain limit. [4]

In Figure 1, you can see that transistor density has continued to climb according to Moore's Law. However, power consumption and CPU have reached a certain limit.

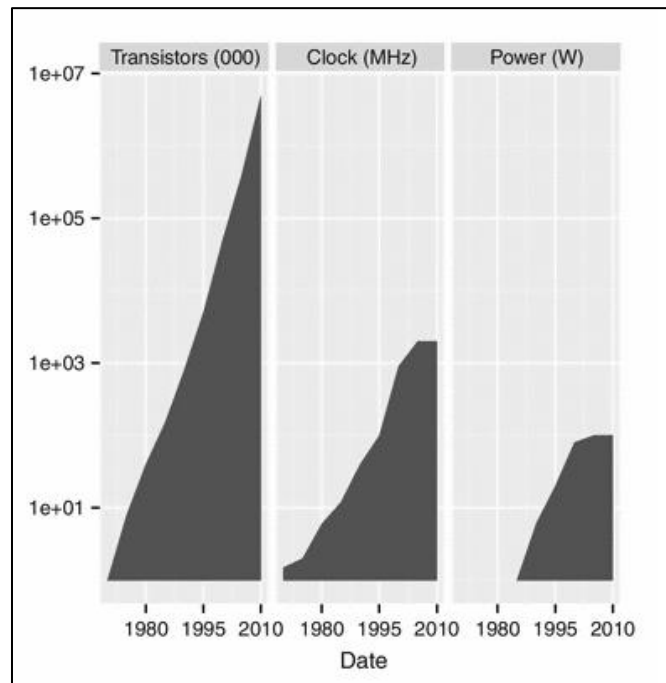


Figure 1: Transistor, clock, and power in Intel CPUs over time.

Meanwhile, the physical memory capacity is still growing; for example, Windows 7 64-bit memory limit is up to 192 GB.

Version	Limit on X86	Limit on X64
Windows 7 Ultimate	4 GB	192 GB
Windows 7 Enterprise	4 GB	192 GB
Windows 7 Professional	4 GB	192 GB

Figure 2: This table specifies the limits on physical memory for Windows 7

For the above, using a memory consuming approach is more appropriate than a CPU consuming one.

1.5 Our Hypothesis

The proposed POS Tagger for Arabic Text is based on searching for each word of the text in our lists of Verbs and Articles. Nouns are found by eliminating Verbs and Articles.

For example, a simple sentence like (حضر زيد إلى المدرسة) can be tagged by comparing the first word (حضر) with our Articles list. The word will be tagged "*Article*" if it is found. Otherwise, the same word will be compared with the Verb list. The word will be tagged "*Verb*" if it was found. Otherwise, by default, the word will be tagged "*Noun*". Our hypothesis states that, if the word in the text is not found in our lists, then it is a Noun. These comparisons will be made for each of the words in the text until all of them have been tagged.

This approach can be applied to Arabic language because it has static verb derivations and a limited number of words.

1.6 Arabic Verbs and Articles Lists

To apply our method, we must first prepare a list of articles and verbs in the Arabic language. Unfortunately, we only have a list of three-lettered verbs (*over 28 million verbs*) produced as text files by the Arabic Language Template Grammars Component Based Technology [5].

Four-lettered verbs (*about 2,000 verbs*) and articles (*over 4,000 articles*) of the Arabic language have not been employed before. All of our data, both new and old, must be prepared for our method, and we must add the proper prefix to the whole data.

The result was a total of 112 million verbs and articles combined, which are used in our comparisons to prove our hypothesis.

As will be discussed in sections 2.2.4 and 2.2.5, using a program that we have developed, a list of all four-lettered verbs and articles will be produced.

1.7 Thesis Layout

The next chapter, Chapter Two, "*Literature Review*," will provide a brief yet better understanding of the Arabic language, together with its rich and complex morphology, and also survey some of the similar work previously performed on the application of POS tagging approaches to Arabic.

In Chapter Three, "*Methodology*," we will discuss the methodology we applied to prove our thesis, and also provide the necessary methodology to design the system as well as analyze its different categories. The implementation of our model will be presented along with the evaluation method that will be used to test the proposed model fully.

Finally, in Chapter Four, "*Results and Conclusions*," a full test of the proposed model will be conducted and a comparison with the two other systems will be undertaken. Finally, a conclusion of the results at hand will be presented.

Chapter II

Literature Review

2.1 Introduction

POS tagging is one of the most important natural language problems studied by researchers. The significance of POS tagging for language processing is the large amount of information it provides about a word and its neighbors [6].

POS tagging is the process of assigning a part-of-speech or other syntactic class marker to each word in the corpus [6]. It is, in other words, the process of assigning a tag from a limited set of tags (*Tag Set*) to a word. The number of tags in a tag set depends on both the language and the intended application. If we talk about tagging, then we always mean some tag set, perhaps implicitly.

There are many methods associated with POS tagging. Most of the modern methods use some form of machine learning. In this chapter, we will present an introduction to Arabic Language and Arabic Morphology, then outline the POS tagging approaches used for Arabic.

2.2 The Arabic Language

Arabic is ranked sixth in the world's league table of languages. With 22 countries defining Arabic as their official language, it has an estimated 186 million native speakers and is spoken by at least 250 million people. As the language of the Quran, the holy book

of Islam, it is also widely used throughout the Muslim world. The Arabic language belongs to the Semitic group of languages, which also includes Hebrew and Amharic, the main language of Ethiopia.

Arabic is one of the few languages that exhibits diglossia, which is the separation between the spoken language (*dialects*) and the formal language. However, the formal language is Modern Standard Arabic (**MSA**), which is used in written texts and spoken in formal settings [7].

Arabic is written using the Arabic alphabet, from right to left, and there are no capital letters. Arabic has two genders: masculine and feminine. Arabic letters are linked together to form words, so some letters may change shape according to their location in the word.

There are 29 letters in the Arabic language, with seven letter sounds that do not exist in the English language (ح، خ، ص، ض، ط، ع، غ). Arabic uses vocalized symbols to determine the sound of a word, and these symbols may or may not be written down, resulting in a very high level of ambiguity.

In addition to the singular and plural constructs, Arabic has an added form called the "*dual*", which indicates precisely two of something. Arabic has only three verb tenses: the past, the present, and the imperative.

2.2.1 Arabic Morphology

Morphology is concerned with the structure of words. It is almost inconceivable for a natural language application not to employ morphological knowledge [8].

Like any other Semitic language, Arabic is highly inflected. Based on what is normally called a "*Constant Root System*", words are derived from a root and pattern, combined with prefixes, suffixes, and circumfixes.

The root, which contains the seed meaning of the word, consists of 3-4 constants, that are called radicals, and the pattern or a "*template*" is a sequence of variables. Patterns are simply all of the different variations of prefixes, infixes, and suffixes that can be allocated to any given root.

Arabic compound words are created by assigning the root radicals to the pattern variables. The combinations of the same root with a different pattern may result in a different meaning.

Sometimes, prefixes and/or suffixes are attached to words. Those affixes may modify several features of the word, including its number (*singular, dual, plural or collective*), gender (*masculine, feminine or no gender*), possession, definiteness, case (*nominative, accusative, or genitive*), tense (*past, present, and future*) and more.

Arabic affixes have the feature of concatenating with each other according to predefined linguistic rules, which increases the overall number of affixes [9].

2.2.2 Arabic Vowel Marks

In Arabic, there are three kinds of vowels:

- The three vowel letters, which are (ا), (و) and (ي). These are used for long vowels.
- The "Hamza".
- The vowel marks which are used to denote short vowels.

To distinguish short vowels from long ones when words are read, Arabic script uses vowel marks. This is implemented by writing the marks over or under a letter.

In modern day Arabic text, these marks are usually not written because Arabic readers can always guess them. Therefore, in our approach, we will consider a tagging system for non-vocalized Arabic text (*a text with no vowel marks*).

2.2.3 Arabic as a Templatic Language

Perhaps the most interesting aspect of Arabic is its regularity of form. Arabic morphemes "words" generally derive from three root radicals to which various affixes (*prefixes, suffixes, and infixes*) can be attached to create a word[10].

2.2.4 Four-Lettered Verb List

As will be described later in Figure 3, a verb is divided into two main types: complete (تام), as in (عسكر), and deficient (معتل), as in (وسوس).

A complete verb, as in (عسكر), is again divided into two types: intransitive (لازم), as in (بعثر), and transitive (متعد), as in (تزلزل).

A transitive verb, as in (تزلزل), is divided into two types: passive (مبني للمجهول) and active (مبني للمعلوم).

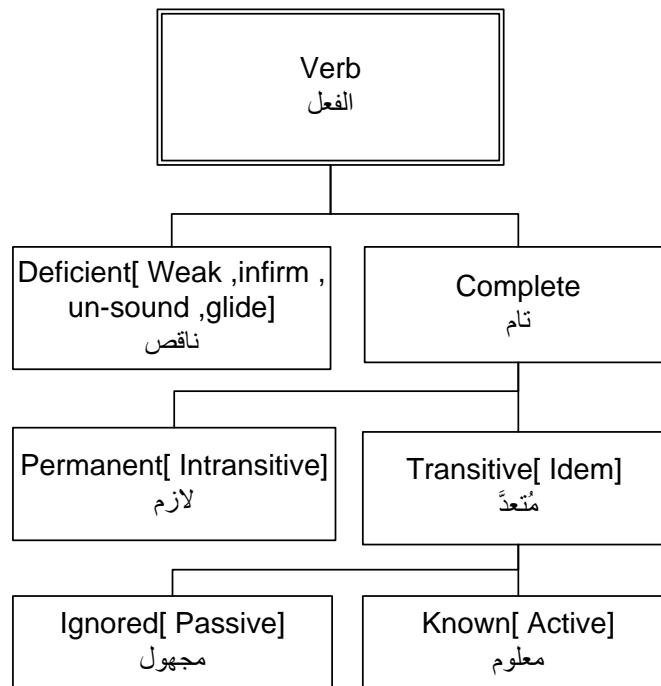


Figure 3: Arabic Verbs

All o these types must be considered in order to create a complete list of four-lettered Arabic verbs. Figure 4 describes the morphology rules used to generate all Arabic verbs.

		Number	Personal Pronoun		
			غائب 3 rd	مخاطب 2 nd	متكلم 1 st
Masculine	ذكر Male	Singular مفرد	هو	أنت	أنا
		Dual مثنى	هما	أنتما	نحن
		Plural جمع	هم	أنتم	نحن
	أنثى Feminine	Singular مفرد	هي	أنتِ	أنا
		Dual مثنى	هما	أنتما	نحن
		Plural جمع	هن	أنتن	نحن

Figure 4: Morphology rule attributes for Arabic verbs

For example, or the four-lettered Arabic verb derivation, we considered the verb (عسكر).

As shown in Figure 5, a complete derivation of the verb (عسكر) was presented.

	الماضي	المضارع	المضارع المحزوم	المضارع المنصوب	المضارع المؤكد	الأمر	الأمر المؤكد
أنت	عسكرت	تعسكر	تعسكر	تعسكر	تعسكرون	عسكر	عسكرون
أنت	عسكرت	تعسكرون	تعسكري	تعسكري	تعسكرون	عسكري	عسكرون
أنتما	عسكرتما	تعسكران	تعسكرا	تعسكرا	تعسكران	عسكرا	عسكران
أنتم	عسكركم	تعسكرون	تعسكروا	تعسكروا	تعسكرون	عسكروا	عسكرون
أنتن	عسكرتن	تعسكرون	تعسكرون	تعسكرون	تعسكرنان	عسكرون	عسكرنان
أنا	عسكرت	أعسكر	أعسكر	أعسكر	أعسكرون		
نحن	عسكرنا	نعسكر	نعسكر	نعسكر	نعسكرون		
هو	عسكر	يعسكر	يعسكر	يعسكر	يعسكرون		
هي	عسكرت	تعسكر	تعسكر	تعسكر	تعسكرون		
هما	عسكرا	يعسكران	يعسكرا	يعسكرا	يعسكران		
هم	عسكروا	يعسكرون	يعسكروا	يعسكروا	يعسكرون		
هن	عسكرون	يعسكرون	يعسكرون	يعسكرون	يعسكرنان		

Figure 5: An example of a four-lettered Arabic verb derivation.

In our approach, which is based on comparisons, all of the verbs and articles in the Arabic text must be denoted in our lists in order to define Nouns by elimination. Therefore, it is necessary to add the proper prefix and suffix. Many of the prefixes and

suffixes are generated within derivations of verbs and articles. However, we have added the missing prefixes as shown in the figure below.

```
for line in infile:
    if len(line)>2:
        outfile.write(line + u'\r\n')
        outfile.write(u"و" + line + u'\r\n')
        outfile.write(u"ف" + line + u'\r\n')
        outfile.write(u"ل" + line + u'\r\n')
        outfile.write(u"س" + line + u'\r\n')
    else:
        outfile.write(line)
```

Figure 6: The function used to add prefixes to Verbs and Articles.

Naturally, since we are focusing on tagging non-vocalized Arabic text, many derivations of the same verb will be duplicated. Therefore, we will have a redundancy problem. In Figure 5, after extensive revision, it is clear that most of the verb (عسكر) derivations are the same. Deleting these duplicates is mandatory.

```
def del_dups(seq):
    seen = {}
    pos = 0
    for item in seq:
        if item not in seen:
            seen[item] = True
            seq[pos] = item
            pos += 1
    del seq[pos:]
```

Figure 7: The function employed to delete duplication.

Finally, by using a program that we developed, we found and listed the total of over 2,000 four-lettered Arabic verbs, with all of their unique derivations.

*(***** as shown in appendix A)*

2.2.5 Articles List

In Arabic language, there are about 80 articles, known as: (حروف المعاني) [11].

نحن	أنا	هن	هي	هم	هما	هو	أنتن	أنت	أنتم	أنتما	أنت	الضمير
أينا	أبي	أيهن	أيها	أيهم	أيهما	أيه	أيكن	أيك	أيكم	أيكما	أيك	آي
إلينا	إلي	إلهن	إليها	إليهم	إليهما	إليه	إليكن	إليك	إليكم	إليكما	إليك	إلي
أننا	أني	أنهن	أنها	أنهم	أنهما	أنه	أنكن	أنك	أنكم	أنكما	أنك	أن
إننا	إني	إنهن	إنها	إنهم	إنهما	إنه	إنكن	إنك	إنكم	إنكما	إنك	إن
أينا	أبي	أيهن	أيها	أيهم	أيهما	أيه	أيكن	أيك	أيكم	أيكما	أيك	أي

Figure 8: An example of Arabic articles derivation.

Applying the morphology rules used to generate all Arabic verbs to all Arabic articles using a program that we have developed, keeping in mind the redundancy problem, we managed to list a total of over 3,000 articles with all of their unique derivations.

*(***** as shown in appendix B)*

2.3 POS Tagging Approaches used for Arabic

2.3.1 SVM

Diab [12, 13] applied a Support Vector Machine (**SVM**) to Arabic POS tagging and tokenization. The SVM-POS tagger achieved an accuracy level of 95.49%. The Arabic TreeBank, consisting of 4,519 sentences, was used for the training and testing. It employs the LDC's POS tagset, which consists of 24 tags³⁵.

