

Multi- Relational Latent Morphology-Semantic “MOR-PHOSEM” Analysis Model For Extracting Qura’nic Concept A New Innovative For Sustainable Society

Dr Asma Abdul Rahman^{a*}, Dr Mohammad Hamiruce Marhaban^b, Colonel Dr Khairool Amali Ahmad^c, Dr Ahmad Abdul Rahman^d, Dr Fakroul Ridzuan Hashim^e, Maryam Abdul Rahman^f

^a Assoc.Prof, Department of Education in Linguistics, Faculty of Major Languages studies, Islamic Science University of Malaysia, Nilai, Negeri Sembilan, Malaysia

^bProfessor, Faculty of Engineering, Islamic Science University of Malaysia, Nilai, Negeri Sembilan, Malaysia

^cAssoc.Prof, Faculty of Engineering, National Defence University of Malaysia, Sungai Besi, Kuala Lumpur, Malaysia

^dSenior Lecturer, Center for Languages studies and generic development, University Malaysia Kelantan, Jeli, Kelantan, Malaysia

^e Assoc. Prof, Faculty of Engineering, National Defence University of Malaysia, Sungai Besi, Kuala Lumpur, Malaysia

^fDepartment of Arabic studies and Islamic civilization, Faculty of Islamic studies, The National University of Malaysia

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: Al Quran is a divine text which represents the purest and most authentic form of the classical Arabic language. To understand the meaning of each verse, a deep knowledge of Arabic linguistic is essential. Therefore, our scholars have made their efforts by engaging themselves in the works of explaining al Quran’s words, interpreting its meanings into Arabic and other languages. Currently, more people are interested in knowing the content of al-Quran, especially for non-Muslim, after 9/11 tragedy. Thus, a flexible model that can represent Qur’anic concept is required for people to understand the content of the Quran. In this research, we propose a Multi-Relational Latent Morphology-Semantic Analysis Model (MORPHOSEM) based on a combination of Arabic Semantic and six multiple relations between words, which are synonym, antonym, hypernym, hyponym, homonym and meronym, to precisely extract Qur’anic concept. The existing literatures focus only on very limited relationships between words which could not extract the in-depth concept of Qur’anic without considering the importance Arabic Semantic. Therefore, the objectives of this research are: (1) to analyses and categorize Quranic words according to Arabic Semantic patterns, (2) to propose a new model for extracting Quranic concept using MORPHOSEM, (3) to investigate semantic relationships between Qur’anic words, and (4) to validate the proposed model with Arabic linguistic, and Qur’anic experts. This research will be conducted qualitatively through content analysis approach a new innovative technological technique. It is expected that the model will come out with a precise analysis for extracting Qur’anic concept. This will be very significant in enhancing the overall Quran’s un-derstanding among the society in Malaysia and Muslim’s world for sustainable society.

Keywords: Multi-Relational, Latent, Morphology-Semantic, Model Extracting, Qur’anic, Sustainable society

1. Introduction

In linguistic study, a concept is referred as a mental structure, which totally depends on relationships between linguistic representatives amongst words and phrases. However, the interchangeable relationship could happen due to the inherent conciseness of natural language. Therefore, many researchers try to solve this problem using ontology approach and Latent Semantic Analysis (LSA) (Chang & Meek, 2013; Ozcan & Aslandogan, 2004), but it needs some considerations before it can be implemented on Arabic, especially in al-Quran, which represents a special classical Arabic language.

2. Problem Statement

A variety of problems have emerged in Malaysia community because of the rapid developments and the flow of humanitarian and scientific knowledge. Thus, that people unable to adapt to the huge number of methods and means of modern education technology. The impact of foreign media and audio-visual and print and unsecured computer network bring about strange behaviour, not only in Malaysia social aspects, but also Arabic language education procedures. The reform of this problem calls for more attention to be paid on theories education teaching methods, to the quality and quantity the rules, principles, methods and curricula of teaching and evaluation. This may includes putting emphasis moral, cultural and heritage education based on the Qur’an and Sunnah.

