A New Scientific Formulation of Tajweed Rules for E-Learning of Quran Phonological Rules

Yahya O. Mohamed Elhadj¹, Mohamed Aoun-Allah², Imad A. Alsughaiyer² and Abdallah Alansari³

¹Information Technology Deanship, ²College of Computer & Information Sciences, ³College of Arabic Language, Al-Imam Muhammad Ibn Saud Islamic University, Riyadh, Kingdom of Saudi Arabia

1. Introduction

Recitation of the Holy Quran is governed by a variety of rules called "Tajweed rules" (Correct pronunciation of the Holy Quran). Reciting the Holy Quran in the appropriate way is very important for all Muslims and is indispensable in Islamic worshipping such as prayers. So, teaching how to recite it correctly was transmitted, since its revelation to the prophet (PBUH), orally from teachers to learners throughout generations. Such a method has been considered as the only way to learn it until the twentieth century, where technology produced recording systems and electronic devices that are able to keep both text and sound of the Quran with tajweed rules. Since then, it becomes possible to listen Quranic recitations recorded from authentic reciters. Many computer-programs have been then appeared to assist novice learners by listening while following corresponding text on the screen. However, efforts spent by the computer scientists in general for the sake of the noble Quran are still limited and have been concentrated only on the direct application of the Information Technology techniques, such as storing, listening, searching, etc, without using more elaborated techniques in the domain.

This work is a part of a project aiming to build a computerized-environment for learning the Holy Quran and its sciences (Computerized teaching of the Holy Quran "CTHQ") (Elhadj et al., 2010a). Our objective in the CTHQ project was to improve computerization of the Holy Quran by introducing advanced techniques and methodologies. Four main tracks have been designed to carry out this project. In the first track, an environment for teaching how to memorize the Holy Quran in a manner similar to the usual way followed in the Quranic incubulation schools was proposed (Alsughayir & Elhadj, 2006; Elhadj, 2010). In the second track, automatic speech recognition technologies have been used to teach how to recite the Holy Quran correctly (Alghamdi et al., 2007; Elhadj et al., 2012, 2010b, 2009; Elhadj, 2009a, 2009b). In the third track, techniques for determining the similarity (tashaboh) between verses (ayahs) of the noble Quran were investigated (Alsughayir & Ohali, 2007). Computer-tools have been developed for analyzing the text of the noble Quran based on complete words and their stems in order to link similar verses (Alsughayir et al., 2009; Elhadj et al., 2009a, 2009b).
In the fourth track, our focus was on the automatic processing of tajweed rules by proposing a mathematical formulation of these rules that can be easily processed by a machine (Elhadj & Aoun-Allah, 2011; Elhadj et al., 2009e). In this chapter, we discuss how this kind of formulation was proposed, its advantages and benefits, and how it was programmed in an efficient plug-in engine. The development of an e-learning system for tajweed rules that uses this engine will also be presented. A general integrated environment for self-learning of the holy Quran and its sciences including Tajweed will be also briefly introduced.

This chapter proceeds as follows. Sections 2, 3, and 4 deal with Tajweed where the first one presents an introduction. Section 3 surveys previous efforts to computerization of Tajweed whereas Section 4 concentrates on our new formulation. We discuss in Section 5 a prerequisite to our engine of automatic processing of Tajweed rules, which is detailed in Section 6. Our Tajweed Learning System (ETaj) will be presented in the Section 7 and a couple of its important interfaces will be given in Section 8. Finally, before concluding this chapter, we shortly present a global learning environment for Quran and its sciences composed of several subsystems including Tajweed teaching system.

2. Overview of Tajweed rules

Reading of the Holy Quran is quite different from the reading of a normal Arabic text due to the pronunciation rules that have to be respected during the recitation. As an example, consider the duration of vowels that may vary from 2 to 6 times that of a single consonant depending on the context. Indeed, there are two kinds of vowels in the Arabic language, short vowels which are keystrokes placed either above or below the preceding consonant and long vowels which denotes a certain repetition of short vowels. There are also other tajweed rules related to:

a. types of nasalization (called "ghunnah"),
b. heaviness and lightness (called respectively "tafkheem" and "tarqeeq"), which means making some sounds emphatic or non-emphatic,
c. types of stops which means making a voiceless break at a Quranic word for a brief moment (called "waqf"),
d. degrees of vibrations or unrest (called "qalqalah") which means producing the voiced stop consonants with a schwa-like sound at the end,
e. etc.