2.3.2 SVM + Morphological Analyzer

In Habash [14], a Support Vector Machine (**SVM**) was applied with the support of a morphological analyzer to produce all possible analyses. Their POS evaluation shows an accuracy of 97.6% on ATB1 and an accuracy of 95.7% on ATB2, both of which are based on gold standard tokenization.

2.3.3 Statistical and Rule-based

In Khoja [15], a system is developed, using a combination of both statistical and rule-based techniques. It uses a simple tagset. A corpus of 50,000 Modern Standard Arabic

words (*an extract from the Saudi Al-Jazirah newspaper, dated 03/03/1999*) was tagged using this tagset 36. It achieved an accuracy of around 90%.

2.3.4 Stanford PoS Tagger

This was originally developed for use with English language at Stanford University. The tagger is based on the maximum entropy model. The improved version, which was published in 2003, adds support for other languages together with speed and usability improvements.

The latest version comes with trained models for Chinese, German and Arabic. It claims to have a 96.42% accuracy for Arabic. The tagger was trained using the training component of the Arabic Penn Treebank (ATB). It uses augmented Bies mapping of ATB tags [16].

2.3.5 Hidden Markov Model (HMM)

In Al Shamsi [17], the proposed Hidden Markov Model (HMM) POS tagger has been tested and achieved a performance of 97%. It used a very simple POS tag set of 55 tags. The training was performed on a special small corpus consisting of a 9.15 MB corpus of native Arabic articles. The authors used a stemmer for segmenting and separating affixes from the stem to produce prefix, stem, and suffix parts.

2.3.6 Brill (Transformation) + Morphological Analyzer

In AlGahtani [3], transformation-based learning was used, as implemented in the Brill tagger [18], for POS tagging of Arabic, with segment-based tags. They used the Buckwalter morphological analyzer [19], and their approach was evaluated on the whole ATB as well as on ATB1. For ATB1, they achieved a POS tagging accuracy of 96.9%. Using the whole ATB, the accuracy was 96.1%, even though large parts of the Treebank are duplicated, so that it is likely that parts of their test set were actually present in the training set [3].

2.3.7 Rules-based and Memory-based

In Tlili-Guiassa [20], a hybrid of rule-based and memory-based learning methods was used. Their method is based firstly on rules that are automatically learned from the training corpus (*that consider the post-position, ending of a word and patterns*) and then the anomalies were corrected by adopting a Memory-Based Learning Method (**MBL**). Secondly, by checking the exceptions to the rules, more information was made available to the learner for treating these exceptional cases. The accuracy level was 85%. The tag set was derived from that of Khoja [15].

2.3.8 Statistical

In Mohamed [21], two approaches were used. Their first approach used complex tags that described full words and did not require any word segmentation. The second approach was segmentation-based, using a segment based on machine learning. They showed that word-based POS tagging can yield better results than segment-based tagging (93.93% vs. 93.41%). Combining both methods resulted in a word accuracy of 94.37%. The POS tag set of the Penn Arabic Treebank was used and two sections of the ATB (**P1V3** and **P3V1**), since those two sets do not contain duplicate sentences. This data set contained approximately 500,000 words.

2.3.9 AraMorph

The Buckwalter Arabic morphological analyzer is one of the best-known Arabic morphological analysis and POS tagging systems. It was developed by LCD (*Linguistic Data Consortium*) in both Perl and Java.

The components of the Buckwalter Arabic Morphological Analyzer are the morphology analysis algorithm, and the data, that primarily consist of three Arabic/English lexicon files: **dictPrefixes** contains 299 entries, **dictSuffixes** contains 618 entries, and **dictStems** contains 82,158 entries, representing 38,600 lemmas.

These lexicons are supplemented by three morphological compatibility tables that are used to control the **prefix-stem** combinations (*1,648 entries*), **stemsuffix** combinations (*1,285 entries*), and **prefix-suffix** combinations (*598 entries*). The algorithm of the morphology analysis and POS tagging is imbedded in the code. It uses the three lexicon files and three compatibility tables in order to perform the morphological analysis and tagging of Arabic words [19].

Chapter III

Methodology

3.1 Introduction

Any linguistic or literary investigation of texts can benefit from computational technology. Evidently, the use of computational dictionaries, concordances and indexes, augmented by sophisticated search tools, can be extremely useful for scholars who are interested in such investigations. As a first step toward achieving this goal, we must first find a much faster and reliable system that can tag any non-vocalized Arabic text.

In this chapter, we will discuss the methodology that we applied to prove our thesis. As stated before (*in chapter one*), the proposed POS Tagger for Arabic Text is based on searching for each word in a text in our lists of Verbs and Articles. Nouns are tagged through the elimination of Verbs and Articles.

Our methodology can be described as a template-based tagging system, because every derivation of the Arabic words is produced from a template. For example, a simple sentence like (حضر زيد إلى المدرسة) can be tagged by comparing the first word (حضر) with our Articles list, which contains every Arabic article and all of their template derivations. The word will be tagged "Article" if it is found. Otherwise, the same word will be compared with the Verb list, which contains every Arabic verb and all of their template derivations. The word will be tagged "Verb" if it is found. Otherwise, the word

will be, by default, tagged as "*Noun*". These comparisons will be made for every word until they have all been tagged.

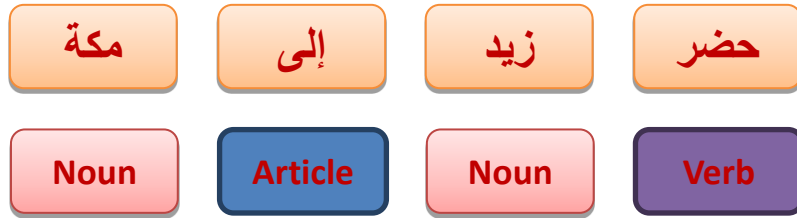


Figure 9: An example of tagging an Arabic sentence.

3.2 Verbs and Articles Lists

To apply our method, as discussed in section 1.6, we must first prepare a list of the articles and verbs of the Arabic language. The comparisons that we will make to tag a non-vocalized Arabic text must relate to whole words (including the *prefix*, *infix* and *suffix*).

Many of the prefixes and suffixes in our verbs and articles lists are generated within the derivations of the verbs and articles. However, we have added the missing prefixes through a program to each and every Verb and Article.

3.3 The Design

A methodological approach to building the model design is required in order to continue. The methodology for this design is based on a database of Arabic verbs and articles, internal buffers, a comparison algorithm, and a raw, non-vocalized Arabic text.

First, the verbs and articles must be uploaded, and then both the Arabic verbs as well as the Arabic articles must be arranged into arrays according to their length. When the arrangement has been completed, each word of the raw, non-vocalized Arabic text is compared with both Articles and Verbs arrays to produce the resulting tagged text.

The proposed model includes the following steps:

Step-1 Uploading Articles and Verbs:

In this step, the proposed design reads the database of Articles and Verbs and stores them in the buffer. According to the length of each Article and Verb, the system will store them into arrays. For example, the verb "حضر" will be stored in the three-lettered array. (See Figure 10)

Less Than Two Letters	16
Two Letters	2,190
Three Letters	42,466
Four Letters	447,101
Five Letters	2,075,181
Six Letters	5,700,889
Seven Letters	12,166,765
Eight Letters	19,839,791
Nine Letters	24,240,970
Ten Letters	22,405,515
Eleven Letters	15,291,729
More Than Eleven Letters	10,360,471
Total	112,573,084

Figure 10: The arrays used in the buffer and the total number of elements.

Step-2 Uploading the Raw Text File:

At this point, the proposed design reads the raw, non-vocalized Arabic text file word by word and stores them in the buffer. Then, the system will identify the length of each raw word.

Step-3 Comparison Algorithm:

In this step, using a binary search algorithm, according to the length of each non-vocalized raw word, the system will compare it with the proper length array of the Articles. If a match is found, the raw, non-vocalized Arabic word will be tagged "Article". Otherwise, the system will compare it with the proper length array of Verbs. If the raw word is found, it will be tagged "Verb".

Step-4 Tagging Nouns:

By default, if the system does not find the word from the raw Arabic text in either the Verbs and Articles lists, it will be tagged "*Noun*".

These steps are repeated as a cycle for each word in the raw Arabic text file. In Figure 11, the proposed system is shown with all of its components, followed by a complete notation of each component.

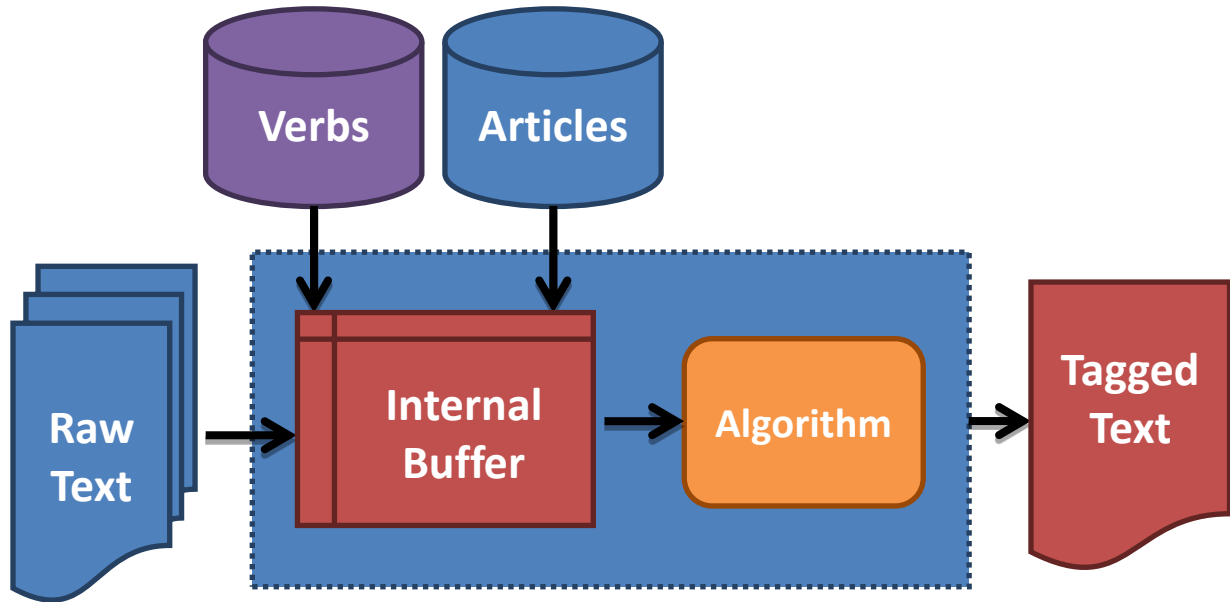


Figure 11: Our proposed approach: the Template-Based Approach to POS Tagging

Raw Text: A non-vocalized, traditional Arabic text that can be found, for example, in newspapers, books and on Arabic websites.

Verbs List: A list of all Arabic verbs and all of their derivations that we have developed by combining three-lettered verbs (*from Ba-Aziz* [5]) with the four-lettered verbs that we have found.

Articles List: A list of all of the Arabic articles and all of their derivations that we have found and developed.

Internal Buffer: Different variables (*or arrays*) that are used to contain the uploaded lists of verbs and articles, which are used to compare the data according to the length of the words.

Algorithm: A binary search searches sorted lists using a divide and conquer technique. On each iteration, the search domain is halved, until the result is found. Since our Verbs and Articles lists are sorted, using a binary search algorithm is the optimal choice.

Tagged Text: The result of the proposed model. Each word of the input is tagged with one of the following tags: *Article*, *Verb* or *Noun*.

3.4 Implementation

The model will be built based on the requirements mentioned in the previous sections. That will include building a program based on the initial analysis. In the initial phases of the research, it was hoped that an existing Arabic three-lettered verbs list would be available. The verbs list was provided as a text file by the Arabic Language Template grammars' component-based technology [5].

The four-lettered verbs (*about 2,000 verbs*) and the articles (*over 4,000 articles*) of the Arabic language have never been handled before. All of our data, both new and old, must be prepared for our method; we must add the proper prefixes to the whole data.

The result produced a total of over 112 million verbs and articles, which are used in our comparisons to prove our hypothesis.

3.4.1 Model Setup

The model was built using Python 3.4 language, involving the 64-bit version, on a Windows 7 64-bit platform using an Intel i7 processor with 16G bytes of memory. The system programming consisted of over 800 lines with over 10 functions.

3.4.2 Uploading the Articles and Verbs

To upload Articles and Verbs according to their length, the program first removed any duplications then sorted and stored them into arrays.

```
for item in newWordLine:
    if len(item)<2: lessArray.append(item+'\n')
    elif len(item)==2: twoArray.append(item+'\n')
    elif len(item)==3: threeArray.append(item+'\n')
    elif len(item)==4: fourArray.append(item+'\n')
    elif len(item)==5: fiveArray.append(item+'\n')
    elif len(item)==6: sixArray.append(item+'\n')
    elif len(item)==7: sevenArray.append(item+'\n')
    elif len(item)==8: eightArray.append(item+'\n')
    elif len(item)==9: nineArray.append(item+'\n')
    elif len(item)==10: tenArray.append(item+'\n')
    elif len(item)==11: elevenArray.append(item+'\n')
    elif len(item)>11: moreArray.append(item+'\n')

lessArray = del_dups(lessArray)
lessArray = sort_array(lessArray)
outfile = codecs.open("lessArray-I" + ".txt", encoding="utf-8" ,mode='w')
for word in lessArray:
    outfile.write(word)
outfile.close()
lessArray=[]
print ("Done lessArray file..")

twoArray = del_dups(twoArray)
twoArray = sort_array(twoArray)
outfile = codecs.open("twoArray-I" + ".txt", encoding="utf-8" ,mode='w')
for word in twoArray:
    outfile.write(word)
outfile.close()
twoArray=[]
print ("Done twoArray file..")
```

Figure 12: Reading the Verbs and Articles files and then removing any duplications, before sorting and storing them according to length.

3.4.3 Uploading the Raw Text File

The program now reads the raw, non-vocalized Arabic text file, word by word, and stores them in the buffer.

```
filename = "D:\Test\Quran.txt"
if filename != '':
    f1 = open(filename,'r')
    strfiletext = f1.read()
    f1.close()
```

Figure 13: Uploading the raw text file.

3.4.4 Comparison Algorithm

According to the length of each non-vocalized raw word, the program will compare it using a binary search algorithm with the proper length array of Articles and Verbs.

```
def binary_search(a, x):
    i = bisect_left(a, x)
    if i != len(a) and a[i] == x:
        return i
    #raise ValueError
    return -1
with codecs.open('Result-Report-Quran.txt', encoding='utf-8',mode='w') as outfile:
    for item in strQuran:
        if binary_search(strOrderedParticles,item)>=0:
            numArticle+=1
            outfile.write("The word (" +item+")\tFound as ARTICLE\r\n")
        elif len(item)<2 and binary_search(lessArray,item)>=0:
            numVerb+=1
            outfile.write("The word (" +item+")\tFound As VERB in lessArray \r\n")
        elif len(item)==2 and binary_search(twoArray,item)>=0:
            numVerb+=1
            outfile.write("The word (" +item+")\tFound As VERB in twoArray \r\n")
        elif len(item)==3 and binary_search(threeArray,item)>=0:
            numVerb+=1
            outfile.write("The word (" +item+")\tFound As VERB in threeArray \r\n")
        elif len(item)==4 and binary_search(fourArray,item)>=0:
            numVerb+=1
            outfile.write("The word (" +item+")\tFound As VERB in fourArray \r\n")
        elif len(item)==5 and binary_search(fiveArray,item)>=0:
```

Figure 14: Comparison Algorithm.