The uniqueness of Arabic Semantic makes the combination with LSA and ontology more significant to extract Qur’anic concept. It is because Phonological analysis is a basic step in various applications including text mining, information retrieval (IR), machine translation, automatic summarization, and Arabic learning systems (Desouki, 2011). LSA on the other hand has limitation, such as difficulty in differentiating fine-grained relations between lexical semantics, synonyms, antonyms, and hypernyms. Due to that limitation, various studies have been carried out to overcome this weakness of LSA introducing Multi-Relational Latent Semantic Analysis using synonym,

antonym and hypernym (Chang & Meek, 2013) but it has not yet been applied on Arabic Text such as al-Quran. Furthermore, Qura'nic ontology studies are lacking of accuracy of language because it depends on the translated version of al-Quran and uses aplastic noun (Azman Ta'a, Abidin, Abdullah, Ali, & Ahmad, 2013; Hikmat Ullah Khan, Muhammad Saqlain, Shoaib, & Sher, 2013; Maha Al-yahya & Hend Al-khalifa, n.d.). The existing literature shows that there are still gaps and problems in extracting Qur'anic concept as have been mentioned previously, due to inappropriate approaches that had been used ignoring the importance of Arabic Semantic and LSA in deriving the meaning of Qur'anic word. Therefore, in order to address the above mentions problem, we propose a new model in extracting Qur'anic concept based on Arabic Semantic and Multi-Relational Latent Semantic Analysis, using combination of six multiple relations between words, which are synonym, antonym, hypernym, hyponym, homonym and meronym.

3.Objective of the research

- 2.1 To analyze and categorize Qur'anic words according to Arabic Semantic patterns.
- 2.2 To investigate semantic relationships between Qur'anic words.
- 2.3 To propose a new model for extracting Qur'anic concept using MORPHOSEM.
- 2.4. To validate the proposed model.

4. Research Questions

- 3.1 How to analyze and categorize Qur'anic words according to Arabic Semantic patterns?
- 3.2 What is result of investigate semantic relationships between Qur'anic words?
- 3.3 How to propose a new multi-relational Latent Phono-Semantic Analysis for extracting Qur'anic concept?
- 3.4 What is the subject matter expert's opinion about the proposed model?

5. Literature Reviews

Arabic is considered one of the phonologically complex languages. Each word can be derived from roots which have, in most cases, three letters by applying templates construct stems and then attaching them to prefixes and suffixes to obtain a very large number of different surface forms (Amr El-Desoky Mousa, Ralf Schl'uter, 2012). Besides that, a single root can be transformed into different word with different pattern, vocalism, and pronunciation (Bassam Al-Salemi; Mohd. Juzaidin Ab Aziz, 2011). The uniqueness of Arabic Semantic makes the combination with LSA and ontology more significant to extract Quranic concept.

Generally, LSA attempts to reveal the hidden conceptual relationships among words and phrases based on linguistic usage patterns. Usually it will be presented in taxonomic structure consisting of a hierarchy word and its relationship, such as synonym, antonym, hyperonym/hyponym, meronym/homonym (member, substance, and part), entailment, cause, attribute, and similarity. LSA has multiple uses in various fields such as an instrument of text summarization and summary evaluation (Steinberger & Ježek, 2004), an approach to Source-Code Plagiarism Detection and Investigation (Cosma, 2008), and Term Prediction instrument (Zhao & Callan, 2010). The following diagram illustrates the relationship between the component parts of the semantic model used by Martin Bryan (2003):

However, LSA has limitation such as difficulty in differentiating fine-grained relations between lexical semantics, synonyms, antonyms, and hypernyms. Due to that limitation, various studies have been carried out to overcome this weakness of LSA by introducing the notion of polarity. The recent LSA research combines multiple relations between words by constructing three relationships which are known as Multi-Relational Latent Semantic Analysis using synonym, antonym, and hypernym (Asma Abdul Rahman 2007-2017).

المقصود بما	الرموز
جملة	ج
طرف اسمي	طس
طرف فعلي	طف
اسم	س
فعل / فاعل / مفعول به	ف / فا / مف
ضمير	ض
مضاف / مضاف إليه	ضف / ضف إليه
شبه جملة	شج
أداة النسخ	د/نسخ
أداة التعريف (ال)	تع
مبتدأ / خبر	م / خ

