Available online at www.refaad.com
Gen. Lett. Math., 11(1) (2021), 10- 11
Research Article

General Letters in Mathematics (GLM)

Journal Homepage: http://www.refaad.com/views/GLM/home.aspx

ISSN: 2519-9277 (Online) 2519-9269 (Print)

Mathematical formalism on "Arabic Language DNA"

Aiman Eid Al-Rawajfeh[a*], Zuhier Altawallbeh[a]

[a] Tafila Technical University, Jordan

Abstract

In this work we formalize new findings on formal Arabic language by constructing equivalence classes on letters depending on inversion principle. This equivalent relation furnishes a conjecture that Arabic language has a DNA-like inversion controller, called language DNA (LDNA) (Al-Rawajfeh, 2020, Ref [2]) and it can be even considered as a part of the full language DNA that can be discovered by more investigations similar to this work.

Keywords: *Language DNA (LDNA); Arabic; equivalence classes; inversion; mathematical formalism.*

MSC2010: *03B65, 03C07, 11F23.*

1. Introduction

The text of the Holy Qur'an [1] represents ultimately the best unchanged text that applied the Arabic language rules, meanings, developments, and extension. It is eventually the best source of Arabic literature for applying theories, especially linguistics theories, or in deeper understanding, the best source to extract new theories.

In this work, we formalize the new findings of Al-Rawajfeh [2] onto the Arabic language that is called "Language DNA" (LDNA). LDNA depends on inversion principle, by defining an equivalent relation on the Arabic alphabet set $\sum$. Similar to the formation of DNA in human beings and its role, LDNA gives us the rule of inversion an Arabic letter to another and was formed by folding the Arabic alphabet set depending on some independent cases (sub-systems) that are stated in the text of the Holy Qur'an to build the whole system of the letter-to-letter relation [2]. This relation furnishes to set a great conjecture in Arabic language as it will be shown down. In general, an alphabet of the formal language is a finite set of elements or letters, and a string is a sequence of such elements [4]. In logical sources, words are sometimes called formulas [3].

2. Mathematical Formalism

The Arabic language L is a subset of the set of all words (strings) $\sum^*$ over the Arabic alphabet set $\sum$. Let us write all the Arabic letters as their equivalent English sounds symbols as follows.

$$\sum = \{A, B, T, \theta, J, \text{H}, X, D, \text{Đ}, R, Z, S, \int, \text{ş}, \text{ḍ}, \text{Ţ}, \text{ẓ}, \text{Ç}, \text{Ɣ}, F, Q, K, L, M, N, H, W, Y\}$$

$$\sum = \{ \text{ا}, \text{ب}, \text{ت}, \text{ث}, \text{ج}, \text{ح}, \text{خ}, \text{د}, \text{ذ}, \text{ر}, \text{ز}, \text{س}, \text{ش}, \text{ص}, \text{ض}, \text{ط}, \text{ظ}, \text{ع}, \text{غ}, \text{ف}, \text{ق}, \text{ك}, \text{ل}, \text{م}, \text{ن}, \text{ه}, \text{و}, \text{ي} \}$$

It was founded in [2] that if some letters replaced by others in a word taken from the Qur'an, then it will lead to the same meaning or a nearby meaning. We checked the meaning of these new words, after replacing some letters by others, either by usage of these new words in other verses and positions in the Qur'an, or by using dictionaries about Arabic language to check their meanings. In addition, some of the new words have no meaning known in Arabic language up to date, which means that they might be used in the future for next generations.

*Corresponding author
Email address: aimanr@yahoo.com (Aiman Eid Al-Rawajfeh)
doi: 10.31559/glm2021.11.1.2
Received: 21 Jun 2021 Revised: 18 Aug 2021 Accepted: 23 Aug 2021

More precisely and mathematically, we define the relation R on $\sum$ by $x \sim_R y$ by replacing the letter x by the letter y in a given word. The following lemma formalizes this kind of relation between letters as follows.

Lemma: If the relation R on $\sum$ is given by $M \sim_R B$, $F \sim_R$ Ħ , ṣ $\sim_R S$, $Q \sim_R J$, $L \sim_R T$, $K \sim_R$ ɵ, ɣ $\sim_R X$, ʕ $\sim_R D$, ẓ $\sim_R$ Đ, Ṭ $\sim_R R$, ḍ $\sim_R Z$, $H \sim_R N$, $Y \sim_R W$, $A \sim_R A$, $A \sim_R Y$, $B \sim_R M$, Ħ $\sim_R F$, $S \sim_R$ ṣ , $J \sim_R Q$, $T \sim_R L$, ɵ $\sim_R K$, $X \sim_R$ ɣ, $D \sim_R$ ʕ, Đ $\sim_R$ ẓ, R$\sim_R$Ṭ, $Z \sim_R$ ḍ, $\int \sim_R \int$, $N \sim_R H$, $W \sim_R Y$, $W \sim_R A$, $Y \sim_R A$ and $\sigma \sim_R \sigma$ for all $\sigma \in \sum$, then R is an equivalent relation on $\sum$.