3.4.5 Tagging Nouns

By default, if the program did not find the word from the raw Arabic text in either the Verbs or Articles lists, it will be tagged "*Noun*".

```
elif len(item)==9 and binary_search(nineArray,item)>=0:
    numVerb+=1
    outfile.write("The word (" +item+")\tFound As VERB in nineArray \r\n")
elif len(item)==10 and binary_search(tenArray,item)>=0:
    numVerb+=1
    outfile.write("The word (" +item+")\tFound As VERB in tenArray \r\n")
elif len(item)==11 and binary_search(elevenArray,item)>=0:
    numVerb+=1
    outfile.write("The word (" +item+")\tFound As VERB in elevenArray \r\n")
elif len(item)>11 and binary_search(moreArray,item)>=0:
    numVerb+=1
    outfile.write("The word (" +item+")\tFound As VERB in moreArray \r\n")
else:
    numNoun+=1
    outfile.write("The word (" +item+")\tFound As NOUN\r\n")
```

Figure 15: Tagging nouns by default.

The word	(بسم)	Found As VERB in threeArray
The word	(الله)	Found As NOUN
The word	(الرحمن)	Found As NOUN
The word	(الرحيم)	Found As NOUN
The word	(قل)	Found As VERB in twoArray
The word	(أعوذ)	Found As VERB in fourArray
The word	(رب)	Found As NOUN
The word	(الناس)	Found As NOUN
The word	(ملك)	Found As VERB in threeArray
The word	(الناس)	Found As NOUN
The word	(إله)	Found As NOUN
The word	(الناس)	Found As NOUN
The word	(من)	Found as ARTICLE
The word	(شر)	Found As VERB in twoArray
The word	(الوسواس)	Found As NOUN
The word	(الخناس)	Found As NOUN
The word	(الذي)	Found as ARTICLE
The word	(يوسوس)	Found As NOUN
The word	(في)	Found as ARTICLE
The word	(صدور)	Found As NOUN
The word	(الناس)	Found As NOUN
The word	(من)	Found as ARTICLE
The word	(الجنة)	Found As NOUN
The word	(والناس)	Found As NOUN

Figure 16: Part of the result.

3.4.6 Python 3.4

Python is a powerful computer language that is used in a variety of applications. It emphasizes code readability, a huge standard library and, above all, it is open source. Using Python 3.4 enabled us to benefit from deploying readymade components' features.

3.5 Evaluation methods

There exist different ways of evaluating a tagger; among the most common are the accuracy or success rate.

$$\text{Success rate} = \frac{\text{Correctly tagged tokens}}{\text{Total number of tokens}} \times 100$$

The success rate is defined over all tags and evaluated by comparing the tags assigned by a tagger with each word assigned to its best tag.

Chapter IV

Results and Conclusions

4.1 Introduction

In this chapter, we will discuss the results of our research, and the fulfillment of our objectives.

To evaluate our proposed method, we used pre-tagged words from "The Quranic Arabic Corpus", making a total of 78,245 words [22], with our method, the Template-based tagging approach compared with (AraMorph) a rule-based tagging approach [19] and the Stanford Log-linear Part-Of-Speech Tagger [16].

Human Tagged		AraMorph		Correctnes	Stanford		Correctnes	Our Method		Correctnes
الم	N	الم	V	0	الم	V	0	الم	N	1
ذلك	A	ذلك	U	U	ذلك	A	1	ذلك	N	0
الكتاب	N	الكتاب	U	U	الكتاب	N	1	الكتاب	N	1
لا	A	لا	U	U	لا	A	1	لا	A	1
ريب	N	ريب	N	1	ريب	N	1	ريب	N	1
فيه	A	فيه	A	1	فيه	N	0	فيه	A	1
هدى	N	هدى	V	0	هدى	N	1	هدى	V	0
للمتقين	N	للمتقين	A	0	للمتقين	N	1	للمتقين	N	1
الذين	A	الذين	U	U	الذين	A	1	الذين	A	1
يؤمنون	V	يؤمنون	U	U	يؤمنون	V	1	يؤمنون	V	1
بالخوب	N	بالخوب	A	0	بالخوب	N	1	بالخوب	N	1
ويقيمون	V	ويقيمون	U	U	ويقيمون	N	0	ويقيمون	V	1
الصلاة	N	الصلاة	U	U	الصلاة	N	1	الصلاة	N	1
ومما	A	ومما	U	U	ومما	N	0	ومما	N	0
رزقناهم	V	رزقناهم	V	1	رزقناهم	N	0	رزقناهم	N	0
ينفقون	V	ينفقون	U	U	ينفقون	V	1	ينفقون	V	1
والذين	A	والذين	U	U	والذين	N	0	والذين	A	1
يؤمنون	V	يؤمنون	U	U	يؤمنون	V	1	يؤمنون	V	1

Figure 17: Part of the results.

4.2 AraMorph

The Buckwalter Arabic morphological analyzer is one of the best-known Arabic morphological analysis and POS tagging systems. It was developed by LCD (*Linguistic Data Consortium*) in both Perl and Java.

The components of the Buckwalter Arabic Morphological Analyzer are the morphology analysis algorithm and the data, that primarily consist of three Arabic/English lexicon files: **dictPrefixes** contains 299 entries, **dictSuffixes** contains 618 entries, and **dictStems** contains 82,158 entries, representing 38,600 lemmas.

These lexicons are supplemented by three morphological compatibility tables that are used to control **prefix-stem** combinations (1,648 entries), **stemsuffix** combinations (1,285 entries), and **prefix-suffix** combinations (598 entries). The algorithm for the morphology analysis and POS tagging is imbedded in the code. It uses the three lexicon files and the three compatibility tables in order to perform morphological analysis and the tagging of Arabic words [19].

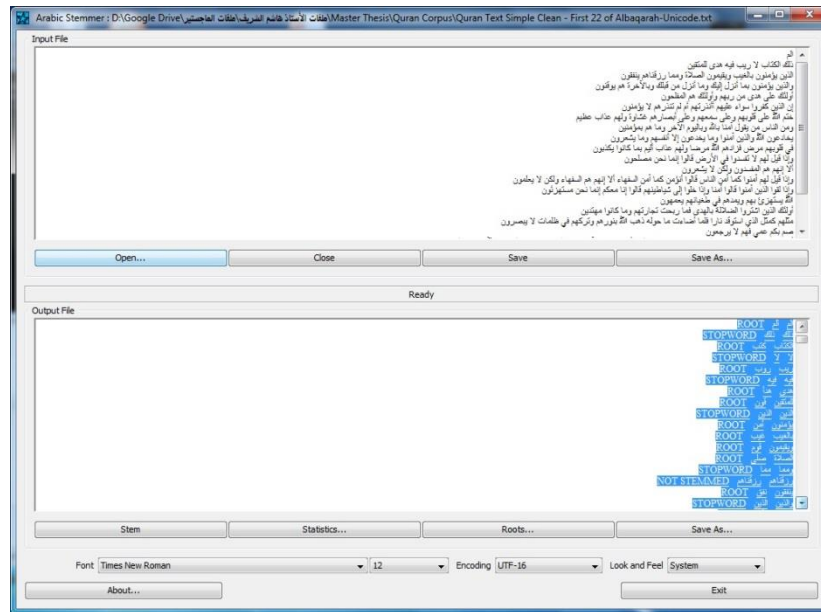


Figure 18: Screenshot of the AraMorph system

4.3 The Stanford POS Tagger

The Stanford part-of-speech tagger uses both preceding and following tag contexts via a dependency network representation, lexical features, including jointly conditioning multiple consecutive words, the effective use of priors in conditional log linear models, and the fine-grained modeling of unknown word features. By combining these ideas, the tagger gives a 97.24% accuracy level on the **Penn Treebank WSJ**, an error reduction of 4.4% on the best previous single automatically learned tagging result [16].

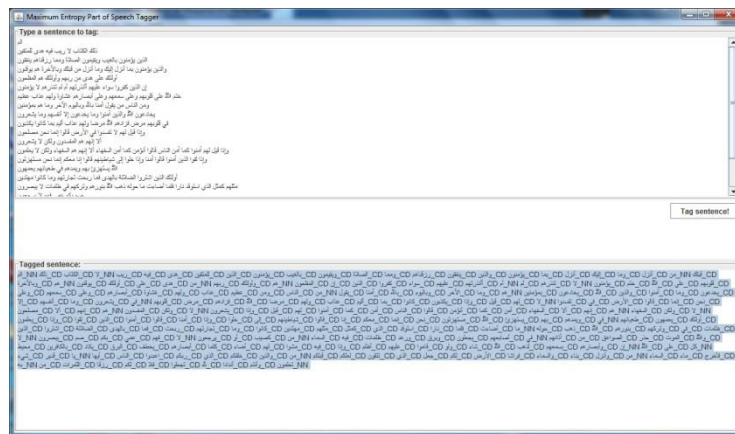


Figure 19: Screenshot of the Stanford system

4.4 The Quranic Arabic Corpus

The Quranic Arabic Corpus is an online, annotated linguistic resource with multiple layers of annotation, including morphological segmentation, part-of-speech tagging, syntactic analysis using dependency grammar and a semantic ontology [22].

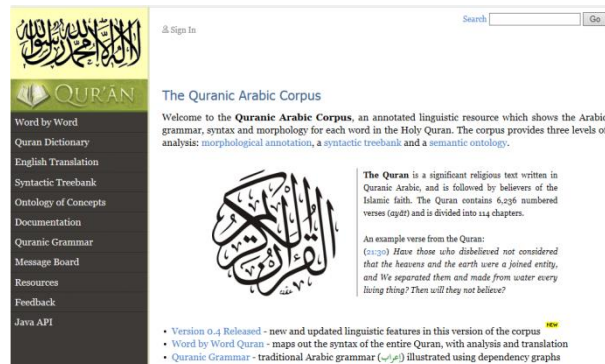


Figure 20: Screenshot of the Quranic Arabic Corpus website

4.5 Results

The AraMorph results produced 14% wrongly-tagged words and Stanford produced 32% wrongly-tagged words, while our method produced only 12% wrongly-tagged words out of all of the words in the "*The Quranic Arabic Corpus*". AraMorph tagged 46% of words as unknown words, while Stanford and our method tagged none as unknown. Finally, AraMorph produced 40% correctly-tagged words and Stanford produced 68% correctly-tagged words, while our method produced 68,501 correctly-tagged words (88%).

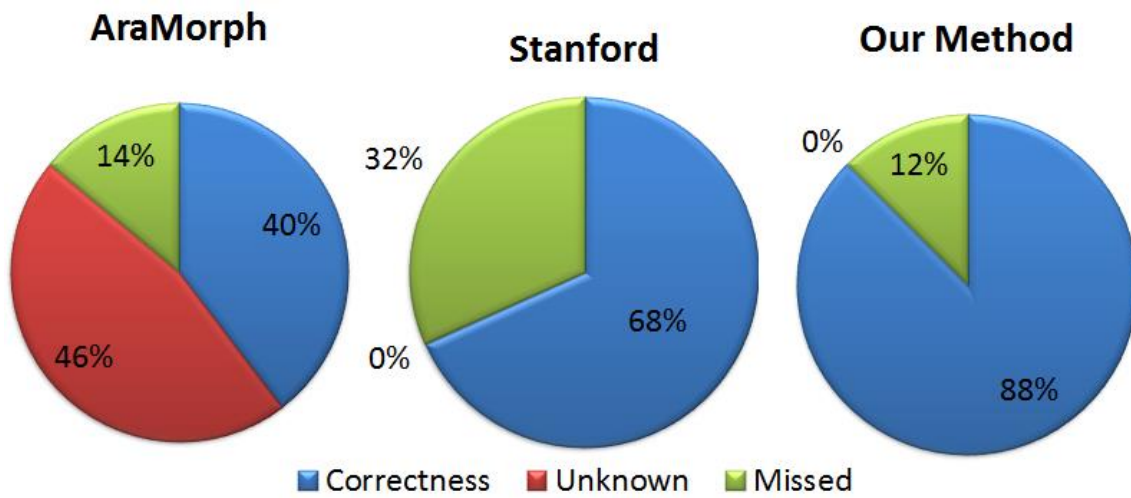


Figure 21: Proof of the validity of the methodology

	AraMorph	Stanford	Our Method
Correctness	40%	68%	88%
Unknown	46%	0%	0%
Missed	14%	32%	12%

Figure 22: The Comparison Table

4.6 Conclusions

We conclude that, because Arabic language has static verb derivations and a limited number of words, a list of all of the Verbs and Articles in Arabic language has been created. Our hypothesis states that, if the word from the text was not found in our list, it is a Noun. Therefore, Nouns are found through the elimination of Verbs and Articles.

After gathering over 112 million words (*verbs and articles*), a model was built to apply our method and test it. Our method has the highest number of correct tags compared with two other well-known Arabic POS tagging systems.

We conclude that our method, compared to AraMorph (a rule-based Arabic tagger) and the Stanford POS tagger, is far more efficient and accurate. We have proved that the simplicity of our design, with its large-scale data comparisons, is far more effective than the complexity of the other methods.

LIST OF REFERENCES

- [1] T.D. Ndie, C. Tangha, F.E. Ekwoge, MDA (Model-Driven Architecture) as a Software Industrialization Pattern: An Approach for a Pragmatic Software Factories, *Journal of Software Engineering and Applications*, 3 (2010) 561.
- [2] G.G. Chowdhury, Natural language processing, *Annual review of information science and technology*, 37 (2003) 51-89.
- [3] S. AlGahtani, W. Black, J. McNaught, Arabic part-of-speech tagging using transformation-based learning, in: *Proceedings of the 2nd International Conference on Arabic Language Resources and Tools*, Cairo, 2009, pp. 66-70.
- [4] M. Benson, Landscape: History, Present Barriers, and The Road Forward, in: *The Art of Software Thermal Management for Embedded Systems*, Springer, 2014, pp. 13-45.
- [5] B. Ba-Aziz, Arabic Language Template Grammars Component Based Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia, (2009).
- [6] D. Jurafsky, J.H. Martin, *Speech & Language Processing*, Pearson Education India, 2000.
- [7] M.R. Zughouli, A.M.i. Abu-Alshaar, English/Arabic/English machine translation: A historical perspective, *Meta: Journal des traducteursMeta:/Translators' Journal*, 50 (2005) 1022-1041.