- (١) قواعد الحذف Deletion المراد به حذف عنصر من عناصر الجملة لغرض ما. فالرموز المشيرة إلى ذلك كالتالي:
- $$a + b \rightarrow b / a \quad a + b \leftarrow a / b$$
- (٢) قواعد الإحلال Replacement المراد به إبدال عنصر بعنصر آخر، وقد يكون إبدال كلمة بكلمة، وإبدال محل بمحل. فالرموز المشيرة إلى ذلك كالتالي:
- $$a \rightarrow b \quad a \leftarrow b$$
- وهذا يحدث غالبا في تحويل الجملة المبنية للمعلوم إلى المبنية للمجهول حين يُبدل المفعول به محل الفاعل، فصار نائباً للفاعل.
- (٣) قواعد التوسيع (التسديد) Expansion المراد به تمديد أو توسيع جملة بطرف من أطراف البنية اللغوية. فالرموز المشيرة إلى ذلك كالتالي:
- $$a \rightarrow b + c \quad a \leftarrow b + c$$
- وهذا يكون غالبا بإضافة عبارة إلى جملة لإطالتها. فمثال:
- (٤) قواعد الزيادة Addition المراد به إضافة عنصر جديد إلى المسند أو المسند إليه مع الحفاظ على المعنى أو المقصود الأصلي. فالرموز المشيرة إلى ذلك كالتالي:
- $$a \rightarrow b + a \quad a \leftarrow b + a$$
- وهذا يكون غالبا بإضافة نعت أو توابع أخرى أو أدوات مثل، إنَّ وأحواتها، وكان وأحواتها، ولام التوكيد، وهلم جرا إلى جملة كاملة.
- وهما نلاحظ أن قواعد التوسيع (التسديد) وقواعد الزيادة متقاربة. فَمَا قواعد التوسيع لأجل تكوين الجملة واسعة الأفكار. وأما قواعد الزيادة فإنها تهم كثيرا بالحفاظ على المعنى الأصلي، فالإضافة لا تكون مغيرة للجملة.

Figure 1: Multi-Relational Latent MORPHOSEM formula Analysis by Asma Abdul Rahman

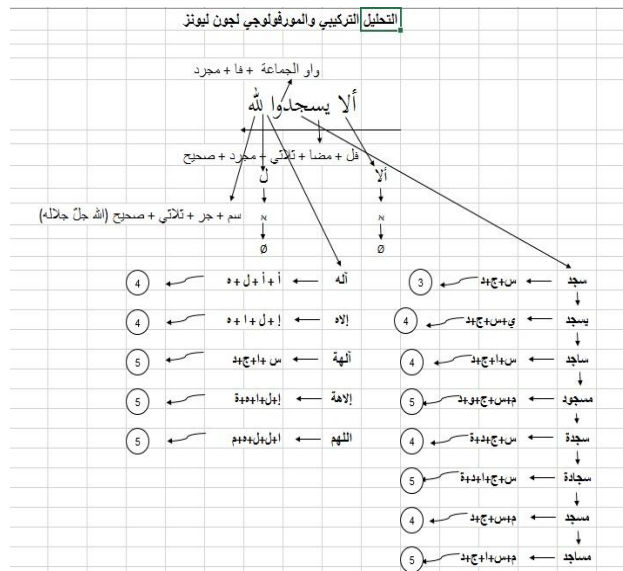
4.1 The above model encodes the raw data in a 3-way tensor to encode multiple word relations, which are synonym, antonym & Hypernym. Each slice captures a particular relation and is in the format of the document-term matrix in LSA. The tensor decomposition method was applied to generalize the representation and discover unseen relations between words.

4.2 Although this recent work tries to explore concept-based information access via ontology approach and LSA on natural language (Ozcan & Aslandogan, 2004) it has not yet been done in al-Quran, which represents the purest and most authentic form of the classical Arabic language (Asma Abdul Rahman, 2003-2018). Muslims believe that words of Holy Quran are divine and eternal. No alteration is whatsoever possible as Allah Almighty Himself has taken into His Hand.

4.3 Ontology on the other hand, is defined as a description of the concepts and relationships that can exist for an agent or a community of agents (Ozcan & Aslandogan, 2004). The use of ontology facilitates identification of concepts and their linguistic representatives, given a key concept. It is like a dictionary or glossary, but with greater detail and structure.

4.4 The existing research show two types of approaches in Qur'anic language computational models based on ontological approach. First, traditional approach which is based on models of Roman language done by Hikmat Ullah Khan, Muhammad Saqlain, Shoaib, & Sher (2013) using English Translation of Holy Quran by Pickthall, and Azman Ta'a, Abidin, Abdullah, Ali, & Ahmad, (2013) using Syammil Al-Quran Miracle the Reference. These two studies lack accuracy of language because it depends on the translated version of al-Quran.