Proof: Firstly, since it is given that $\sigma \sim_R \sigma$ for all $\sigma \in \sum$, the relation R is reflexive. By this property, we mean that if we fix the letter σ in a given word from Qur'an, regardless that we fix some other letters or even have been changed under R, then the new word will give the same meaning. Secondly, it is given that $M \sim_R B$ and $B \sim_R M$, $F \sim_R$ Ħ and Ħ $\sim_R F$, ṣ $\sim_R S$ and $S \sim_R$ ṣ, $Q \sim_R J$ and $J \sim_R Q$, $L \sim_R T$ and $T \sim_R L$, $K \sim_R$ ɵ and ɵ $\sim_R K$, ɣ $\sim_R X$ and $X \sim_R$ ɣ, ʕ $\sim_R D$ and $D \sim_R$ ʕ, ẓ $\sim_R$ Đ and Đ $\sim_R$ ẓ, Ṭ $\sim_R R$ and $R \sim_R$ Ṭ, ḍ $\sim_R Z$ and $Z \sim_R$ ḍ, $H \sim_R N$ and $N \sim_R H$, $Y \sim_R W$ and $W \sim_R Y$, $A \sim_R Y$ and $Y \sim_R A$ which means that R is symmetric. Finally, from the given relations between letters that transitivity holds on $\sum$. This completes the proof.

Because that the Arabic trilateral root of a word defines the underlying meaning of the word, without loss of generality if $\sigma_i, \sigma_i' \in \sum$ where $1 \leq i \leq 6$ and $\sigma_1 \sim_R \sigma_1', \sigma_2 \sim_R \sigma_2', \sigma_3 \sim_R \sigma_3'$ then we define another relation R'on $\sum'$ by saying that $\sigma_1\sigma_2\sigma_3 \sim_{R'} \sigma_1'\sigma_2'\sigma_3'$. The proof of the following lemma is straightforward and thus omitted.

Lemma: The relation R' defined above on $\sum^*$ is an equivalent relation.

In the published paper [2] and by checking many words from the Qur'an [1], it was discovered that these words can be classified by the equivalent classes arises from the relation R'. So, this is a clear indicator that the Arabic language can be somehow formalized by the given relation to be as a language DNA or even a part of full language DNA. So, from the Qur'an and from Al-Rawajfeh's discovery in [2], we can formalize the Language DNA theorem or conjecture as follows.

Conjecture: Let R' be the relation as it is shown above. Then the quotient set $\sum^*/R' = L$.

The inversion of the letters, according to the LDNA finding, gives the corresponding word that has the same meaning. For example, the word "KaBoRa" which means "grown up or aged" in the verse 2:266, can be inverted by Arabic LDNA to the corresponding words, as shown in Table 1. The reader can be referred to [2], for more examples on every set of letters and its inversion.

Table (1): The application of the Arabic LDNA to the word "KaBoRa" which means "grown up or aged."

The inverted letter	The new word	The meaning
K → θ	θaBoRa	Mortal or exhausted
B → M	KaMoRa	Covered (by white hair, by weakness)
R → ḍ	KaBo ḍa	Either disappeared or can be used with the language extension
K → θ and B → M	θaMoRa	Yielded or generated (kids)
K → θ and R → ḍ	θaBo ḍa	Discouraged, feeble, thwarted
B → M and R → ḍ	KaMo ḍa	Either disappeared or can be used with the language extension
K → θ, B → M and R → ḍ	θaMo ḍa	Softened, weakened or become thin (his skin or bone)

3. Conclusion

A part of the so-called Arabic Language DNA (LDNA) was formulated mathematically. This work is important in the road to discover a full LDNA of the Arabic language and introduce it to other languages for possible application. By inverting at least one letter of the word, another word with similar meaning can be produced. The produced word is either in use, or available in the dictionary but not in use or it can be considered as a future word that has a potential possibility to be used in future applications according to the language extension principle.

References

[1] The Holy Qur'an.

[2] A. Al-Rawajfeh. New findings: Inversion and language DNA theory. Journal of The Iraqi University, 47 (1) (2020), 278-284.

[3] W. Rautenberg, A Concise Introduction to Mathematical Logic, Springer Science+Business Media, (2010).

[4] S. C. Reghizzi, Formal Languages and Compilation, Texts in Computer Science, Springer, (2009).