- [8] S. Wintner, Hebrew computational linguistics: Past and future, *Artificial Intelligence Review*, 21 (2004) 113-138.
- [9] I.A. Al-Sughaiyer, I.A. Al-Kharashi, Arabic morphological analysis techniques: A comprehensive survey, *Journal of the American Society for Information Science and Technology*, 55 (2004) 189-213.
- [10] Z.A. Seikaly, *The Arabic Language: The Glue That Binds the Arab World*, AMIDEAST Publications, (2007).
- [11] W. Wright, C.P. Caspari, *A grammar of the Arabic language*, Cosimo, Inc., 2011.
- [12] M. Diab, K. Hacioglu, D. Jurafsky, Automatic tagging of Arabic text: From raw text to base phrase chunks, in: *Proceedings of HLT-NAACL 2004: Short Papers*, Association for Computational Linguistics, 2004, pp. 149-152.
- [13] M. Diab, Second generation AMIRA tools for Arabic processing: Fast and robust tokenization, POS tagging, and base phrase chunking, in: *2nd International Conference on Arabic Language Resources and Tools*, 2009.
- [14] N. Habash, O. Rambow, Arabic tokenization, part-of-speech tagging and morphological disambiguation in one fell swoop, in: *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, Association for Computational Linguistics, 2005, pp. 573-580.
- [15] S. Khoja, APT: Arabic part-of-speech tagger, in: *Proceedings of the Student Workshop at NAACL*, 2001, pp. 20-25.

- [16] K. Toutanova, D. Klein, C.D. Manning, Y. Singer, Feature-rich part-of-speech tagging with a cyclic dependency network, in: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1, Association for Computational Linguistics, 2003, pp. 173-180.
- [17] F. Al Shamsi, A. Guessoum, A hidden Markov model-based POS tagger for Arabic, Des Journées internationales d'Analyse statistique des Données Textuelles (JADT-8), Besançon, (2006) 31-42.
- [18] E. Brill, Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging, Computational linguistics, 21 (1995) 543-565.
- [19] T. Buckwalter, Arabic Morphological Analyzer (AraMorph), in, Version, 2002.
- [20] Y. Tlili-Guiassa, Hybrid Method for Tagging Arabic Text, Journal of Computer science, 2 (2006).
- [21] E. Mohamed, S. Kübler, Arabic Part of Speech Tagging, in: LREC, 2010.
- [22] K. Dukes, The Quranic Arabic Corpus, in, 2009.

Appendix A

Four-Lettered Verbs List

أدحرج	أححصص	أحرجم	أبعثر
أزعزع	أزحزح	أرفرف	أدمدم
أشماز	أشقلب	أشققش	أزلزل
أعربد	أطمأن	أطقطق	أصرصر
أكفهر	أككب	أعسكر	أعسعس
بعثرا	بعثر	أوسوس	أنقق
بعثرما	بعثرتم	بعثرتا	بعثرت
بعثروا	بعثرنا	بعثرن	بعثرتن
تبعثرن	تبعثران	تبعثر	بعثري
تخرجمان	تخرجم	تبعثرين	تبعثرون
تححصص	تخرجمين	تخرجمون	تخرجمن
تححصصين	تححصصون	تححصصن	تححصصان
تدحرجون	تدحرجن	تدحرجان	تدحرج
تدمدمن	تدمدما	تدمدم	تدحرجين
تدرفرفان	تدرفرف	تدمدمين	تدمدمون
تدحرج	تدرفرفين	تدرفرفون	تدرفرفن
تدحرجين	تدحرجون	تدحرجن	تدحرجان
تزعزعون	تزعزعن	تزعزعان	تزعزع
تزلزلن	تزلزلان	تزلزل	تزعزعين
تشققشقان	تشققشق	تزلزلين	تزلزلون
تشقلب	تشقققين	تشقققون	تشقققن
تشقلبين	تشقالبون	تشقلب	تشقالبان
تشمازون	تشمازن	تشمازان	تشماز
تصرصرن	تصرصران	تصرصر	تشمازين
تطقطقان	تطقطق	تصرصرين	تصرصرون
تطمأن	تطقطقين	تطقطقون	تطقطقن
تعربد	تطمأنين	تطمأنون	تطمأنان
تعربدين	تعربدون	تعربدن	تعربدان

تعسعون	تعسعن	تعسعان	تعسس
تعسكرن	تعسكران	تعسكر	تعسسين
تكبكان	تككب	تعسكرين	تعسكرون
تكفهر	تكبكين	تككبون	تكبكين
تكفهرين	تكفهرن	تكفهرن	تكفهران
تتقنقون	تتقنقن	تتقنقان	تتقنق
توسوسن	توسوسان	توسوس	تتقنقين
حرجما	حرجم	توسوسين	توسوسون
حرجمتما	حرجتم	حرجمتا	حرجمت
حرجموا	حرجمنا	حرجمن	حرجمتن
حصصت	حصصا	حصص	حرجمي
حصصتن	حصصتما	حصصتم	حصصتا
حصصي	حصصوا	حصصنا	حصصن
دحرجتا	دحرجت	دحرجا	دحرج
دحرجن	دحرجتن	دحرجتما	دحرجتم
دمدم	دحرجي	دحرجوا	دحرجنا
دمدمتم	دمدمتا	دمدمت	دمدما
دمدما	دمدمن	دمدمتن	دمدمتما
رفرفا	رفرف	دمدمي	دمدموا
رفرفتما	رفرفتم	رفرفتا	رفرفت
رفرفوا	رفرفنا	رفرفن	رفرفتن
زحزحت	زحزحا	زحزح	رفرفي
زحزحتن	زحزحتما	زحزحتم	زحزحتا
زحزحي	زحزحوا	زحزحنا	زحزحن
زعزعتا	زعزعت	زعزعا	زعزع
زعزعن	زعزعتن	زعزعتما	زعزعتم
زلزل	زعزعي	زعزعوا	زعزعنا
زلزلتم	زلزلتا	زلزلت	زلزلا
زلزلنا	زلزلن	زلزلتن	زلزلتما
شقشقا	شقشق	زلزلي	زلزلوا
شقشقتما	شقشقتم	شقشقتا	شقشقت

شَقَشَقْتَن	شَقَشَقْن	شَقَشَقْنَا	شَقَشَقُوا
شَقَشَقِي	شَقَلَب	شَقَلَبَا	شَقَلَبْت
شَقَلَبْنَا	شَقَلَبْتَم	شَقَلَبْتَمَا	شَقَلَبْتَن
شَقَلَبْن	شَقَلَبْنَا	شَقَلَبُوا	شَقَلَبِي
شَمَاز	شَمَازَا	شَمَازْت	شَمَازْتَا
شَمَازْتَم	شَمَازْتَمَا	شَمَازْتَن	شَمَازْن
شَمَازْنَا	شَمَازُوا	شَمَازِي	شَمَزَا
صِرَصِر	صِرَصِرَا	صِرَصِرْت	صِرَصِرْتَا
صِرَصِرْتَم	صِرَصِرْتَمَا	صِرَصِرْتَن	صِرَصِرْن
صِرَصِرْنَا	صِرَصِرُوا	صِرَصِرِي	طَقَطَق
طَقَطَقَا	طَقَطَقْت	طَقَطَقْنَا	طَقَطَقْتَم
طَقَطَقْتَمَا	طَقَطَقْتَن	طَقَطَقْن	طَقَطَقْنَا
طَقَطَقُوا	طَقَطَقِي	طَمَأْن	طَمَأْنَا
طَمَأْنَت	طَمَأْنَا	طَمَأْنْتَم	طَمَأْنْتَمَا
طَمَأْنْتَن	طَمَأْنُوا	طَمَأْنِي	طَمْنَا
عَرَبِد	عَرَبِدَا	عَرَبِدْت	عَرَبِدْتَا
عَرَبِدْتَم	عَرَبِدْتَمَا	عَرَبِدْتَن	عَرَبِدْن
عَرَبِدْنَا	عَرَبِدُوا	عَرَبِدِي	عَسْعَس
عَسْعَسَا	عَسْعَسْت	عَسْعَسْنَا	عَسْعَسْتَم
عَسْعَسْتَمَا	عَسْعَسْتَن	عَسْعَسْن	عَسْعَسْنَا
عَسْعَسُوا	عَسْعَسِي	عَسْكَر	عَسْكَرَا
عَسْكَرْت	عَسْكَرْتَا	عَسْكَرْتَم	عَسْكَرْتَمَا
عَسْكَرْتَن	عَسْكَرْن	عَسْكَرْنَا	عَسْكَرُوا
عَسْكَرِي	فَأَبْعَثَر	فَأَحْرَجَم	فَأَحْصَحَص
فَأَدْحَرَج	فَأَدْمَدَم	فَأَرْفَرَف	فَأَزْحَزَح
فَأَزْعَزَع	فَأَزْلَزَل	فَأَشْقَشَق	فَأَشْقَلَب
فَأَشْمَاز	فَأَصْرَصِر	فَأَطَقَطَق	فَأَطْمَأْن
فَأَعْرَبِد	فَأَعْسَعَس	فَأَعْسَكَر	فَأَكْبَكَب
فَأَكْفَهَر	فَأَنْقَنَق	فَأَوْسَوْس	فَبِعْثَر
فَبِعْثَرَا	فَبِعْثَرْت	فَبِعْثَرْتَا	فَبِعْثَرْتَم
فَبِعْثَرْتَمَا	فَبِعْثَرْتَن	فَبِعْثَرْن	فَبِعْثَرْنَا

فطقطقتم	فطقطقتما	فطقطقتن	فطقطقن
فطقطقنا	فطقطقوا	فطقطقي	فطمان
فطمانا	فطمانت	فطمانتا	فطمانتم
فطمانتما	فطمانتن	فطمانوا	فطماني
فطمناً	فعربد	فعربدا	فعربدت
فعربدتا	فعربدتم	فعربدتما	فعربدتن
فعربدن	فعربدنا	فعربدوا	فعربدي
فعسعس	فعسعسا	فعسعست	فعسعستا
فعسعستم	فعسعستما	فعسعستين	فعسعسن
فعسعسنا	فعسعسوا	فعسعسي	فعسكر
فعسكرا	فعسكرت	فعسكرتا	فعسكرتم
فعسكرتما	فعسكرتن	فعسكرون	فعسكرا
فعسكروا	فعسكري	فككب	فككبنا
فككببت	فككببتا	فككببتم	فككببتما
فككببتن	فككبن	فككبنا	فككبوا
فككببي	فكفهر	فكفهرا	فكفهرت
فكفهرتا	فكفهرتم	فكفهرتما	فكفهرتن
فكفهرن	فكفهرنا	فكفهروا	فكفهري
فنبعثر	فنحرجم	فنحصص	فندحرج
فندمدم	فنزرف	فنزحزح	فنززع
فنزلزل	فنشقشق	فنشقلب	فنشماز
فنصرصر	فنططق	فنطمان	فنربد
فنسعس	فنسكر	فنقنق	فنقنقا
فنقنقت	فنقنقتا	فنقنقتم	فنقنقتما
فنقنقتن	فنقنقن	فنقنقنا	فنقنقوا
فنقنقي	فنككب	فنكفهر	فنقنق
فوسوس	فوسوس	فوسوسا	فوسوست
فوسوستا	فوسوستم	فوسوستما	فوسوستن
فوسوسن	فوسوسنا	فوسوسوا	فوسوسي
فبيعثر	فبيعثران	فبيعثرن	فبيعثرون
فيحرجم	فيحرجمان	فيحرجمن	فيحرجمون

لبعثرا	لبعثرت	لبعثرتا	لبعثرتم
لبعثرتما	لبعثرتن	لبعثرن	لبعثرنا
لبعثروا	لبعثري	لبعثر	لبعثران
لتبعثرن	لتبعثرون	لتبعثرين	لتحرجم
لتحرجمان	لتحرجمن	لتحرجمون	لتحرجمين
لتحصحص	لتحصحصان	لتحصحصن	لتحصحصون
لتحصحصين	لتدحرج	لتدحرجان	لتدحرجن
لتدحرجون	لتدحرجين	لتدمدم	لتدمدمان
لتدمدمن	لتدمدمون	لتدمدمين	لتدرفرف
لتدرفرفان	لتدرفرفن	لتدرفرفون	لتدرفرفين
لتدحرجح	لتدحرجحان	لتدحرجحن	لتدحرجحون
لتدحرجحين	لتدعزع	لتدعزعان	لتدعزعن
لتدعزعون	لتدعزعين	لتدزلزل	لتدزلزلان
لتدزلزلن	لتدزلزلون	لتدزلزلين	لتدششق
لتدششقان	لتدششقن	لتدششقون	لتدششقين
لتدششقب	لتدششقلبان	لتدششقلبن	لتدششقلبون
لتدششقلبين	لتدشماز	لتدشمازان	لتدشمازان
لتدشمازون	لتدشمازين	لتدصرصر	لتدصرصران
لتدصرصرن	لتدصرصرون	لتدصرصرين	لتدطقطق
لتدطقطقان	لتدطقطقن	لتدطقطقون	لتدطقطقين
لتدطمأن	لتدطمأنان	لتدطمأنون	لتدطمأنين
لتدعربد	لتدعربدان	لتدعربدن	لتدعربدون
لتدعربدين	لتدعسعس	لتدعسعسان	لتدعسعسن
لتدعسعسون	لتدعسعسين	لتدعسكر	لتدعسكران
لتدعسكرن	لتدعسكرون	لتدعسكرين	لتدككب
لتدككبان	لتدككبن	لتدككبون	لتدككبين
لتدكفهر	لتدكفهران	لتدكفهرن	لتدكفهرون
لتدكفهرين	لتدتنقنق	لتدتنقنقان	لتدتنقنقن
لتدتنقنقون	لتدتنقنقين	لتدسوس	لتدسوسان
لتدسوسن	لتدسوسون	لتدسوسين	لحرجم
لحرجما	لحرجمت	لحرجمتا	لحرجمتم

لجرمتما	لجرمتن	لجرمن	لجرمنا
لجرموا	لجرمي	لحصص	لحصصا
لحصصت	لحصصتا	لحصصتم	لحصصتما
لحصصتن	لحصصن	لحصصنا	لحصصوا
لحصصي	لدرج	لدرجا	لدرجت
لدرجتا	لدرجتم	لدرجتما	لدرجتن
لدرجن	لدرجنا	لدرجوا	لدرجي
لدمدم	لدمدما	لدمدمت	لدمدمتا
لدمدمتم	لدمدمتا	لدمدمتن	لدمدمن
لدمدما	لدمدموا	لدمدمي	لررف
لررفا	لررفت	لررفتا	لررفتم
لررفتما	لررفتن	لررفن	لررفنا
لررفوا	لررفي	لررح	لررحا
لررحت	لررحتا	لررحتم	لررحتما
لررحتن	لررحن	لررحنا	لررحوا
لررحي	لرزع	لرزعا	لرزعت
لرزعتا	لرزعتم	لرزعتما	لرزعتن
لرزعن	لرزعنا	لرزعوا	لرزعي
لزلزل	لزلزلا	لزلزلت	لزلزلتا
لزلزلم	لزلزلتما	لزلزلتن	لزلزلن
لزلزلنا	لزلزلوا	لزلزلي	لشقشق
لشقشقا	لشقشقت	لشقشقتا	لشقشقتم
لشقشقتما	لشقشقتن	لشقشقتن	لشقشقتنا
لشقشقوا	لشقشقي	لشقلب	لشقلبا
لشقلبت	لشقلبتا	لشقلبتم	لشقلبتما
لشقلبتن	لشقلبن	لشقلبنا	لشقلبوا
لشقلبي	لشماز	لشمازا	لشمازت
لشمازتا	لشمازتم	لشمازتما	لشمازتن
لشمازن	لشمازنا	لشمازوا	لشمازي
لشمزأ	لصرصر	لصرصرا	لصرصرت
لصرصرتا	لصرصرتم	لصرصرتما	لصرصرتن