4.5 Second, modern approach is based on an authoritative and rich source of Arabic language, i.e. the Holy Quran. This study uses lexicon ontology development based on the Unified Process for Ontology (UPON), an ontological engineering approach (Maha Al-yahya & Hend Al-khalifa, n.d.). The study focuses mainly on nouns from the "time" semantic field, which is aplastic noun - having no forms except one.



وَلَا تَجْهَرُوا لَهُ، بِالْقَوْلِ كَجَهْرِ بَعْضِكُمْ لِبَعْضٍ أَن تَحْبَطَ أَعْمَالِكُمْ وَأَنتُمْ لَا تَشْعُرُونَ ﴿٢٠٤﴾

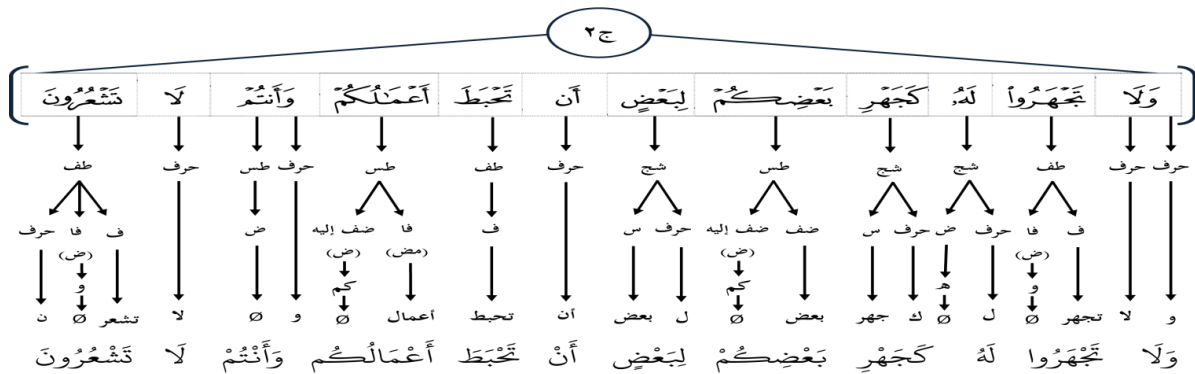


Figure 2: Model of MORPHOSEM in Qur’anic Ontology by Asma Abdul Rahman

4.6 In the above model, semantic dimensions need to be determined, and then words will be organized into a hierarchical classification with general concepts at the top, and specific at the bottom. Words in the hierarchy are associated with components via ontological relations. This classification structure of the ontology implies that the deeper word moves into the hierarchy, the more arguments the componential formula will have, and therefore the meaning narrows. In contrast, words at higher levels have fewer arguments in their componential formula, and therefore the meaning broadens. This model however did not include the root of words which is the appropriate to Arabic language.

4.7 The existing literature shows that there are still gaps in extracting Qur’anic concept because the aspect of Arabic Semantic and LSA does not fully utilize. Therefore, in order to address the above mentions problem, we propose a new model in extracting Qur’anic concept based on Arabic Semantic and Multi-Relational Latent Semantic Analysis, using combination of six multiple relations between words, which are synonym, antonym, hypernym, hyponym, homonym and meronym.

6.Methodology

This research is focusing on designing a multi-relational LSA model in extracting Qur’anic concept based on combination of Arabic Semantic and six multiple relations between words, which are synonym, antonym, hypernym, hyponym, homonym, and meronym. The researcher will used text analysis method as follow:

For more specifically, this approach will be conducted in the following techniques as follows:

1st Phase: Phonological Root Analysis

This phase will be conducted in following steps:

1. The researcher will identify digital document file of Qur’anic text. The text’s authentication will be verified by expert.

