E-Fluency Measurement of Arabic Scripts

Mais Saad Safoq

Computer Science Department, Science College, Karbala University, Karbala, IRAQ

Abstract: Fluency is the ability to speak or write a sentence easily and accurately, in which speaker should be able to express himself smoothly to be understood by any receiver. Fluency of context of the Arabic text depends on many standards to be accepted, these standards are related to the words, combination of words in the text or the grammar context of the sentence. In this study we propose a system to check the fluency of Arabic Scripts, the system is build using a particular set of standards which represents the factors for measuring the fluency of the text. We chose four types of standards exist in Arabic Language Grammars, these are the Brutal words, Disharmony of words letters, the repetition of words and the percentage of letters in the paragraph. The system shows a good performance in detecting the non-fluency words and it works fine for all of the scripts we used.

Keywords: Arabic Language Fluency; Disharmony; Brutal Words; Percentage of letters; Appearance of words.

1. Introduction

Fluency in Arabic language is the sense of the statement clearness, in the context it is the obviously meaning of words; the familiar used words among the Arabs.

It's associated with the description of the word, speech, and the speaker. As they said: a fluent word, and the words of fluent and fluent speaker [1].

The conditions of fluency and eloquence are the concision and clearness in addition to the disposal of curiosity in speech in order to express large feelings with a few words.

In other words, the speech must be clear and evident and should not need thinking to be extracted and concentration to understand it whether it was poetry or prose [2].

The scientists of Arabic language have different opinion to differentiate between fluency and eloquence. Some of them say that fluency specialized in words and eloquence represents the sense of text. This means that the word can be well spoken and it is not consider eloquent unless the sentence explain the desired effect then it becomes eloquent speech.

They decided to set the difference between the two so that the term is said to be fluent but is not eloquent, while speech is said to be eloquent and fluent. Eloquence in more general and comprehensive, so every eloquent is fluent but not all fluent text is eloquent [2].

While fluency meant more in the integrity of speech based on grammar and is free from disharmony with the truth of its meaning. The important requirements of Eloquence is to clarify what referred to by the text, valid indication of the desired goal, good arrangement of the sentences, creativity in the way of simile acting and then brief details of the subject, taking into consideration putting the separation and connectivity marks where they are needed, all of that should be considered under the condition of eloquence [1].

2. Literature Review

Many researchers have been investigated the fluency of Arabic language, this researches were focused on the oral reading fluency for individuals. There are many indicators for measuring the reading ability of an individual; this measurement is mainly based on a human interactive and their ability of understanding the reader without investigating the real fluency rules in Arabic language.

In [3], the experiment was performed on a selected group of students in the school to measure students' ability to read and by measuring the language fluency. The examiner prepared a test for the students to examine the oral reading rate by measuring the speed at which the child reads words aloud by calculating the total number of actual words read per minute, and the oral reading accuracy in which the students should be able to recognize word in order to get a proper comprehension of a text. Additional factors are used such as prosody and oral reading comprehension in order to investigate the impact of bilingual education on students' oral reading fluency in Arabic.

In [4], the researcher focuses in the reading fluency of the first grades students since it represents a crucial period for revising student's skills for reading accurately; screening in early grade is important to determine students in danger for failing to develop their skills and the fluent word recognition and reading connected text, and this is crucial for reading success in later grades.

The research is based on two basic skills factors in identifying the growth of first-grade language reading skills; it compared two curriculum-based assessments of letter knowledge: Letter Naming Fluency (LNF) and Letter Sound Fluency (LSF) in the Arabic language.

In [5], the research suggests that naming speed ability is a key component in reading acquisition in Arabic. This study revealed a significant relationship between naming speed measures and reading fluency in Arabic language. The measurement of the test is comprised of two subtests which are based on automatical naming: one for letters and

one test for numbers, the children were supposed to name the five letters and digits preceded the timed trial.

3. Fluency of Arabic Text

The Fluency is the simple expression for the subject with smooth words that can be understood by any listener. Language fluency is the degree to which one is consider a fluent in a language; it means the flow in which sounds, words and phrases are joined together when speaking quickly [6].