To preserve the Holy Quran from any alteration in its pronunciation and then to guarantee its perfect reading, early Muslim scholars described Quranic recitation methods very accurately by textual rules as well as sound ones. These descriptions are classified into categories depending on some sound properties or features. These categories are often summarized as follows:

a. category "Noon sakinah" & "Tanween" (باب النون الساكنة والتنوين)
b. category "Meem sakinah" (باب الميم الساكنة)
c. category "Qalqalah" (باب القلقة)
d. category "Tafkheem" & "Tarqeeq" (باب التفخيم والترقية)
e. category “Al Mad” (lengthening)
f. etc.
Each category has a set of rules governing the pronunciation of underlying sounds. More details regarding the classification of Tajweed rules can be obtained from the official site of King Fahd Complex for the Printing of Holy Quran at this link: www.qurancomplex.org.

3. Previous efforts related to computerization of Tajweed

Teaching how to apply Tajweed rules during Quranic recitation has been done through teachers who pronounce the Quranic sounds accurately. With the era of computers, it becomes possible to computerize the learning process of Tajweed, but this need an appropriate description of its rules. To the best of our knowledge, there was no attempt to describe them by a scientific algorithm, which could be processed automatically by a machine. This was a conclusion from a large survey we conducted on the Quranic software currently available on the market, which was published in (Alsughayeir & Elhadj, 2006). The methodology employed to conduct this study consisted of collecting a maximum number of Quranic applications, either for desktop, web-based or hand-held applications. Each application was separately studied and evaluated in order to know the characteristics and services it offers.

It appeared that softwares developed for the sake of the Holy Quran are still very limited either in their objectives or in the term of availability and relevance of features. Regarding the Tajweed by itself, it was only given as a small part of the surveyed programs. It is offered, at almost all programs, in a classical manner as textual lessons with some graphical explanation of sound production. No dynamic interaction was proposed to allow, neither efficient learning nor any kinds of intelligent processing.

4. New formulation of Tajweed rules

In order to find a scientific formulation of tajweed rules, we studied them thoroughly with the assistance of an expert in the domain and finally we came out with the conclusion that almost all Tajweed rules described textually by scholars could be written in a scientific manner and consequently automatically treated by a computer.

Our study concluded that to extract the tajweed rule for any letter the maximum number of words concerned by the rule is two; i.e. either the previous or the following word. And in terms of letters, there are at most 6 letters concerned by the rule which are either preceding or following the letter itself. Moreover, many rules are applicable only to a letter at the end of a word which is easily detectable by comparing the next letter to a space character.

For instance, in the category of “Noon Sakinah” & “Tanween” if any letter “Noon Sakinah” or diacritic “Tanween” appears at the end of a word followed by any of these characters "ي ى ى ى ى ى" in the following word then the 2 letters should be assimilated. Moreover, this category is subdivided into two sections where the assimilation could be with “Ghunnah” (a sound effect) or without when the letter is followed respectively by “ي ى ى ى ى ى" or “ي ى ى ى ى ى". We proposed a whole scientific formulation of such textual description of rules into machine readable rules as indicated in the example in Table 1 taking in consideration the following notations:

a. “C” means the character treated and “S” its diacritic
b. “C+1” or “C-1” represent respectively the character (or characters separated by a comma) immediately following (rep. preceding) the character treated.
c. And in the same way we define “C+2”, “C+3”,... or “C-2”, “C-3”,... according to the extent of the rule.

d. “*” is used as a wildcard character to replace any letter or diacritic.

e. “Space” represents a space character (between two words).

f. “Text” is human comprehensible message to be displayed to the user if needed.

And finally, the “Ruling” of the rule which denotes its name.