2. The digital document will be analyzed using corpus analysis using Wordsmith version 2.1,
3. Each word in text will be categorized into two families of verb and noun, excluding particles.
4. The researcher will identify root of each word in these two categories, then categorize the words belong to similar root in another sub-family
5. To enhance the confidence of investigation, the result then will be triangulated with three Arabic referred dictionaries:
 1. Lisan al-Arab by Ibn Manzuur,
 2. Mu'jaam al Waseet by Majma' Luhghah al-Arabiyyah in Cairo, and
 3. al-Mu'jam al-Asasi by al-Munazzamah al-Arabiyyah lil Tarbiyyah wa Thaqafah wal 'Ulum

2nd Phase: To Propose A New Model for Extracting Qur'anic Concept

Based on previous 1st phase result, the researcher will follow the next steps:

1. Create and propose a model using six multiple relations between words in Latent Semantic Analysis as follow:

- i.Synonym: a word or phrase that means exactly or nearly the same as another word or phrase in the same language
- ii.Antonym: a word opposite in meaning to another
- iii.Hypernym: a word with a broad meaning constituting a category into which words with more specific meanings fall
- iv.Hyponym: a word of more specific meaning than a general or super ordinate term applicable to it

homonym: A concept of which this concept forms a part VI. Meronym: A term that denotes part of something

2. Related connections of word to conceptual meaning then will be identified 3. Linking the related connections to key concept in tree-form drawing. 4. Describing the design of extracting Qur'anic concept based on ontological approach in Latent Semantic Analysis. **3rd Phase:** Investigating Semantic Relationships between Qur'anic Words

The proposed model will be validated by testing a concept of "Sight" in Quran as a sample. (an expected analysis is shown as below)**4th Phase:** Validating the proposed model

The result then will be evaluated by two groups of experts:

1. An Arabic linguist who will examine Phonological Root Analysis and the merging process with multi-relational latent semantic analysis, and
 2. Qur'anic experts who examine holistically the concept of Sight in al-Quran, according to Islamic perspective
- the evaluation will be conducted in focus group interview. Any suggestion or feedbacks will be taken into consideration to improve the model.

A sample of proposed model in using MORPHOSEM to extract Qur'anic concept

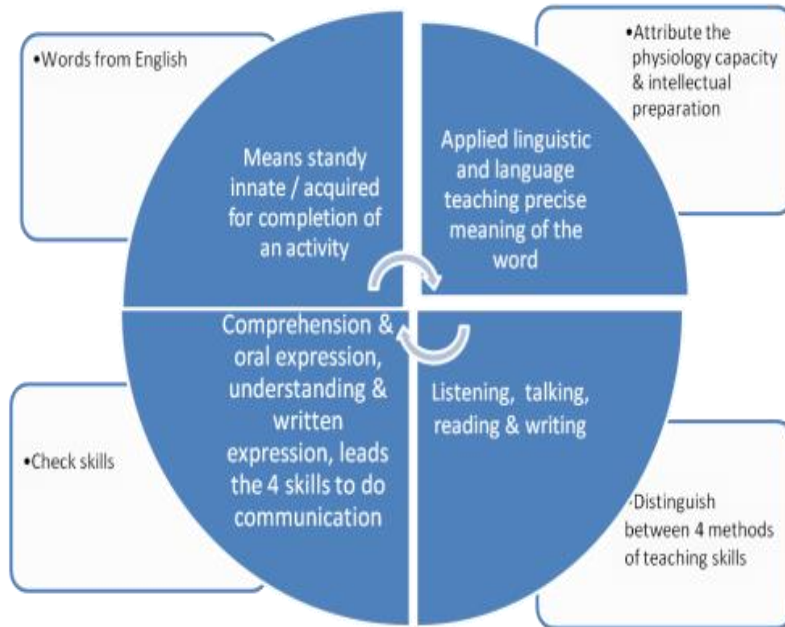


Figure 3: (Morphosem) Alm For Arabic Linguistics Learning Skills, Ass For Arabic Linguistics Speaking Skills, Ars For Arabic Linguistics Reading Skills, Aws For Arabic Linguistics Writing Skills