The fluency can be determined in three forms, in term of word fluency, speak fluency or text fluency [1, 7].

Word fluency is the pure word which should be exclusive of the flowing conditions:

Disharmony of the characters, that are not heavy on hearing the word or difficult on the tongue to spell it, like (هعخع).

The strange use of the word, when the word is not understandable or not used in Arabic language so that it will be difficult to understand the intended meaning, like (تكأكأتم , مسرّج , مُسْتَشْزِرَات, مُسْحَنْفِرَة , بُعَاق , مُشْمَخِرّ).

If it is contrary to measurement: word to be abnormal, unlike the morphological law contriver from the words of the Arabs.

It is hatred to be heard, when the word is a brutal, rejected by the ears, same as it reject annoying sounds, like (الجرشى , جَرْدَحْل , جُحَيْشُ).

In conclusion, if there was something in the word of these four conditions, the word is considered not fluent, eloquent people must avoid these things. [1]

Fluency of speak or text, text can be fluent if it was pure and eliminated of the following things:

If the text includes non-fluent words, it considered not fluent.

The combination of words in the text has disharmony. When its words are not harmonious, this make it heavy for hearing and make it difficult for pronunciation, for example (وقـبرحـرب بـمكان قفـرُ وليس قـربَ قـبر حـرب قـبرُ.).

The weakness of its grammar, the context text/speech does not apply the grammar of Arabic language.

Repetition of the word, repeating a single word twice or more in the same text line (this is happened in poor poetry).

It is the complexity of verbal, when the words are arranged unlike the order of meanings.

 Intangible meaning Complexity: that the meaning is hidden in the structure of context due to poor simile needed to the realization of the human mind in order to understand the meaning [1].

Percentage of the letters in the text, scientists divided the list of characters into sets and each set should appear in limited percentage in the text [8].

The fluency speakers must be able to express the meaning, with fluent words, and this ability is acquired in long time training and exercises [1].

Disharmony of text: The language scientists have been agreed on a condition for the word to be plausibly acceptable, the word should be free of repulsion letters: disharmony the letters may came from diverging outlets, for example one letter is given orally from the lip, and the other from the throat, like "بعاق". It may also caused by the close outlets letters, word such as "هعخع", here we can note that the "الهاء" and "العين" or the whole letters of the word "الهاء" and "العين" and "الخاء" the whole annular convergent, and therefore it became difficult and heavy. But most scientists agree that good taste is the major talent to know the level of disharmony in the words, and the evidence for that is the "Quran" [9].

4. The Proposed System

At the mention of fluency one can think about good talking and speech, but in fact fluency could be measured in one word or in a written text paragraph. In this research, we have designed a system to measure the fluency for the written text based on a specific set of known standards for fluency in Arabic language, the diagram of the system is shown in figure 1.

We have chosen four standards of fluency rules in Arabic language; they represent the main factors affecting the fluency in Arabic script, next we will list these set of factors and their representation in the system.
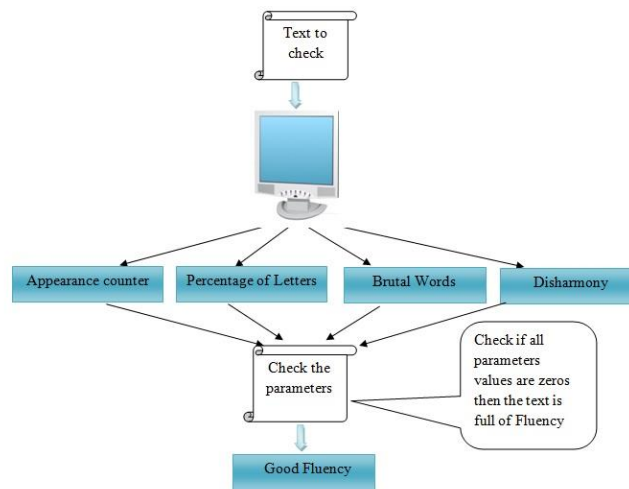
Figure 1: The Fluency system diagram