<table>
<thead>
<tr>
<th>C = * ; S = ٰ;</th>
<th>C = ن ; S = ۳;</th>
</tr>
</thead>
<tbody>
<tr>
<td>C+1 = Space; S = *;</td>
<td>C+1 = Space; S = *;</td>
</tr>
<tr>
<td>C+2 = ي،ي،ي،ي； S = *;</td>
<td>C+2 = ي،ي،ي،ي； S = *;</td>
</tr>
<tr>
<td>Text = یُدغم التنوین بعنة في حروف (وي،وي)</td>
<td>Text = تدغم التنوين الساكنة بعنة في حروف (وي،وي)</td>
</tr>
<tr>
<td>Ruling = IdghamGhunnah</td>
<td>Ruling = IdghamGhunnah</td>
</tr>
</tbody>
</table>

(Table 1.A) (Table 1.B)

<table>
<thead>
<tr>
<th>C = م ; S = ۳;</th>
<th>C-1 = * ; S = ۳;</th>
</tr>
</thead>
<tbody>
<tr>
<td>C+1 = Space; S = *;</td>
<td>C = ی； S = *;</td>
</tr>
<tr>
<td>C+2 = م ； S = *;</td>
<td>C+1 = Space ; S = *;</td>
</tr>
<tr>
<td>Text = یُدغم العيم في الميم ادغامًا شفوية</td>
<td>Text = یُدغم التنوین بعنة في حروف (وي،وي)</td>
</tr>
<tr>
<td>Ruling = IdghamGhunnah</td>
<td>Ruling = IdghamGhunnah</td>
</tr>
</tbody>
</table>

(Table 1.C) (Table 1.D)

### Table 1. Rules describing a couple of cases of assimilation with Ghunnah

Table 1.A (resp. Table 1.B) describes a rule stating that if any character with a “tanween” diacritic (resp. a “Noon” character with “Sakn” diacritic) at the end of a word followed by a word starting by any of “ي،ي،ي،ي” letters, then there is assimilation with Ghunnah. Table 1.D describes a special case of the rule in Table 1.B where “tanween” is followed by a “lengthening” at the end of the word. Finally, Table 1.C deals with a “Meem” character with “Sakin” diacritic as last letter in a word followed by “Meem” character with any diacritic as first letter in the next word and where the ruling is the same as previously.

Using this technique of writing “tajweed” rules, we finished describing those of 5 most important categories (or chapters): “Al Mad” (lengthening), “qalqala” (unrest), “Noon Sakinah” & “Tanween”, “Meem Sakinah”, and “Tafkheem” & “Tarqeeq”. This results in a total number of almost 200 rules. Moreover, any advanced tajweed rule could be easily added by writing it as described earlier in the rules text file.

Notice that we faced many difficulties in writing our rule set as described above. The most important difficulty was the overlap of rules. This situation happens when more than one rule could be triggered for one letter. For instance, in the word “اسماء” we have a lengthening of obligatory 4 cycles but when it is the last word in a verse it changes to optionally 2, 4 or 6 cycles. To overcome this overlap, we added to each rule a priority in order to avoid the conflict between triggered rules and by choosing only the rule with the highest priority.
5. Need of a textual version of the Holy Quran with full diacritics

The Holy Quran is generally written in a special font which could not be edited in text editor software. Consequently, a text editable version of the Holy Quran that contains no missing diacritic was necessary to get our automatic processing works correctly. However, this textual editable version as we need it was almost inexistent. Therefore, we looked for the most agreed version to which we manually added full diacritics and then it was given to scholars for revision and validation.

6. Automatic processing of Tajweed rules: Taj Engine (TajE)

The Tajweed rule set described in section 4 is interpreted by an inference engine (TajE) able to deal with our rulings formulation and that can be used alone (See Figure 1) or integrated with other components for both identification and verification of Tajweed rules in any Quranic verse (or even any Arabic diacritized text).

Fig. 1. Structure of the Taj Engine

This inference engine works like an expert system inference engine in forward chaining and one level inference. In other words, for each character of the treated text, our engine tests all rules once to find the rule or those rules that could be triggered. Conflicting rules are treated by the rule priority explained in Section 4.

As presented earlier, TajE engine could be used either for identification or for verification of Tajweed rulings. The first option (identification) aims to extract Tajweed rulings from Quranic verses. It is very helpful in different situation, especially to assist a student during his memorization of the Holy Quran.

The second option (verification) of the TajE engine is to ensure if a Tajweed ruling really occurred in a specific place in the Quranic verse (See Figure 2). This is also a very important property as it can be used in an interactive manner for teaching and correcting Tajweed exercises. A detailed typical exercise scenario is described in the following section.
Our main objective is to design an easy to use e-learning system for Tajweed using our developed TajE engine. It is intended to be used by students to help them learning, verifying their knowledge, and to train them on Tajweed for the noble intention of correctly reciting the Holy Quran.