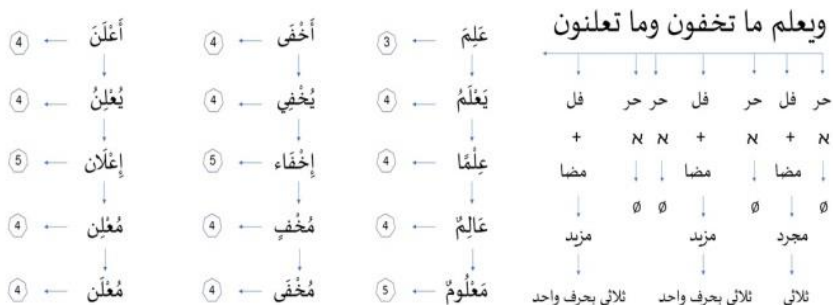
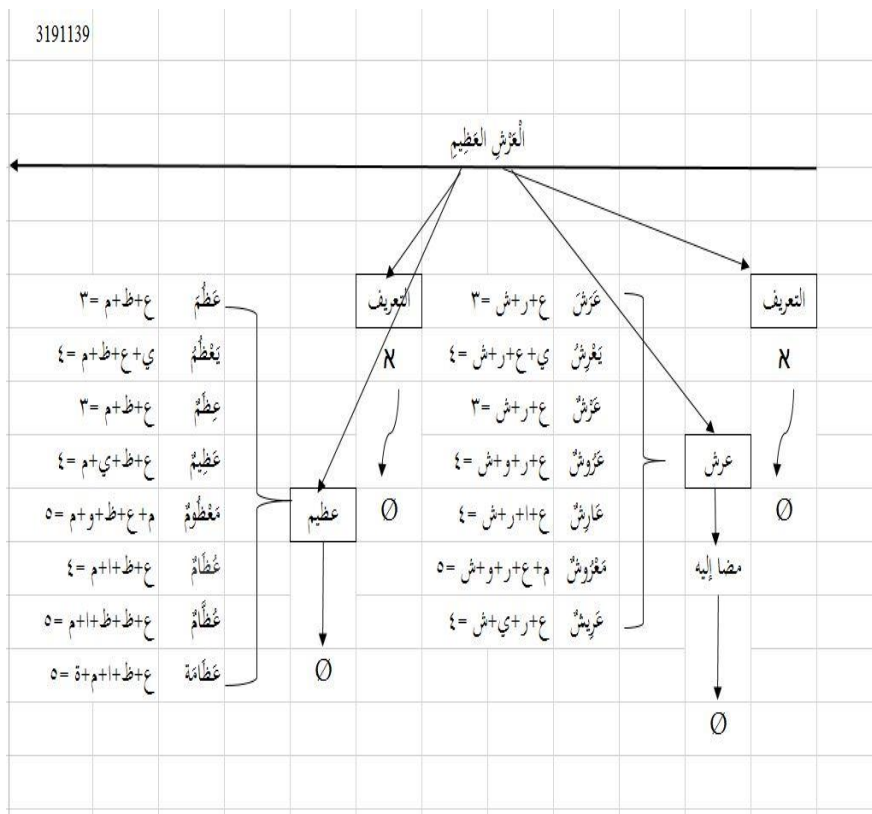


Figure 4: How to analyze Qur'anic verses based on MORPHOSEM by Asma Abdul Rahman

Figure 5: How to analyze Qur’anic verses based on MORPHOSEM by Asma Abdul Rahman

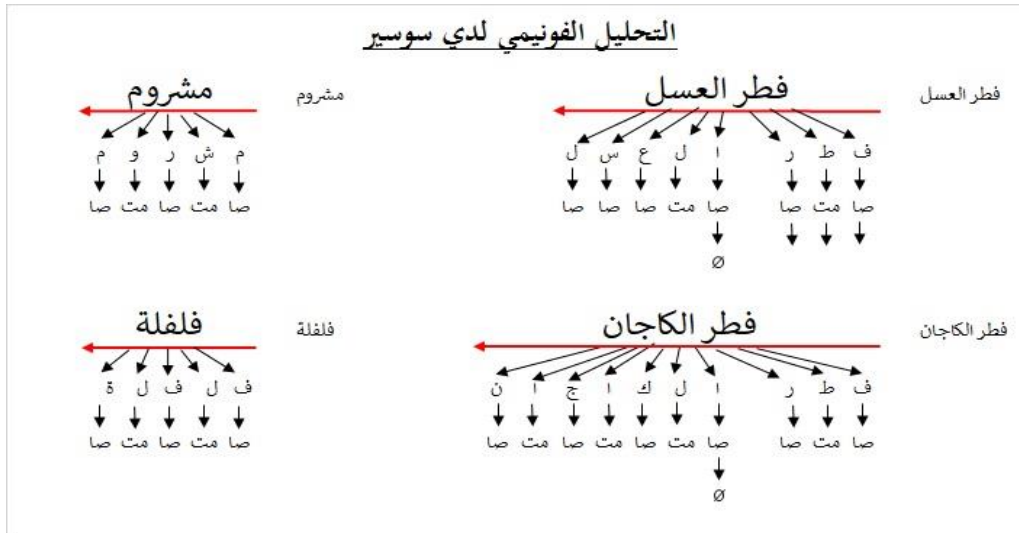


Figure 6: Extracting Quranic MORPHOSEM concept from various multiple relationships between words by Asma Abdul Rahman

أَيْحَسِبُ الْإِنْسَانُ أَنْ يَجْمَعَ عِظَامَهُ.