لصر صرن	لصر صرنا	لصر صروا	لصر صري
لطقق	لطققا	لطققت	لطققتا
لطققتم	لطققتما	لطققتن	لطققتن
لطققنا	لطققوا	لطققي	لطمأن
لطمأنا	لطمأنت	لطمأنتا	لطمأنتم
لطمأنتما	لطمأنتن	لطمأنوا	لطمأني
لطمناً	لعربد	لعربدا	لعربدت
لعربدتا	لعربدتم	لعربدتما	لعربدتن
لعربدن	لعربدنا	لعربدوا	لعربدي
لعرعس	لعرعسا	لعرعست	لعرعستا
لعرعستم	لعرعستما	لعرعستن	لعرعسن
لعرعسنا	لعرعسوا	لعرعسي	لعرسكر
لعرسكرا	لعرسكرت	لعرسكرتا	لعرسكرتم
لعرسكرتما	لعرسكرتن	لعرسكرن	لعرسكنا
لعرسكروا	لعرسكري	لعركب	لعركبأ
لعركببت	لعركببنا	لعركبتم	لعركببنا
لعركببتن	لعركببن	لعركبنا	لعركبوا
لعركببي	لعركهر	لعركهرا	لعركهرت
لعركهرتا	لعركهرتم	لعركهرتما	لعركهرتن
لعركهرن	لعركهنا	لعركهروا	لعركهري
لنبرعشر	لنبرجم	لنبرحصص	لنبرحرج
لنبردم	لنبرفر	لنبرحزح	لنبرعزع
لنبرزل	لنبرشق	لنبرقلب	لنبرشماز
لنبرصر	لنبرطق	لنبرطمان	لنبربد
لنبرعس	لنبرسكر	لنبرنق	لنبرنقا
لنبرنقت	لنبرنقتا	لنبرنقتم	لنبرنقتما
لنبرنقتن	لنبرنقن	لنبرنقنا	لنبرنقوا
لنبرنقي	لنبركب	لنبركه	لنبرنق
لنبروسوس	لنبروس	لنبروسا	لنبروسست
لنبروسستا	لنبروسستم	لنبروسستما	لنبروسستن
لنبروسن	لنبروسنا	لنبروسوا	لنبروسي

ليبعثرون	ليبعثرن	ليبعثران	ليبعثر
ليحرجمون	ليحرجمن	ليحرجمان	ليحرجم
ليحصحصون	ليحصحصن	ليحصحصان	ليحصحص
ليدحرجون	ليدحرجن	ليدحرجان	ليدحرج
ليدمدمون	ليدمدمن	ليدمدمان	ليدمدم
ليرفرفون	ليرفرفن	ليرفرفان	ليرفرف
ليزحزحون	ليزحزحن	ليزحزحان	ليزحزح
ليزعزعون	ليزعزعن	ليزعزعان	ليزعزع
ليزلزلون	ليزلزلن	ليزلزلان	ليزلزل
ليشقشقون	ليشقشقن	ليشقشقان	ليشقشق
ليشقلبون	ليشقلبن	ليشقلبان	ليشقلب
ليشمازون	ليشمازن	ليشمازان	ليشماز
ليصرصرون	ليصرصرن	ليصرصران	ليصرصر
ليطقطقون	ليطقطقن	ليطقطقان	ليطقطق
ليعربد	ليطمانون	ليطمانون	ليطمأن
ليعسعس	ليعربدون	ليعربدن	ليعربدان
ليعسكر	ليعسعسون	ليعسعسن	ليعسعسان
ليكبكب	ليعسكرون	ليعسكرون	ليعسكران
ليكهفر	ليكبكبون	ليكبكببن	ليكبكببن
لينقنق	ليكهفرون	ليكهفرون	ليكهفيران
ليوسوس	لينقنقون	لينقنقن	لينقنقان
نبعثر	ليوسوسون	ليوسوسن	ليوسوسان
ندمدم	ندحرج	نحصحص	نحرجم
نزلزل	نزعزع	نزحزح	نرفرف
نصرصر	نشماز	نشقلب	نششق
نعسعس	نعربد	نطمأن	نطقطق
نقنقت	نقنقا	نقنق	نعسكر
نقنقتن	نقنقتما	نقنقتم	نقنقتا
نقنقي	نقنقوا	نقنقنا	نقنقن
نوسوس	ننقنق	نكهفر	نكبكب
وأدحرج	وأحصحص	وأحرجم	وأبعثر

وأدمم	وأرفرف	وأزحزح	وأزعزع
وأزلزل	وأششق	وأشقلب	وأشماز
وأصرصر	وأططق	وأطمان	وأعربد
وأعسعس	وأعسكر	وأكبكب	وأكفهر
وأنقنق	وأوسوس	وبعثر	وبعثرا
وبعثرت	وبعثرتا	وبعثرتم	وبعثرتما
وبعثرتن	وبعثرن	وبعثرنا	وبعثروا
وبعثري	وتبعثر	وتبعثران	وتبعثرن
وتبعثرون	وتبعثرين	وتحرجم	وتحرجمان
وتحرجمن	وتحرجمون	وتحرجمين	وتحصحص
وتحصحصان	وتحصحصن	وتحصحصون	وتحصحصين
وتدحرج	وتدحرجان	وتدحرجن	وتدحرجون
وتدحرجين	وتدمدم	وتدمدمان	وتدمدمن
وتدمدمون	وتدمدمين	وترفررف	وترفررفان
وترفررفن	وترفررفون	وترفررفين	وترحزح
وترحزحان	وترحزحن	وترحزحون	وترحزحين
وترعزع	وترعزعان	وترعزعن	وترعزعون
وترعزعين	وترلزل	وترلزلان	وترلزلن
وترلزلون	وترلزلين	وتشقشق	وتشقشقان
وتشقشقين	وتشقشقون	وتشقشقين	وتشقشقلب
وتشقشلبان	وتشقشلبين	وتشقشلبون	وتشقشلبين
وتشماز	وتشمازان	وتشمازن	وتشمازون
وتشمازين	وتصرصر	وتصرصران	وتصرصرن
وتصرصرون	وتصرصرين	وتططق	وتططقان
وتططقن	وتططقون	وتطططين	وتطمأن
وتطمأنان	وتطمأنون	وتطمأنين	وتعربد
وتعربدان	وتعربدن	وتعربدون	وتعربدين
وتعسعس	وتعسعسان	وتعسعسن	وتعسعسون
وتعسعسين	وتعسكر	وتعسكران	وتعسكرن
وتعسكرون	وتعسكرين	وتكبكب	وتكبكبان
وتكبكين	وتكبكبون	وتكبكبين	وتكفهر

وتكفهران	وتكفهرن	وتكفهرين	وتكفهرين
وتتقنق	وتتقنقن	وتتقنقان	وتتقنقون
وتتقنقين	وتوسوسان	وتوسوس	وتوسوسن
وتوسوسون	وحرجم	وتوسوسين	وحرجما
وحرجمت	وحرجمتم	وحرجمتا	وحرجمتما
وحرجمتن	وحرجمنا	وحرجمن	وحرجموا
وحرجمي	وحصصا	وحصص	وحصصت
وحصصتا	وحصصتما	وحصصتم	وحصصتن
وحصصن	وحصصوا	وحصصنا	وحصصي
ودحرج	ودحرجت	ودحرجا	ودحرجتا
ودحرجتم	ودحرجتن	ودحرجتما	ودحرجن
ودحرجنا	ودحرجي	ودحرجوا	ودمدم
ودمدما	ودمدمتا	ودمدمت	ودمدتم
ودمدمتما	ودمدمن	ودمدمتن	ودمدمنا
ودمدموا	ورفرف	ودمدمي	ورفرفا
ورفرفت	ورفرقتم	ورفرفتا	ورفرقتما
ورفرفتن	ورفرفنا	ورفرفن	ورفرفوا
ورفرفي	وزحزح	وزحزح	وزحزحت
وزحزحتا	وزحزحتما	وزحزحتم	وزحزحتن
وزحزحن	وزحزحوا	وزحزحنا	وزحزحي
وزعزع	وزعزعت	وزعزعا	وزعزعتا
وزعزعتم	وزعزعتن	وزعزعتما	وزعزعن
وزعزعنا	وزعزعي	وزعزعوا	وززلزل
وززلزلا	وززلزلنا	وززلزلت	وززلزلتم
وززلزلتما	وززلزلن	وززلزلتن	وززلزلنا
وززلزلوا	وسوس	وزلزلي	وسوسا
وسوست	وسوستم	وسوستا	وسوستما
وسوستن	وسوسنا	وسوسن	وسوسوا
وسوسي	وشقشقا	وشقشق	وشقشقت
وشقشقتا	وشقشقتما	وشقشقتم	وشقشقتن
وشقشقتن	وشقشقوا	وشقشقنا	وشقشقي

وشقلاب	وشقلابا	وشقلابت	وشقلابتا
وشقلابتم	وشقلابتما	وشقلابتن	وشقلابتین
وشقلابنا	وشقلابوا	وشقلابي	وشماز
وشمازا	وشمازت	وشمازتا	وشمازتم
وشمازتما	وشمازتن	وشمازن	وشمازنا
وشمازوا	وشمازي	وشمزأ	وصرصر
وصرصرا	وصرصرت	وصرصرتا	وصرصرتم
وصرصرتما	وصرصرتن	وصرصرن	وصرصرنا
وصرصروا	وصرصري	وطقطق	وطقطقا
وطقطقت	وطقطقتا	وطقطقتم	وطقطقتما
وطقطقتن	وطقطقن	وطقطقنا	وطقطقوا
وطقطقي	وطمان	وطمانا	وطمانت
وطمانتا	وطمانتم	وطمانتما	وطمانتن
وطمانوا	وطماني	وطمنأ	وعربد
وعربدا	وعربدت	وعربدتا	وعربدتم
وعربدتما	وعربدتن	وعربدن	وعربدنا
وعربدوا	وعربدي	وعسعا	وعسعسا
وعسعست	وعسعستا	وعسعستم	وعسعستما
وعسعستن	وعسعسن	وعسعسنا	وعسعسوا
وعسعسي	وعسكر	وعسكرا	وعسكرت
وعسكرتا	وعسكرتم	وعسكرتما	وعسكرتن
وعسكرن	وعسكرنا	وعسكروا	وعسكري
وككبب	وككببا	وككببت	وككببتا
وككببتم	وككببتما	وككببتن	وككببن
وككببنا	وككببوا	وككببي	وككفر
وككفرا	وككفرت	وككفرتا	وككفرتم
وككفرتما	وككفرتن	وككفرن	وككفرنا
وككفروا	وككفري	ونبعثر	ونحرجم
ونححصص	وندحرج	وندمدم	ونررفرف
ونزحزح	ونزعزع	ونزلزل	ونشقشق
ونشقلب	ونشماز	ونصرصر	ونطقطق

ونطمأن	ونعربد	ونعسعس	ونعسكر
ونقنق	ونقنقا	ونقنقت	ونقنقتا
ونقنقتم	ونقنقتما	ونقنقن	ونقنقن
ونقنقنا	ونقنقوا	ونقنقي	ونكبكب
ونكفهر	وننقنق	ونوسوس	ووسوس
ووسوسا	ووسوست	ووسوستا	ووسوستم
ووسوستما	ووسوستن	ووسوسن	ووسوسنا
ووسوسوا	ووسوسي	ويبعثر	ويبعثران
ويبعثرن	ويبعثرون	ويحرجم	ويحرجمان
ويحرجمن	ويحرجمون	ويحصحص	ويحصحصان
ويحصحصن	ويحصحصون	ويدحرج	ويدحرجان
ويدحرجن	ويدحرجون	ويدمدم	ويدمدمان
ويدمدمن	ويدمدمون	ويرفرف	ويرفرفان
ويرفرفن	ويرفرفون	ويزحزح	ويزحزحان
ويزحزحن	ويزحزحون	ويزعزع	ويزعزعان
ويزعزعن	ويزعزعون	ويزلزل	ويزلزلان
ويزلزلن	ويزلزلون	ويشققشقق	ويشققشقان
ويشققشققن	ويشققشقون	ويشقلاب	ويشقلابان
ويشقلابن	ويشقلابون	ويشمازان	ويشمازان
ويشمازان	ويشمازون	ويصرصر	ويصرصران
ويصرصرن	ويصرصرون	ويطقطق	ويطقطقان
ويطقطقن	ويطقطقون	ويطمان	ويطمانان
ويطمانون	ويعربد	ويعربدان	ويعربدن
ويعربدون	ويعسعس	ويعسعسان	ويعسعسن
ويعسعسون	ويعسكر	ويعسكران	ويعسكرن
ويعسكرون	ويكبكب	ويكبكبان	ويكبكبن
ويكبكبون	ويكفهر	ويكفهران	ويكفهرن
ويكفهرون	وينقنق	وينقنقان	وينقنقن
وينقنقون	ويوسوس	ويوسوسان	ويوسوسن
ويوسوسون	ييعثر	ييعثران	ييعثرن
ييعثرون	يحرجم	يحرجمان	يحرجمن

يحصص	يحصصان	يحصص	يحر جمون
يدحصون	يدحرجان	يدحرج	يحصصون
يدمدمون	يدمدمان	يدمدم	يدحرجون
يرفرمون	يرفرفان	يرفرف	يدمدمون
يرفرفون	يرفرحان	يرفرح	يرفرفون
يزحزون	يزعزان	يزعزع	يزحزون
يزلزون	يزلزلان	يزلزل	يزعزون
يشقشقون	يشقشقان	يشقشق	يزلزون
يشقشقبون	يشقشقبان	يشقشقب	يشقشقون
يشمازون	يشمازان	يشماز	يشقشقبون
يصرصرون	يصرصران	يصرصر	يشمازون
يطقطقون	يطقطقان	يطقطق	يصرصرون
يعربدون	يعربدان	يعربدان	يطقطقون
يعسعون	يعسعون	يعسعون	يعربدون
يعسكرون	يعسكرون	يعسكرون	يعسعون
يكبكبون	يكبكبان	يكبكب	يعسكرون
يكفهرون	يكفهران	يكفهران	يكبكبون
ينقنقون	ينقنقان	ينقنقان	يكفهرون
يوسوسون	يوسوسان	يوسوسان	ينقنقون

Appendix B

The Articles List

اتخذ	اتخذوا	اتخذت	اتخذتم
اتخذت	اتخذت	اتخذت	اتخذتم
اتخذوا	اتخذوا	اتخذوا	اتخذتم
اللاتي	اللاتي	اللاتي	اللاتي
انفك	انفك	انفك	انفك
انفككت	انفككت	انفككت	انفككت
انفككتكم	انفككتكم	انفككتكم	انفككتكم
انفككتنا	انفككتنا	انفككتنا	انفككتنا
إذا	إذا	إذا	إذا
إذما	إذما	إذما	إذما
إذماكن	إذماكن	إذماكن	إذماكن
إذماهم	إذماهم	إذماهم	إذماهم
إذن	إذن	إذن	إذن
إلا	إلا	إلا	إلا
إلاكن	إلاكن	إلاكن	إلاكن
إلاهم	إلاهم	إلاهم	إلاهم
إلاي	إلاي	إلاي	إلاي
إليكم	إليكم	إليكم	إليكم
إليه	إليه	إليه	إليه
إليها	إليها	إليها	إليها
إيا	إيا	إيا	إيا
إياكن	إياكن	إياكن	إياكن
إياها	إياها	إياها	إياها
إياي	إياي	إياي	إياي
أصبح	أصبح	أصبح	أصبح
أصبحتم	أصبحتم	أصبحتم	أصبحتم
أصبحنا	أصبحنا	أصبحنا	أصبحنا
أضحيت	أضحيت	أضحيت	أضحيت
أضحيتكم	أضحيتكم	أضحيتكم	أضحيتكم
أضحيتنا	أضحيتنا	أضحيتنا	أضحيتنا
أضحوا	أضحوا	أضحوا	أضحوا
ألكم	ألكم	ألكم	ألكم
أله	أله	أله	أله
ألهما	ألهما	ألهما	ألهما
أم	أم	أم	أم
أمسى	أمسى	أمسى	أمسى