The proposed system has been designed to include the following features:

a. **General Rules of Recitation and its Ethics:** these are stored guidelines about general rules of recitation like position of stops and their rulings. Reading ethics like “Estiatha” and “Basmalah”, etc. are also included.

b. **Tajweed Rulings:** this option offers to the student rich text lessons about Tajweed rulings along with examples in text and audio formats. The student could interact with by pausing, stopping or replaying the audio files.

c. **Exercises:** they cover all Tajweed categories and are taken by student using TajE engine described in the previous section. A typical scenario is given below. Another option for examination is offered to students to pass examinations in a manner very similar to exercises but in a less helpful way and by assigning a final mark to the learner.

Exercises and examination options of ETaj system are implemented as follows:

a. A student starts by choosing the Tajweed category he wants to practice or be tested on.

b. According to his choice, all rulings under the category will be displayed with a specific color associated to each ruling.

c. Besides, a verse or a couple of verses chosen randomly will be displayed too. We notice here that the selection of verses is totally random among all Quran verses and having a fixed minimum number of the selected category rulings. This property offers to our system a great dynamic aspect. The selection process is either done on the fly by searching the Holy Quran using first option of the TajE engine to extract verses with the minimum number of rules, or by selecting them from a database of ruling occurrences already filled using our TajE engine beforehand. Obviously, the first technique is slower than the second but requires less storage space. However, we chose the second
technique for the simple reason that is the fastest way since a non-negligible treatment is done beforehand and also to guarantee a good response time for the system.

d. At this moment, the student is invited to select letters where Tajweed rulings appear. This process is achieved by the learner by choosing a color (i.e. a category ruling) and then selecting the appropriate letters. The student is free to change or to remove his selections until he decides to validate his choices.

e. Once validated, choices are treated automatically by TajE engine and a report is displayed to the student containing the incorrect and missing selections along with the necessary explanations (i.e. the field “Text” in the rule, See section 4).

Figures 3 and 4 show respectively the structure of the proposed ETaj system and its architecture.

Moreover, we aim to add to the system a new feature allowing the user to specify the “narration” which will influence all the above features. We mention that our current system supports only “Hafs” narration.

8. Most important interfaces of the ETaj system

ETaj has been designed to be easy to use and to allow great interaction with the user. It is implemented in the .Net framework using C# and JavaScript as well as other advanced technologies. Our database is implemented using MySql as a database management system as it is free and it offers very excellent features and performance. The most important interfaces of the system are introduced in the following sections.

Fig. 3. Structure of the ETaj system
Fig. 4. Architecture of the ETaj system

If we access the system, the following menu will appear to the user:

Fig. 5. Main menu of the ETaj system
From this main menu, we can navigate to different options in the system. If we click, for example one of the tajweed rulings, a page like the following is displayed:

Fig. 6. Tajweed rulings page

Three main parts are enclosed in this page: one for the tajweed category and its rulings that can be navigated one after another (Figure 6 (a)), the content of the current ruling (Figure 6 (b)), and examples of the ruling (Figure 6 (c)):

Fig. 6(a). Tajweed category and its rulings
Now, if we click on the "exercises" option, we will be redirected to a page where we can choose the category of tajweed we want to practice (Figure 7):

Once the category is selected, a page like that of Figure 8 is displayed and the user is asked to navigate the category rulings and to select their occurrences in the ayah that has been randomly chosen. This is done in the same manner as explained previously in section 7.
After finishing determining places of ruling occurrences, a page result (see Figure 9) is displayed.

Fig. 9. Page Result


The Tajweed learning system is a part of a fully integrated environment developed for self-learning of the holy Quran and its sciences. This environment comprises four subsystems in addition to the ETaj system: subsystem for learning how to recite and memorise the holy Quran, subsystem for studying the similarity between Quranic terms and verses, subsystem providing the most relevant features that learners of the holy Quran may need such as (Erab"، Nuzoul”， Tafseer”، Ejaz"، الاعجاز العلمي“، التفسير“، إجازة”، etc.), and finally a reach library providing authentic sources from which the relevant features are taken. The three main subsystems will be briefly described in the following sections to get picture of this environment.