<p>أَيْحَسِبُ الْإِنْسَانُ الْكَافِرُ أَنْ نَقْدِرَ عَلَى جَمْعِ عِظَامِ الْإِنْسَانِ.</p>		<p>البنية العميقة</p>
الحذف	<p>أَيْحَسِبُ الْإِنْسَانُ \emptyset أَنْ نَقْدِرَ عَلَى جَمْعِ عِظَامِ الْإِنْسَانِ.</p>	<p>أَيْحَسِبُ الْإِنْسَانُ (الْكَافِرُ) أَنْ نَقْدِرَ عَلَى جَمْعِ عِظَامِ الْإِنْسَانِ.</p>
الحذف والإحلال	<p>أَيْحَسِبُ الْإِنْسَانُ أَنْ نَجْمَعَ عِظَامَ الْإِنْسَانِ.</p>	<p>أَيْحَسِبُ الْإِنْسَانُ أَنْ (نَقْدِرَ عَلَى) (جَمْعِ) عِظَامِ الْإِنْسَانِ.</p>
النسخ	<p>أَيْحَسِبُ الْإِنْسَانُ أَنْ نَجْمَعَ عِظَامَ (____) (____).</p>	<p>أَيْحَسِبُ الْإِنْسَانُ أَنْ نَجْمَعَ عِظَامِ (الْإِنْسَانِ).</p>
<p>أَيْحَسِبُ الْإِنْسَانُ أَنْ يَجْمَعَ عِظَامَهُ.</p>		<p>البنية السطحية</p>

Figure 7: Extracting Qur’anic MORPHOSEM concept from various multiple relationships between in the structure of sentence by Asma Abdul Rahman

7. Finding and New Result Novelty

6.1 Policies for government agencies to help them curb and solve the educational system in teaching and learning Arabic linguistics training and practices in tricking consumers into thinking that particular product by apply a new innovative method for whole level of students and society. This is an invention; thus, no comparable product existed in the market. Impact on human being and socio economic. The “(MORPHOSEM)” have improved knowledge and linguistic skills in the production of higher quality of human capital. Furthermore, the researcher was sales from text and reference, more formula books, CDs and Courses, training, workshop generate additional income for.

6.2 Harmonized society and community where Muslim and non-Muslim can seat together enjoying their communication with multiple lingual; that Muslim be it Malaysian or international could be unity of the nation local or foreign without any doubt or miscommunication.

6.3 Academically closing the gap on communication between society and community related research, providing the aspect of language, culture, and society perspective.

6.4 As academics, we aim to contribute to closing the gap on miscommunication related research by looking from the aspect of language, culture, and society perspective. The research will result in publications and we target to bring about the knowledge into seminars nationally and internationally and symposium to share with universities and industry players and those who are interested.

We have to find that the research output will useful to attract foreign direct investors, mainly Asian countries, and other potential interested investor from Middle East, and Europe to understand the opportunity in linguistics education, to join in the development of Malaysia and other countries of Renewal Energy (SCORE) specifically in two identified clusters Linguistics Education with a new innovative method, society, cultures –Hub sustainability and Tourism.

8. Conclusion

This Study was studying that relationship exists between morphology and semantic and legacy language old in the Qur'anic structure the of role the investigates latter. The course linguistic semantic the highlights also It. meaning the people in morpheme morphologic semantique the of past augmented trilingual the of variety and multiplicity it which was applied *MORPHOSEM* method by USIM and Nationally and internationally IPTA's , registry with charged and loaded is word Arabic the that demonstrated has been fertile a and action of place large a speaker the gives that fact , It's of investigations and interpretation of space of Classification in a meaning