4.1. Percentage of characters appearance

The scientists of Arabic language divide the characters of the language into seven groups, each of these groups have standard limits that must not exceeded in the text [8]. These groups are:

(ل ، م ، ن ، أ)   its limit is 70%

(هـ ، و ، ت) its limit is 50%

(ب ، ك ، ر ، ف) its limit is 30%

(ع ، ق ، س ، ذ ، د ، ج) its limit is 20%

(خ ، ح ، ش) its limit is 15%

(ص ، ض ، ز) its limit is 10%

(ظ ، ط ، ث ، غ) its limit is 5%

The program trace the text and count this percentages for each group, normally this parameter should appear more in speaking, so a user should recognize the defected sentence if the program detect a high percentage of these groups.

4.2. Repetition of the Word in the Text

The word should not being repeated in the text more than two times; this could make the meaning of the text weak and undesired, such concept usually noticed in the Arabic poetry. The program search the text to count how many times each word repeated in the same sentence, and return the counter of appearance if the word occurred more than two times in the text.

4.3. Brutal and Strange Words

The use of brutal and strange words should not appear in the text; it sometimes represents impoliteness and lack of decency and may cause distracting for thoughts because of the need to search for its meaning in a dictionary or references. In this parameter we suggested to search and highlight such brutal/strange word if it found in a sequential order in the text.

4.4. Disharmony

The disharmony in the Arabic word occurs when the letters of similar sounds and closed outlets or diverge outlets [2], this may cause the word to appear in disharmony for the listener. In this section the heavy letters that could cause disharmony are divided into five groups according to its location of outlet [10,11]. The proposed algorithm works by collecting the permutations of the letters in the group and then the text presented is tested using the set of permutations collected for each group. Figure 2 illustrates the groups of letters which are classified based on the location of their outlets; each set contains the letters which has close outlets, these groups are:

أ , ه , ع , ح , غ , خ

ت, ض , د , ط

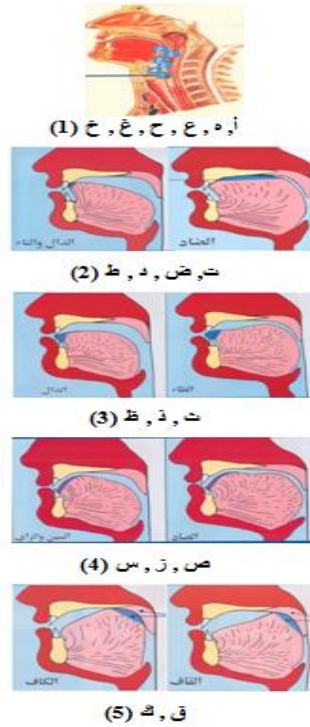ث , ذ , ظ

ص , ز , س

ق , ك

Figure 2: The groups of letters based on the outlets

The algorithm is as follows:

Read the text's paragraphs

Loop for each set of letters

Collect the permutation(letters)

Search for permutation in the text

If (the word contain the permutation) then

the word has disharmony

5. Experimental results

The proposed system has been evaluated using different Arabic Scripts, every time the presented text is tested to check the existence of the four factors. In this section we illustrate the result of three different scripts after testing it in the system.

The text is an essay taken from a journal (figure 3); the result was accepted by all of the fluency factors except for the repetition of the words, the system has collected three different words that were repeated more than three times in the paragraph.

Figure 3: The result of Text 1 (an essay of a journal)

The second experiment was achieved by testing a sermon from the (نهج البلاغة)book, after checking this sermon of Imam Ali the result that it is full of fluency and all the parameters was inexistent see figure 4.