أمسيت	أمسيتم	أمسيتما	أمسيتين
أمسينا	أن	أنا	أنت
أنتم	أنتما	أنتن	أنك
أنكم	أنكما	أنكن	أننا
أنه	أنها	أنهم	أنهما
أنهن	أني	أو	أوشك
أوشكا	أوشكت	أوشكتم	أوشكتما
أوشكتن	أوشكن	أوشكنا	أوشكوا
أولاء	أولاءك	أولاءكم	أولاءكما
أولاءكن	أولاءنا	أولاءه	أولاءها
أولاءهم	أولاءهما	أولاءهن	أولاءي
أي	أيك	أيكم	أيكما
أيكن	أينا	أيه	أيها
أيهم	أيهما	أيهن	أيي
آ	بات	باتا	باتت
باتوا	بت	بتم	بتما
بتن	بتنا	برح	برحا
برحت	برحتم	برحتما	برحتن
برحن	برحنا	برحوا	بل
بلى	تان	تانك	تانكم
تانكما	تانكن	تاننا	تانه
تانها	تانهم	تانهما	تانهن
تاني	ته	تهك	تهكم
تهكما	تهكن	تهنا	تهه
تهها	تههم	تههما	تههن
تهي	تي	تيك	تيكم
تيكما	تيكن	تين	تينا
تينك	تينكم	تينكما	تينكن
تيننا	تينه	تينها	تينهم
تينهما	تينهن	تيني	تیه
تيها	تيهم	تيهما	تيهن

ثمي	ثم	ثمك	ثمكم
ثمكما	ثمكن	ثمنا	ثمه
ثمها	ثمهم	ثمهما	ثمهن
ثمي	جعل	جعلنا	جعلت
جعلتم	جعلتما	جعلتن	جعلن
جعلنا	جعلوا	جلل	جير
جيرك	جيركم	جيركما	جيركن
جيرنا	جيره	جيرها	جيرهم
جيرهما	جيرهن	جيرى	حاشا
حاشاك	حاشاكم	حاشاكما	حاشاكن
حاشانا	حاشاه	حاشاها	حاشاهم
حاشاهما	حاشاهن	حاشاي	حتى
حسب	حسبك	حسبكم	حسبكما
حسبكن	حسبنا	حسبه	حسبها
حسبهم	حسبهما	حسبهن	حسبي
خلا	خلاك	خلاككم	خلاكما
خلاكن	خلانا	خلاه	خلاها
خلاههم	خلاهنا	خلاههن	خلاي
دام	داما	دامت	داموا
دمت	دمتم	دمتما	دمتن
دمن	دمنا	ذا	ذاك
ذاكم	ذاكما	ذاكن	ذان
ذانا	ذانك	ذانكم	ذانكما
ذانكن	ذاننا	ذانه	ذانهما
ذانهم	ذانهما	ذانهن	ذاني
ذاه	ذاها	ذاهم	ذاهما
ذاهن	ذاي	ذه	ذهك
ذهكم	ذهكما	ذهكن	ذهنا
ذهه	ذهها	ذههم	ذههما
ذههن	ذهي	ذي	ذيك
ذيكم	ذيكما	ذيكن	ذين

ذينا	ذينا	ذينا	ذينا
ذينكن	ذينه	ذيننا	ذينها
ذينهم	ذينهين	ذينهنا	ذيني
ذيه	ذيهم	ذيهنا	ذيهما
ذيهين	رام	ذيهي	راما
رامت	رأوا	راموا	رأى
رأيا	رأيتم	رأيت	رأيتما
رأيتن	رأينا	رأين	رب
ربك	ربكما	ربكم	ربكن
ربنا	ربها	ربه	ربهم
ربهما	ربي	ربهن	رمت
رمتم	رمتن	رمتما	رمن
رمتنا	زالا	زال	زالت
زالوا	زعا	زعم	زعمت
زعمتم	زعمتن	زعمتما	زعمن
زعمنا	زلت	زعموا	زلتم
زلتما	زلن	زلتن	زلنا
سوف	صارا	صار	صارت
صاروا	صرتم	صرت	صرتما
صرتن	صرنا	صرن	ظل
ظلا	ظلت	ظلت	ظلتتم
ظلتما	ظللن	ظللتن	ظلنا
ظلوا	ظنا	ظن	ظنت
ظنن	ظننت	ظننا	ظننتم
ظننتما	ظنوا	ظننتن	عدا
عداك	عداكما	عداكم	عداكن
عدانا	عداها	عداه	عداهم
عداهما	عداي	عداهن	عسى
عسيا	عسيتم	عسيتم	عسيتما
عسيتين	عسينا	عسين	عسيوا
عل	علكم	علك	علكما

علمت	علما	علم	علكن
علمن	علمتن	علمتما	علمتم
عله	علنا	علموا	علمنا
علهن	علهما	علهم	علها
عليكم	عليك	علي	على
عليه	علينا	عليكن	عليكما
عليهن	عليهما	عليهم	عليها
عنكما	عنكم	عنك	عن
عنها	عنه	عننا	عنكن
عني	عنهن	عنهما	عنهم
فاتخذتم	فاتخذت	فاتخذنا	فاتخذ
فاتخذنا	فاتخذن	فاتخذتن	فاتخذتما
فالذين	فالذي	فالتي	فاتخذوا
فاللذان	فالتان	فالتائي	فالتاتي
فانفككت	فانفكت	فانفكا	فانفك
فانفككن	فانفككتن	فانفككتما	فانفككتم
فإذا	فإذ	فانفكوا	فانفككنا
فإذماكما	فإذماكم	فإذماك	فإذما
فإذماها	فإذماه	فإذمانا	فإذماكن
فإذماي	فإذماهن	فإذماهما	فإذماهم
فإلاكم	فإلاك	فإلا	فإذن
فإلاه	فإلانا	فإلاكن	فإلاكما
فإلاهن	فإلاهما	فإلاهم	فإلاها
فإليكم	فإليك	فإلي	فإلاي
فإليه	فإلينا	فإليكن	فإليكما
فإليهن	فإليهما	فإليهم	فإليها
فإياكم	فإياك	فإيا	فإما
فإياه	فإيانا	فإياكن	فإياكما
فإياهن	فإياهما	فإياهم	فإياها
فأصبحا	فأصبح	فأجل	فإياي
فأصبحتن	فأصبحتما	فأصبحتم	فأصبحت

فأصبحن	فأصبحنا	فأصبحوا	فأضحى
فأضحيا	فأضحيت	فأضحيتم	فأضحيتما
فأضحيتين	فأضحين	فأضحينا	فأضحوا
فأل	فألك	فألكم	فألكما
فالكن	فالنا	فالها	فالها
فالهم	فالهما	فالهن	فأم
فأمست	فأمسو	فأمسى	فأمسيا
فأمسيت	فأمسيتم	فأمسيتما	فأمسيتين
فأمسينا	فأن	فأنا	فأنت
فأنتم	فأنتما	فأنتن	فأنك
فأنكم	فأنكما	فأنكن	فأننا
فأنه	فأنها	فأنهم	فأنهما
فأنهن	فأني	فأو	فأوشك
فأوشكا	فأوشكت	فأوشكتم	فأوشكتما
فأوشكتن	فأوشكن	فأوشكنا	فأوشكوا
فأولاء	فأولاءك	فأولاءكم	فأولاءكما
فأولاءكن	فأولاءنا	فأولاءه	فأولاءها
فأولاءهم	فأولاءهما	فأولاءهن	فأولاءي
فأي	فأيك	فأيكم	فأيكما
فأيكن	فأينا	فأيه	فأيها
فأيهم	فأيهما	فأيهن	فأيي
فآ	فبات	فباتا	فباتت
فباتوا	فبت	فبتم	فبتما
فبتن	فبتنا	فبرح	فبرحا
فبرحت	فبرحتم	فبرحتما	فبرحتن
فبرحن	فبرحنا	فبرحوا	فبل
فبلى	فتان	فتانك	فتانكم
فتانكما	فتانكن	فتاننا	فتانه
فتانها	فتانهم	فتانها	فتانهن
فتاني	فته	فتهك	فتهكم
فتهكما	فتهكن	فتهنا	فتهه

فتيها	فتيهم	فتيها	فتيها
فتي	فتي	فتي	فتيكم
فتيكن	فتين	فتينا	فتيكن
فتينكم	فتينكما	فتينكن	فتينكن
فتينه	فتينها	فتينهم	فتيننا
فتينهن	فتيني	فتيه	فتينهما
فتيهم	فتيها	فتيهن	فتيها
فتى	فتئا	فتئت	فتيي
فتئتما	فتئتن	فتئن	فتئتم
فتئوا	فثم	فثمك	فتئنا
فثمكما	فثمكن	فثمنا	فثمكم
فثمها	فثمهم	فثمها	فثمه
فثمي	فجعل	فجعلنا	فثمهن
فجعلتم	فجعلتما	فجعلتن	فجعلت
فجعلنا	فجعلوا	فجلل	فجعلن
فجيرك	فجيركم	فجيركما	فجير
فجيرنا	فجير ه	فجيرها	فجيركن
فجيرهما	فجيرهن	فجيري	فجيرهم
فحاشاك	فحاشاكم	فحاشاكما	فحاشا
فحاشانا	فحاشاه	فحاشاها	فحاشاكن
فحاشاهما	فحاشاهن	فحاشاي	فحاشاهم
فحسب	فحسبك	فحسبكم	فحتى
فحسبكن	فحسبنا	فحسبه	فحسبكما
فحسبهم	فحسبهما	فحسبهن	فحسبها
فخلا	فخلاك	فخلاكم	فحسبي
فخلاكن	فخلانا	فخلاه	فخلاكما
فخلاهم	فخلاهما	فخلاهن	فخلاهما
فدام	فداما	فدامت	فخلاي
فدمت	فدمتم	فدمتما	فداموا
فدمن	فدمننا	فذا	فدمتن
فذاكم	فذاكما	فذاكن	فذاك

فذان	فذاننا	فذانك	فذانكم
فذانكما	فذانكن	فذاننا	فذاننه
فذانها	فذانهم	فذانهما	فذانهن
فذاني	فذاه	فذاها	فذاهم
فذاهما	فذاهن	فذاي	فذه
فذهك	فذهكم	فذهكما	فذهكن
فذهنا	فذهه	فذهها	فذههم
فذهما	فذهن	فذهي	فذي
فذيك	فذيكم	فذيكما	فذيكن
فذين	فذيننا	فذينك	فذينكم
فذينكما	فذينكن	فذيننا	فدينه
فذينها	فذينهم	فذينهما	فدينهن
فذيي	فذيه	فذيها	فذيهم
فذيهما	فذهين	فذيي	فرا م
فرا ما	فرا مت	فرا موا	فرا و ا
فراى	فرا يا	فرا يت	فرا يتم
فرا يتما	فرا يتن	فرا ين	فرا ينا
فرب	فربك	فربكم	فربكما
فربكن	فربنا	فربه	فربها
فربهم	فربهما	فربهن	فربي
فرمت	فرمتم	فرمتما	فرمتن
فرمن	فرمنا	فرال	فرالا
فزالت	فزالوا	فز عم	فز عما
فز عمت	فز عتم	فز عمتما	فز عمتن
فز عمن	فز عمنا	فز عموا	فز لت
فز لتم	فز لتما	فز لتن	فز لن
فز لنا	فسوف	فصار	فصارا
فصارت	فصاروا	فصرت	فصرتم
فصرتما	فصرتن	فصرن	فصرنا
فظل	فظلا	فظلت	فظلت
فظللتم	فظللتما	فظللتن	فظللن

فظنا	فظن	فظلوا	فظلنا
فظننت	فظننا	فظنن	فظننت
فظنوا	فظننن	فظننتما	فظننتم
فعداكما	فعداكم	فعداك	فعدا
فعداها	فعداه	فعدانا	فعداكن
فعداي	فعداهن	فعداهما	فعداهم
فعسيتم	فعسيت	فعسيا	فعسى
فعسينا	فعسين	فعسينن	فعسينتما
فعلكم	فعلك	فعل	فعسيوا
فعلما	فعلم	فعلكن	فعلكما
فعلمنن	فعلمتما	فعلمتم	فعلمت
فعلنا	فعلموا	فعلمنا	فعلمن
فعلهما	فعلهم	فعلها	فعله
فعليك	فعلي	فعلى	فعلهن
فعلينا	فعلين	فعليكما	فعليكم
فعليهما	فعليهم	فعليهما	فعليه
فعنكم	فعنك	فعن	فعليهن
فعنه	فعننا	فعنكن	فعنكما
فعنهن	فعنهما	فعنهم	فعنها
ففتنت	ففتنا	ففتنى	فعني
ففتنن	ففتننن	ففتنتما	ففتنتم
ففيك	ففي	ففتنوا	ففتننا
ففينا	ففين	ففيكما	ففيكم
ففيهما	ففيهم	ففيها	ففيه
فكاد	فقد	ففيي	ففيهن
فكان	فكادوا	فكادت	فكادا
فكانك	فكانوا	فكانت	فكانا
فكاننا	فكانن	فكانكما	فكانكم
فكانهما	فكانهم	فكانها	فكانه
فكدتم	فكدت	فكاني	فكانهن
فكدنا	فكدن	فكدتن	فكدتما

فكر بتم	فكر بت	فكر با	فكر ب
فكر بنا	فكر بن	فكر بتن	فكر بتما
فكنا	فكن	فكلا	فكربوا
فكنتن	فكنتما	فكنتم	فكننت
فلست	فلات	فلا	فكي
فلسنا	فلستن	فلستما	فلستم
فلعلكما	فلعلكم	فلعلك	فلعل
فلعلها	فلعله	فلعلنا	فلعلكن
فلعلي	فلعلهن	فلعلهما	فلعلهم
فلكنكما	فلكنكم	فلكنك	فلكن
فلكنها	فلكنه	فلكننا	فلكنكن
فلكني	فلكنهن	فلكنهما	فلكنهم
فلو	فلن	فلما	فلم
فلولاكما	فلولاكم	فلولاك	فلولا
فلولاها	فلولاه	فلولانا	فلولاكن
فلولاي	فلولاهن	فلولاهما	فلولاهم
فلوماكما	فلوماكم	فلوماك	فلوما
فلوماها	فلوماه	فلوماننا	فلوماكن
فلوماي	فلوماهن	فلوماهما	فلوماهم
فليتكما	فليتكم	فليتك	فليت
فليتها	فليته	فليتنا	فليتكن
فليتي	فليتهن	فليتهما	فليتههم
فما	فليسوا	فليس	فليس
فمك	فمذ	فمن	فمذ
فمننا	فمكن	فمنكما	فمنكم
فمنها	فمنهم	فمنها	فمنه
فنعم	فنحن	فمني	فمنهن
فهاكما	فهاكم	فهاك	فها
فهاها	فهاه	فهاننا	فهاكن
فهاي	فهاهن	فهاهما	فهاهم
فهاكم	فهاك	فها	فهل