The study concluded with the following: - That the Qur'an was revealed in the Arabic miracle language, just as other languages have their various means of constructing the word, and diversifying its connotations, as it depends in the formation of additional structures on clamping by increasing one, two or three letters, and this aspect of the linguistic lesson The focus of interest in morphology, in which the predecessor had a known effort, and that the increased buildings are more meaningful than the abstract ones; And that is because it achieves an increase in the meaning, so an increase in the building comes to an increase in the meaning or a change in it that follows the meaning of the word to give the desired meaning There is a direct proportionality between the formula and the sign, the more the building, the stronger the sign. The buildings of the Trio are more, the addition of the form of "muta'fal", which is the formula for which Ibn Jinni gave in the characteristics six examples, for which Ahmed Mukhtar Omar, Ibrahim al-Samarrai and Shawqi Dhaif gave more than two examples, from what was heard in the heritage or what was emerging in contemporary Arabic words, and their evidence is: the letter The excess in the word among the Arabs, like the letter Meem in this form, is what the original letter is in a bundle of derivation It is clear that the Arabs do this when they want to express new connotations alongside the semantics of abstract words, which makes the extra letter in the words take the rule of the original letters, and by doing so, Ibn Jinni is the first to record this formula, "a formula that is activated," with what he mentioned of examples and from a sound and correct protest. For the examples, the morphological rule can be applied, and a study of the structures of the triple past more, and the structures of the triple past more and more special importance, in terms of listing the buildings and mentioning their most famous meanings, their representation, and sometimes mentioning their evidence. - It reads from the augmented buildings that if the word has a high degree of expressive power and includes an increase, then that power has acquired through that increase many different and many connotations, such as: creation, creation, Ghadan, Ogodden, repressed, and suppressed.

References

1. Amr El-Desoky Mousa, Ralf Schl' uter, H. N. (2012). INVESTIGATIONS ON THE USE OF MORPHEME LEVEL FEATURES IN LANGUAGE MODELS FOR ARABIC LVCSR, 5021–5024.
2. Asma ABDUL Rahman, (2003-2007), LINGUISTICS STUDIES, Publishing Section USIM, Bandar Baru Nilai, Negeri Sembilan.
3. Asma ABDUL Rahman, (2007-2017), A'LIYAT TAHLIL AL-NASS AL-QURANI AL-LUGHAWWY: DIRASAH LUGHAWIYYAH DA-LALIYYAH (Formula Book), Publishing Section USIM, USIM, Bandar Baru Nilai, Negeri Sembilan.
4. Asma ABDUL Rahman, (2018-2019), MODERN LINGUISTICS STUDIES, Publishing Section USIM, USIM, Bandar Baru Nilai, Negeri Sembilan.

5. Azman Ta'a, Abidin, S. Z., Abdullah, M. S., Ali, B. B. M., & Ahmad, M. (2013). AL-QURAN THEMES CLASSIFICATION USING ONTOLOGY, (074), 383–389.
6. Chang, K., & Meek, W. Y. C. (2013). Multi-Relational Latent Semantic Analysis.
7. Cosma, G. (2008). An approach to source-code plagiarism detection investigation using latent
8. Cosma, G. (2008). An approach to source-code plagiarism detection investigation using latent
9. semantic analysis. IEEE Transactions on Computers. Retrieved from <http://eprints.dcs.warwick.ac.uk/401/1/cs-rr-440.pdf>
10. Desouki, S.-G.-. (2011). An Application Oriented Arabic Phonological Analyzer, (27), 7–19.
11. Hikmat Ullah Khan, Muhammad Saqlain, S., Shoaib, M., & Sher, M. (2013). Ontology Based Semantic Search in Holy Quran. International Journal of Future Computer and Communication, 2(6), 570–575. doi:10.7763/IJFCC. 2013.V2.229.
12. Maha Al-yahya, & Hend Al-khalifa. (n.d.). AN ONTOLOGICAL MODEL FOR REPRESENTING SEMANTIC LEXICONS: AN APPLICATION ON TIME NOUNS IN THE HOLY QURAN, 35(2), 21–35.
13. Ozcan, R., & Aslandogan, Y. A. (2004). Concept Based Information Access Using Ontologies and Latent Semantic Analysis, 1–16.
14. Steinberger, J., & Ježek, K. (2004). Using Latent Semantic Analysis in Text Summarization. In Proceedings of ISIM 2004, 93–100.
15. Zhao, L., & Callan, J. (2010). Term necessity prediction. Proceedings of the 19th ACM International Conference on Information and Knowledge Management - CIKM '10, 259. doi:10.1145/1871437.1871474