Figure 4: The result of Text 2 (a sermon for Imam Ali)

The last result was achieved using a script taken from a book that talks about fluency in Arabic language and it actully contain non fluence words as examples. The system was able to detect the influence words that affects two different parameters of the system (these words was higlited by the system), one word represents disharmony word and the other words are a brutal words, figure 5 illustrates the result of this script.
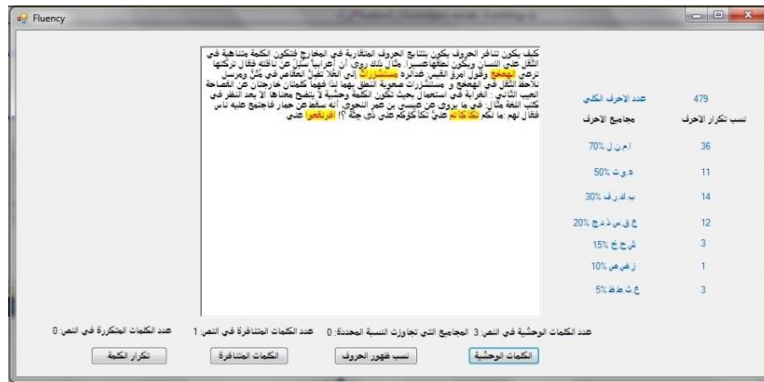

Figure 5: The result of Text 3 (text from a book of fluency)

As we can see in these three samples the percentage factors has not been detected in any of the samples, the percent actually countered differently but none of the groups exceeded the expected limit, in figure 6, a comparison of this factor of two of the samples are shown. As it shown text 2 has detected less percentage for all of the groups than text 1.

Figure 6: percentage of the seven groups

Figure 7 shows the results of the experimented samples, as shown in this plot text 2 has represented in straight line as it indicating no non-fluency words, while in the other two texts the factors have been affected by detecting non-fluency words.
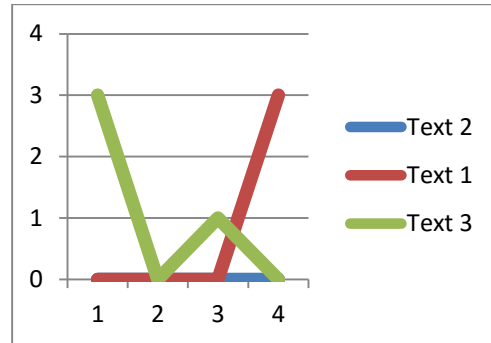

Figure 7: The plot of the experimented samples

6. Conclusion

The Fluency in the Arabic language represents an important indicator for setting up the text, and they affect the meaning and significance of the words. In additional the fluency represents the taste for choosing the words. In this research four of the existing factors are used for measuring the fluency of Arabic script, the system shows a good performance in checking the presented text and it was able to detect any not fluency text.

References

Abo Alfotoh mohammad ghanem , Arabs chamber , Fluency and Eloquence,

http://www.diwanalarab.com/spip.php?page=article&id_article=15513 , (24 /10/ 2015).

AlKafajy Abdulla bin sinan, Secret of Fluency, scientific books publishers, 1982, 84.

Abdelaziz M. Hussien, 2014, The Indicating Factors of Oral Reading Fluency of Monolingual and Bilingual Children in Egypt, Published by Canadian Center of Science and Education, Vol. 7, No. 2, 75-90.

Al-Hmouz Hanan, Mutah University, 2013, THE RELATIONSHIP BETWEEN LETTER FLUENCY MEASURES AND ARABIC GPA, INTERNATIONAL JOURNAL OF SPECIAL EDUCATION, Vol 28, No: 3, 140-149.

Raphiq Ibrahim, 2015, How Does Rapid Automatized Naming (RAN) Correlate with Measures of Reading Fluency in Arabic, Psychology, 6, 269-277.

Wikipedia, Fluency, https://en.wikipedia.org/wiki/Fluency, (25/10/2015).

Amin Loqman Aljabar, Mossel University, The Standard of Regular and Auditory Fluency in AlMozhir book of Alsuty a criticism study, 2007.

Ali Mansour Khialy, Fluency and Eloquence,

https://www.youtube.com/watch?v=broYytUCkjI, (25/10/2015).

Fluency and Eloquence,

http://islamacademy.net/media.php?Item_Id=3338&parentid=460, (25/10/2015).

https://www.youtube.com/watch?v=CX9mYG4HGcg, The Arabic letters which has similar sounds, 12/12/2015.

https://www.youtube.com/watch?v=BZZI6W3bDcA, Pronunciation of Arabic Letters, 12/12/2015.