فهاكما	فهاكن	فهلانا	فهللاه
فهلها	فهلهم	فهلها	فهلها
فهللي	فهلك	فهلكم	فهلكما
فهلكن	فهلنا	فهلله	فهلها
فهلهم	فهلها	فهلهم	فهللي
فهم	فهما	فهن	فهننا
فهنالك	فهنالك	فهنالكما	فهنالك
فهنانا	فهنانه	فهنانا	فهنانهم
فهنانها	فهنانهم	فهناني	فهو
فهيا	فهيا	فوا	فوالتي
فو الذي	فو الذين	فو اللاتي	فو اللاتي
فواللتان	فواللذان	فوجد	فوجدنا
فوجدت	فوجدتم	فوجدتما	فوجدتن
فوجدن	فوجدنا	فوجدوا	في
فيا	فيك	فيكم	فيكما
فيكن	فيها	فيه	فيها
فيهم	فيهما	فيهن	فيها
قد	كاد	كادا	كادت
كادوا	كان	كانا	كانت
كانوا	كانك	كانكم	كانكما
كانكن	كاننا	كانه	كانها
كانهم	كانهما	كانهن	كاني
كدت	كدتم	كدتما	كدتن
كدن	كدنا	كرب	كربا
كربت	كربتم	كربتما	كربتن
كرين	كرينا	كربوا	كلا
كن	كنا	كنت	كنتم
كنتما	كنتن	كي	لا
لات	لاتخذ	لاتخذا	لاتخذت
لاتخذتم	لاتخذتما	لاتخذتن	لاتخذن
لاتخذنا	لاتخذوا	لانفك	لانفكا

لانفكت	لانفكت	لانفكت	لانفكتما
لانفكتن	لانفكتنا	لانفكتن	لانفكتوا
لإذ	لإذما	لإذا	لإذماك
لإذماكم	لإذماكن	لإذماكما	لإذمانا
لإذماه	لإذماهم	لإذماها	لإذماههما
لإذماهن	لإذن	لإذماي	لإلا
لإلاك	لإلاكما	لإلاكم	لإلاكن
لإلانا	لإلاها	لإلاه	لإلاهم
لإلاهما	لإلاي	لإلاهن	لإلي
لإليك	لإليكما	لإليكم	لإليكن
لإلينا	لإليها	لإليه	لإليهم
لإليهما	لإما	لإليهن	لإيا
لإياك	لإياكما	لإياكم	لإياكن
لإيانا	لإياها	لإياه	لإياهم
لإياهما	لإياي	لإياهن	لأجل
لأصبح	لأصبحت	لأصباحا	لأصباحتم
لأصبحتما	لأصبحن	لأصبحتن	لأصبحنا
لأصبحوا	لأضحيا	لأضحى	لأضحيت
لأضحيتم	لأضحيتن	لأضحيتما	لأضحين
لأضحينا	لأل	لأضحوا	لألك
لألكم	لألكن	لألكما	لألنا
لأله	لألهم	لألها	لألها
لألهن	لأمت	لأم	لأمسو
لأمسى	لأمسيت	لأمسيا	لأمسيتم
لأمسيتما	لأمسينا	لأمسيتين	لأن
لأنا	لأنتم	لأننت	لأنتما
لأنتن	لأنكم	لأنك	لأنكما
لأنكن	لأنه	لأننا	لأنها
لأنهم	لأنهن	لأنهما	لأنى
لأو	لأوشكا	لأوشك	لأوشكت
لأوشكتم	لأوشكتن	لأوشكتما	لأوشكن

لأولاءك	لأولاء	لأولئكوا	لأولئكنا
لأولاءنا	لأولاءكن	لأولاءكما	لأولاءكم
لأولاءهما	لأولاءهم	لأولاءها	لأولاءه
لأليك	لأي	لأولاءي	لأولاءهن
لأينا	لأيكُن	لأيكما	لأيكُم
لأيهما	لأيهم	لأيهها	لأيهه
لباتا	لبات	لأبي	لأيهن
لبتم	لبت	لباتوا	لباتت
لبرح	لبتنا	لبتن	لبتما
لبرحما	لبرحتم	لبرحت	لبرحا
لبرحوا	لبرحنا	لبرحن	لبرحتن
لتانك	لتان	لبلى	لبل
لتاننا	لتانكن	لتانكما	لتانكم
لتانهما	لتانهم	لتانهها	لتانهه
لتهاك	لته	لتاني	لتانهن
لتهنا	لتهكن	لتهكما	لتهكم
لتهمها	لتهم	لتهاها	لتهه
لتيك	لتي	لتهي	لتههن
لتين	لتیکن	لتيكما	لتيكُم
لتينكما	لتينكم	لتينك	لتينا
لتينها	لتينه	لتيننا	لتينكن
لتيني	لتينهِن	لتينهما	لتينهَم
لتيهما	لتيهم	لتيهها	لتيهه
لثمك	لثم	لتيي	لتيهن
لثما	لثمكن	لثكما	لثكم
لثهما	لثهم	لثهاها	لثمه
لثعلا	لثعل	لثمي	لثمهن
لثعلتن	لثعلتما	لثعلتم	لثعلت
لجل	لجعلوا	لثعلنا	لثعلن
لجبركما	لجبركم	لجبرك	لجبر
لجبرها	لجبره	لجبرنا	لجبركن

لجيري	لجيرهن	لجيرهما	لجيرهم
لحاشاكما	لحاشاكم	لحاشاك	لحاشا
لحاشاها	لحاشاه	لحاشانا	لحاشاكن
لحاشاي	لحاشاهن	لحاشاهما	لحاشاهم
لحسبكم	لحسبك	لحسب	لحتى
لحسبه	لحسبنا	لحسبكن	لحسبكما
لحسبهن	لحسبهما	لحسبهم	لحسبها
لخلاكم	لخلاك	لخلا	لحسبي
لخلاه	لخلانا	لخلاكن	لخلاكما
لخلاهن	لخلاهما	لخلاهم	لخلاها
لدامت	لداما	لدام	لخلاي
لدمتما	لدمتم	لدمت	لداموا
لذا	لدمنا	لدمن	لدمتن
لذاكن	لذاكما	لذاكم	لذاك
لذاكنم	لذاذك	لذانا	لذان
لذانه	لذاننا	لذانكن	لذانكما
لذانهن	لذانهما	لذانهم	لذانها
لذاهم	لذاها	لذاه	لذاني
لذه	لذاي	لذاهن	لذاهما
لذهكن	لذهكما	لذهكم	لذهك
لذههم	لذهها	لذهه	لذهنا
لذي	لذهي	لذههن	لذههما
لذيكن	لذيكما	لذيكم	لذيك
لذينكم	لذينك	لذينا	لذين
لذينه	لذيننا	لذينكن	لذينكما
لذينهن	لذينهما	لذينهم	لذينها
لذيههم	لذيهها	لذيه	لذيبي
لرام	لذيي	لذينهن	لذيهما
لرأوا	لراموا	لرامت	لراما
لرأيتم	لرأيت	لرأيا	لرأى
لرأينا	لرأين	لرأيتن	لرأيتما

لرب	لربك	لربكم	لربكما
لربيكن	لربنا	لربه	لربها
لربهم	لربهما	لربيهن	لربي
لرمت	لرمتم	لرمتما	لرمتن
لرمن	لرمننا	لزال	لزالا
لزالت	لزالوا	لزعم	لزعما
لزعمت	لزعمتم	لزعمتما	لزعمتن
لزعمن	لزعمننا	لزعموا	لزلت
لزلتم	لزلتمنا	لزلتن	لزلن
لزلنا	لست	لستم	لستما
لستن	لشنا	لسوف	لصار
لصارا	لصارت	لصاروا	لصرت
لصرتم	لصرتما	لصرتن	لصرن
لصرنا	لظل	لظلا	لظلت
لظللت	لظللتم	لظللتما	لظللتن
لظللن	لظلنا	لظلوا	لظن
لظنا	لظنت	لظنن	لظننا
لظننت	لظننتم	لظننتما	لظننتن
لظنوا	لعدا	لعداك	لعداكم
لعداكما	لعداكن	لعدانا	لعداه
لعداها	لعداهم	لعداهما	لعداهن
لعداي	لعسى	لعسيا	لعسيت
لعسيتم	لعسيتما	لعسيتن	لعسين
لعسينا	لعسيوا	لعل	لعاك
لعلكم	لعلكما	لعلكن	لعلم
لعلما	لعلمت	لعلتم	لعلمتما
لعلمتن	لعلمن	لعلمنا	لعلموا
لعلنا	لعله	لعلها	لعلهم
لعلهما	لعلهن	لعلى	لعلي
لعليك	لعليكم	لعليكما	لعليكن
لعلينا	لعليه	لعليها	لعليهم

لعليهما	لعليهن	لعن	لعنك
لعنكم	لعنكما	لعنكن	لعننا
لعنه	لعنها	لعنهم	لعنهما
لعنهن	لعني	لعتي	لعتنا
لعتت	لعتتم	لعتتما	لعتتن
لعتن	لعتنا	لعتوا	لعي
لعيك	لعيكم	لعيكما	لعيكن
لعيها	لعيه	لعيها	لعيهم
لعيهما	لعيهن	لعيي	لقد
لكاد	لكادا	لكادت	لكادوا
لكان	لكانا	لكانت	لكانوا
لكأنك	لكأنكم	لكأنكما	لكأنكن
لكأننا	لكأنه	لكأنها	لكأنهم
لكأنهما	لكأنهن	لكأني	لكدت
لكدتم	لكدتما	لكدتن	لكدن
لكدنا	لكرب	لكربا	لكربت
لكربتم	لكربتما	لكربتن	لكربن
لكربنا	لكربوا	لكلا	لكن
لكنا	لكنت	لكنتم	لكنتما
لكنتن	لكنك	لكنكم	لكنكما
لكنكن	لكننا	لكنه	لكنها
لكنهم	لكنهما	لكنهن	لكني
لكي	للا	للات	للاتي
للأني	للأتان	للأتي	للأذان
للأذي	للأذين	للأست	للأستم
للأستما	للأستن	للأسنا	للأعل
للألك	للألكم	للألكما	للألكن
للألنا	للأله	للألها	للألهم
للألهما	للألهن	للألي	للأكن
للألكك	للألككم	للألككما	للألككن
للأكننا	للأكنه	للأكنها	للأكنهم

لكنهما	لكنهن	للكني	للم
للما	للن	للو	للول
للولاك	للولاكم	للولاكما	للولاكنا
للولانا	للولاه	للولاهما	للولاهم
للولاهما	للولاهن	للولاي	للولما
للولماك	للولماكم	للولماكما	للولماكن
للولمانا	للولماه	للولماها	للولماهم
للولماهما	للولماهن	للولماي	للوليت
للوليتك	للوليتكم	للوليتكما	للوليتكن
للوليتنا	للوليته	للوليتها	للوليتهم
للوليتهما	للوليتهن	للوليتي	للوليس
للوليسا	للوليسوا	لم	لما
لمذ	لمن	لمنذ	لمنك
لمنكم	لمنكما	لمنكن	لمننا
لمنه	لمنها	لمنهم	لمنهما
لمنهن	لمني	لن	لنحن
لنعم	لها	لهاك	لهاكم
لهاكما	لهاكن	لهانا	لهاه
لهاها	لهاهم	لهاهما	لهاهن
لهاي	لهل	لهلا	لهلاك
لهلاككم	لهلاكما	لهلاكنا	لهلانا
لهلاه	لهلاها	لهلاههم	لهلاهنا
لهلاهن	لهلاي	لهلاك	لهلكم
لهلكما	لهلكن	لهلنا	لهله
لهلها	لهلهم	لهلها	لهلهن
لهلي	لهم	لهما	لهن
لهنا	لهناك	لهناكم	لهناكما
لهناكن	لهنانا	لهناه	لهناها
لهناهم	لهناهما	لهناهن	لهناي
لهو	لهي	لهيا	لو
لوا	لوجد	لوجدا	لوجدت

لوجدتم	لوجدتما	لوجدتن	لوجدن
لوجدنا	لوجدوا	لولا	لولاك
لولاكم	لولاكما	لولاكن	لولانا
لولاه	لولاها	لولاهم	لولاهما
لولاهن	لولاي	لوما	لوماك
لوماكم	لوماكما	لوماكن	لوماننا
لوماه	لوماها	لوماهم	لوماهما
لوماهن	لوماي	ليا	ليت
ليتك	ليتكُم	ليتكما	ليتكُن
ليتتا	ليته	ليتها	ليتهم
ليتهما	ليتهن	ليتي	ليس
ليسا	ليسوا	ما	مذ
من	مذ	منك	منكم
منكما	منكن	مننا	منه
منها	منهم	منهما	منهن
مني	نحن	نعم	ها
هاك	هاكم	هاكما	هاكن
هانا	هاه	هاها	هاهم
هاهما	هاهن	هاهنا	هاهناك
هاهناكم	هاهناكما	هاهناكن	هاهنانا
هاهناه	هاهناها	هاهناهم	هاهناهما
هاهناهن	هاهناي	هاي	هتان
هتاك	هتانكم	هتانكما	هتانكن
هتاننا	هتانه	هتانها	هتانهم
هتانها	هتانهن	هتاني	هته
هتهك	هتهكم	هتهكما	هتهكن
هتهنا	هتهه	هتهها	هتههم
هتههما	هتههن	هتهي	هتي
هتيك	هتيكم	هتيكما	هتيكن
هتين	هتيئا	هتينك	هتينكم
هتينكما	هتينكن	هتيننا	هتينه

هتنيها	هتينهم	هتينهما	هتينهن
هتيني	هتية	هتيها	هتيهم
هتيهما	هتيهن	هتيي	هذا
هذاك	هذاكم	هذاكما	هذاكن
هذان	هذانا	هذانك	هذانكم
هذانكما	هذانكن	هذاننا	هذانه
هذانها	هذانهم	هذانهما	هذانهن
هذاني	هذاه	هذاها	هذاهم
هذاهما	هذهن	هذاي	هذه
هذهك	هذهكم	هذهكما	هذهكن
هذهنا	هذهه	هذهها	هذههم
هذههما	هذههن	هذهي	هذي
هذيك	هذيكم	هذيكما	هذيكن
هذين	هذيننا	هذينك	هذينكم
هذينكما	هذينكن	هذيننا	هذينه
هذينها	هذينهم	هذينهما	هذينهن
هذيئي	هذيه	هذيها	هذيهم
هذيهما	هذهين	هذيي	هل
هلا	هلاك	هلاكم	هلاكما
هلاكن	هلانا	هلاه	هلاها
هلاههم	هلاهنا	هلاههن	هلاي
هلك	هلكم	هلكما	هلكن
هلنا	هله	هلها	هلهم
هلها	هلهن	هلي	هم
هما	هن	هنا	هناك
هناكم	هناكما	هناكن	هنانا
هناه	هناها	هناهم	هناهما
هناهن	هناي	هو	هؤلاء
هؤلاءهم	هؤلاءك	هؤلاءكم	هؤلاءكما
هؤلاءكن	هؤلاءنا	هؤلاءه	هؤلاءها
هؤلاءهما	هؤلاءهن	هؤلاءي	هي

هيا	وا	واتخذ	واتخذنا
واتخذت	واتخذتم	واتخذتما	واتخذتن
واتخذن	واتخذنا	واتخذوا	والتي
والذي	والذين	واللاتي	واللاتي
واللتان	واللذان	وانفك	وانفكا
وانفكت	وانفككت	وانفككتم	وانفككتما
وانفككتن	وانفككن	وانفككنا	وانفككوا
وإذ	وإذا	وإذما	وإذماك
وإذماكم	وإذماكما	وإذماكن	وإذمانا
وإذماه	وإذماها	وإذماهم	وإذماههما
وإذماهن	وإذماي	وإذن	وإلا
وإلاك	وإلاكم	وإلاكما	وإلاكن
وإلانا	وإلاه	وإلاها	وإلاهم
وإلاهما	وإلاهن	وإلاي	وإلي
وإليك	وإليكم	وإليكما	وإليكن
وإلينا	وإليه	وإليها	وإليهم
وإليهما	وإليهن	وإما	وإيا
وإياك	وإياكم	وإياكما	وإياكن
وإيانا	وإياه	وإياها	وإياهم
وإياهما	وإياهن	وإياي	وأجل
وأصبح	وأصباحا	وأصبحت	وأصبحتم
وأصبحتما	وأصبحتن	وأصبحن	وأصبحنا
وأصبحوا	وأضحى	وأضحيا	وأضحيت
وأضحيتم	وأضحيتما	وأضحيتن	وأضحين
وأضحينا	وأضحوا	وأل	وأك
وأكم	وأكما	وأكن	وألنا
وآله	وآلها	وآلهم	وآلهما
وآلهن	وأم	وأمست	وأمسو
وأمسى	وأمسيا	وأمسيت	وأمسيتم
وأمسيتما	وأمسيتين	وأمسينا	وأن
وأنا	وأنت	وأنتم	وأنتما

وَأَنْتَن	وَأَنْك	وَأَنْكَمَا	وَأَنْكُن
وَأَنْكَن	وَأَنْه	وَأَنْهَما	وَأَنْهِي
وَأَوْ	وَأَوْشَكَ	وَأَوْشَكْتَا	وَأَوْشَكُنَا
وَأَوْشَكْتُمْ	وَأَوْشَكْتُنَا	وَأَوْشَكْتُمَا	وَأَوْشَكْتُمْ
وَأَوْشَكْنَا	وَأَوْشَكُوا	وَأَوْشَكْتُمْ	وَأَوْشَكْتُمْ
وَأَوْلَاكُمْ	وَأَوْلَاكُمْ	وَأَوْلَاكُمْ	وَأَوْلَاكُمْ
وَأَوْلَاهُ	وَأَوْلَاهُ	وَأَوْلَاهُ	وَأَوْلَاهُ
وَأَوْلَاهُنَّ	وَأَوْلَاهُنَّ	وَأَوْلَاهُنَّ	وَأَوْلَاهُنَّ
وَأَيْكُمْ	وَأَيْكُمْ	وَأَيْكُمْ	وَأَيْكُمْ
وَأَيْه	وَأَيْه	وَأَيْه	وَأَيْه
وَأَيَّهِنَّ	وَأَيَّهِنَّ	وَأَيَّهِنَّ	وَأَيَّهِنَّ
وَبَاتَتْ	وَبَاتَتْ	وَبَاتَتْ	وَبَاتَتْ
وَبَتْنَا	وَبَتْنَا	وَبَتْنَا	وَبَتْنَا
وَبَرَحَا	وَبَرَحَا	وَبَرَحَا	وَبَرَحَا
وَبَرَحْتُنَا	وَبَرَحْتُنَا	وَبَرَحْتُنَا	وَبَرَحْتُنَا
وَبَلَّ	وَبَلَّ	وَبَلَّ	وَبَلَّ
وَتَانَكُمْ	وَتَانَكُمْ	وَتَانَكُمْ	وَتَانَكُمْ
وَتَانَهُ	وَتَانَهُ	وَتَانَهُ	وَتَانَهُ
وَتَانَهُنَّ	وَتَانَهُنَّ	وَتَانَهُنَّ	وَتَانَهُنَّ
وَتَهَكُم	وَتَهَكُم	وَتَهَكُم	وَتَهَكُم
وَتَهَهُ	وَتَهَهُ	وَتَهَهُ	وَتَهَهُ
وَتَهْنَهُنَّ	وَتَهْنَهُنَّ	وَتَهْنَهُنَّ	وَتَهْنَهُنَّ
وَتِيكُمْ	وَتِيكُمْ	وَتِيكُمْ	وَتِيكُمْ
وَتِينَا	وَتِينَا	وَتِينَا	وَتِينَا
وَتِينَكُمْ	وَتِينَكُمْ	وَتِينَكُمْ	وَتِينَكُمْ
وَتِينَهُنَّ	وَتِينَهُنَّ	وَتِينَهُنَّ	وَتِينَهُنَّ
وَتِيَهُ	وَتِيَهُ	وَتِيَهُ	وَتِيَهُ
وَتِيَهُنَّ	وَتِيَهُنَّ	وَتِيَهُنَّ	وَتِيَهُنَّ
وَتَمَكُم	وَتَمَكُم	وَتَمَكُم	وَتَمَكُم
وَتَمَهُ	وَتَمَهُ	وَتَمَهُ	وَتَمَهُ

وتمهن	وتمي	وجد	وجدا
وجدت	وجدتم	وجدتما	وجدتن
وجدن	وجدنا	وجدوا	وجعل
وجعلا	وجعلت	وجعلتم	وجعلتما
وجعلتن	وجعلن	وجعلنا	وجعلوا
وجلل	وجير	وجيرك	وجيركم
وجيركما	وجيركن	وجيرنا	وجيره
وجيرها	وجيرهم	وجيرهما	وجيرهن
وحيري	وحاشا	وحاشاك	وحاشاكم
وحاشاكما	وحاشاكن	وحاشانا	وحاشاه
وحاشاها	وحاشاهم	وحاشاهما	وحاشاهن
وحاشاي	وحتى	وحسب	وحسبك
وحسبكم	وحسبكما	وحسبكن	وحسبنا
وحسبه	وحسبها	وحسبهم	وحسبهما
وحسبهن	وحسبي	وخلا	وخلاك
وخلاكم	وخلاكما	وخلاكن	وخلانا
وخلاه	وخلاها	وخلاهم	وخلاهما
وخلاهن	وخلاي	ودام	وداما
ودامت	وداموا	ودمت	ودمتن
ودمتما	ودمتن	ودمن	ودمنا
وذا	وذاك	وذاكم	وذاكما
وذاكن	وذان	وذانا	وذانك
وذانكم	وذانكما	وذانكن	وذاننا
وذانه	وذانها	وذانهم	وذانهما
وذانهن	وذاني	وذاه	وذاهها
وذاهم	وذاهما	وذاهن	وذاي
وزه	وزهك	وزهكم	وزهكما
وزهكن	وزهنا	وزهه	وزهها
وزههم	وزههما	وزههن	وزهي
وذي	وذلك	وذلكم	وذلكما
وذلكن	وذلكنا	وذلكنا	وذلكنا

وڌينڪم	وڌينڪما	وڌينڪن	وڌينڪا
وڌينه	وڌينهها	وڌينهم	وڌينهها
وڌينهڻ	وڌيني	وڌيه	وڌيها
وڌيهم	وڌيهما	وڌيهڻ	وڌيي
ورام	وراما	ورامت	وراموا
ورأوا	ورأى	ورأيا	ورأيت
ورأيتم	ورأيتما	ورأيتن	ورأين
ورأينا	ورب	وربك	وربكم
وربڪما	وربڪن	وربنا	وربه
وربها	وربهم	وربهما	وربهڻ
وربي	ورمت	ورمتم	ورمتما
ورمتن	ورمن	ورمنا	وزال
وزالا	وزالت	وزالوا	وزعم
وزعما	وزعت	وزعتم	وزعتما
وزعتن	وزعن	وزعنا	وزعموا
وزلت	وزلتم	وزلتما	وزلتن
وزلن	وزلنا	وسوف	وصار
وصارا	وصارت	وصاروا	وصرت
وصرتم	وصرتما	وصرتن	وصرن
وصرنا	وظل	وظلا	وظلت
وظللت	وظللتم	وظللتما	وظللتن
وظللن	وظلنا	وظلوا	وظن
وظنا	وظنت	وظنن	وظننا
وظننت	وظننتم	وظننتما	وظننتن
وظنوا	وعدا	وعداڪ	وعداڪم
وعداڪما	وعداڪن	وعداڪا	وعداڪه
وعداها	وعداهم	وعداهما	وعداهن
وعداي	وعسى	وعسيا	وعسيت
وعسيتم	وعسيتما	وعسيتن	وعسين
وعسينا	وعسيوا	وعل	وعلك
وعلڪم	وعلڪما	وعلڪن	وعلم

وعلما	وعلمت	وعلمتم	وعلمتما
وعلمتن	وعلمن	وعلمنا	وعلموا
وعلنا	وعله	وعلها	وعلهم
وعلها	وعلهن	وعلى	وعلي
وعليك	وعليكم	وعليكما	وعليكن
وعلينا	وعليه	وعليها	وعليهم
وعليهما	وعليهن	وعن	وعنك
وعنكم	وعنكما	وعنكن	وعننا
وعنه	وعنها	وعنهم	وعنهما
وعنهن	وعني	وفتئى	وفتئا
وفتنت	وفنتتم	وفنتئما	وفنتئن
وفنتن	وفنتنا	وفنتوا	وفى
وفيك	وفيكم	وفيكما	وفيكن
وفينا	وفيه	وفيهما	وفيهم
وفيهما	وفيهن	وفيي	وقد
وكاد	وكادا	وكادت	وكادوا
وكان	وكانا	وكانت	وكانوا
وكانك	وكانكم	وكانكما	وكانكن
وكاننا	وكانه	وكانها	وكانهم
وكانهما	وكانهن	وكانى	وكدت
وكدتم	وكدتما	وكدتن	وكدن
وكدنا	وكرب	وكربا	وكربت
وكربتم	وكربتما	وكربتن	وكربن
وكربنا	وكربوا	وكلا	وكن
وكنا	وكنت	وكنتم	وكنتما
وكنتن	وكي	ولا	ولات
ولست	ولستم	ولستما	ولستن
ولسنا	ولعل	ولعلك	ولعلكم
ولعلكما	ولعلكن	ولعلنا	ولعله
ولعلها	ولعلهن	ولعلها	ولعلهن
ولعلي	ولكن	ولكنك	ولكنكم

ولكنكما	ولكنكن	ولكننا	ولكنه
ولكنها	ولكنهم	ولكنهما	ولكنهن
ولكني	ولم	ولما	ولن
ولو	ولولا	ولولاك	ولولاكم
ولولاكما	ولولاكن	ولولانا	ولولاه
ولولاها	ولولاهم	ولولاهما	ولولاهن
ولولاي	ولوما	ولوماك	ولوماكم
ولوماكما	ولوماكن	ولوماننا	ولوماه
ولوماها	ولوماهم	ولوماهما	ولوماهن
ولوماي	وليت	وليتك	وليتكم
وليتكما	وليتكن	وليتنا	وليته
وليتها	وليتهم	وليتها	وليتهن
وليتي	وليس	وليسا	وليسوا
وما	ومذ	ومن	ومذ
ومنك	ومنكم	ومنكما	ومنكن
ومننا	ومنه	ومنها	ومنهم
ومنهما	ومنهن	ومني	ونحن
ونعم	وها	وهاك	وهاكم
وهاكما	وهاكن	وهانا	وهاه
وهاها	وهاهم	وهاهما	وهاهن
وهاي	وهل	وهلا	وهلاك
وهلاككم	وهلاكما	وهلاكنا	وهلانا
وهلاه	وهلاها	وهلاههم	وهلاهنا
وهلاهن	وهلاي	وهلك	وهلكم
وهلكما	وهلكن	وهلنا	وهله
وهلها	وهلهم	وهلهما	وهلهن
وهلي	وهم	وهما	وهن
وهنا	وهناك	وهناكم	وهناكما
وهناكن	وهنانا	وهناه	وهناها
وهناهم	وهناهما	وهناهن	وهناي
وهو	وهي	وهيا	ووا

ووجدتم	ووجدت	ووجدا	ووجد
ووجدنا	ووجدن	ووجدتن	ووجدتما
	يا	ويا	ووجدوا

المستخلص

لا وجود لمجاميع من الأسماء العربية. وعلاوة على ذلك، في أي نص عربي يمكن وجود الأسماء بأشكال مختلفة. في الواقع، عن طريق وضع علامات للأسماء في النص العربي، يمكن في كل جملة تحديد ما إذا كانت تبدأ مع اسم أو فعل.

توصيف أجزاء الكلام (POS) مهمتها وصف كل كلمة في الجملة بالفئة المناسبة، وهو ما يسمى الكلمات الدلالية (اسم، فعل، وحرف). في هذه الأطروحة، نحن نحاول وضع وصف لكلمات النص العربي غير المشكل. تستند طريقتنا المقترحة بالبحث عن كل كلمة من النص العربي في قوائم من أفعال والحروف. يتم العثور على الأسماء من خلال إقصاء الأفعال والحروف. وفرضية بحثنا تنص على أنه إذا لم يتم العثور على كلمة النص في القوائم لدينا، إذن إفتراضياً هي اسم. سيتم إجراء هذه المقارنات لكل كلمة من كلمات النص حتى يتم وصفها جميعاً.

لتطبيق أسلوبنا، قمنا بإعداد قائمة من الحروف والأفعال في اللغة العربية مع ما مجموعه ١١٢ مليون فعل وحرف مجتمعين، والتي تستخدم في المقارنات لإثبات فرضيتنا.

لتقييم أسلوبنا المقترح، استخدمنا كلمات سبق وصفها من "مدونة القرآن الكريم"، بمجموع ٧٨٢٤٥ من الكلمات، مع أسلوبنا، والنهج القائم على العلامات المقولبة (AraMorph) ونهج قائم على القواعد في وضع العلامات (ستانفور).

أخيراً، أنتج (AraMorph) ٤٠% من الكلمات بشكل صحيح وأنتج (ستانفور) ٦٨% من الكلمات بشكل صحيح، في حين أنتج أسلوبنا ٦٨٥٠١ من الكلمات بشكل صحيح الموسومة بمعدل (٨٨%).

استخدام تقنية القوالب لعنونة الأسماء العربية غير المشكّلة

هاشم بن سعد الدين آل غالب الشريف

بحث مقدم للحصول على درجة الماجستير في علوم الحاسبات

إشراف

الدكتور/ خالد بن عمر ثابت

كلية الحاسبات وتقنية المعلومات

جامعة الملك عبدالعزيز

جدة - المملكة العربية السعودية