9.1 The memorization subsystem: E-halaga

The E-halaga system is designed based-on the Quran memorization halaga “ حلقات تخفيف القرآن” approach (E-halaga means electronic halaga). It simulates the real one in having tutors,
supervisor, and registered students. E-halaga allows four main types of users with specific roles: administrator, supervisor, tutor, and student. Roles are distributed hierarchically to allow a great flexibility over the system. The creation of the E-halaga is done by its supervisor, which is in turn created by the system administrator. The supervisor is responsible of running the E-halaga in terms of adding or deleting tutors, distributing learners, etc.

Since the E-halaga is a simulation of the real halaga, the learner have to specify a daily amount (memorization section) to be memorized starting from somewhere in the Quran (starting ayah/page). As the learner progresses, he needs to have in parallel a revision program for the memorized parts to improve their quality (revision section). So, the registration in the E-halaga system requires specifying the following parameters: starting ayah/page, memorization section and its length, revision section and its length (at least the double of the memorization section), and previously memorized sections before joining the system.

Once registered, the learner will obtain a user name and password to access the system. If he logs into the system, he will be redirected to the memorization/revision part where he can find a list of useful options. He can browse his own profile and change it as needed, measure his performance (number of sections perfected, times of failure in every section, duration of memorizing), print out his transcript, etc. The learner can listen to an ideal recitation taken from one of the famous reciters stored in the system as a reference for perfect recitation. After listening to the reference recitation several times, or reciting directly, the learner can test his memory by clicking on an icon to record the section. Then the pages of the section will disappear to test his memory. Clicking on the button of “end recording” will display again the section pages allowing the learner to verify his memorization. The learner can repeat this process (recording and verifying) till the perfection of memorization. Once the learner approves his recording, the audio file is uploaded to the server and the learner will not be able to modify it. The tutor is then notified with an unmarked recorded section of that learner. After the tutor marks the section, the learner can benefit from audio and/or textual remarks explaining his mistakes. Figure 10 shows the architecture of the E-halaga system and we invite the reader to consult (Elhadj, 2010) for more details.

9.2 The similarity search engine

The similarity search engine is a component devoted to determine the similarity between Quranic verses (and Quranic terms). To build this system, we started preparing an authentic fully diacritised textual version of the Holy Quran since such one was not available for the use in the domain of research. Next, the focus was on a manual morphological analysis of this version. Each word in the Holy Quran is split into four parts, prefixes, stem, root and suffixes, and then stored in an indexed database. Words are kept in their original context, which means Quranic verses (Elhadj et al., 2009c, 2009d). A semi-automatic environment for morphological analysis has been developed (see figure 11). Full-text searching techniques were investigated and then computer-programs have been developed for analyzing the text of the noble Quran based on complete words and their stems in order to link similar verses (Alsughayeir & Ohali, 2007; Alsughayeir et al., 2009). Figure 12 shows the engine interface.
A New Scientific Formulation of Tajweed Rules for E-Learning of Quran Phonological Rules

Fig. 10. E-halaga architecture

Fig. 11. Interface of the Semi-automatic Morphological Analyzer
9.3 The Quran sciences subsystem

It is a system letting the user navigates the Holy Quran and its main sciences. The main features offered are taken from authentic books, which were linked with the holy Quran through a full indexing of Quranic ayahs using advanced techniques and methodologies. Information are kept in their original sources (books), but can be easily retrieved and quickly processed. The books can be navigated, downloaded, etc., from a library, which represents another subsystem of this environment. Possibility of updates is well considered.
10. Conclusion

In this chapter, we firstly presented our work related to the textual specification of Tajweed rules in order to come up with a machine-readable formulation of these rules. An appropriate and easily extendable rule set has then been proposed. A fully diacritised textual version of the Holy Quran has also been prepared to be used with our new-tajweed formulations. Next, we focused on the implementation of these formulations in a fast and efficient plug-in Tajweed Engine (TajE) that can be integrated in different kinds of systems for teaching the Holy Quran and its sciences. TajE is able to handle rulings in two different ways: identification and verification. The identification option is intended to extract Tajweed rulings from Quranic verses. The verification option, may serve to ensure that a specific Tajweed ruling really occurred in a certain place in the Quranic verse, which is very useful.

An e-learning system for self learning of tajweed (ETaj) is next built on the top of the tajweed engine (TajE). It provides different options allowing learners to get maximum benefit. The first important option of the system gives general rules and ethics of learning the Holy Quran. The second option offers a well designed component for learning tajweed rulings using a convenient and attractive manner of presenting the content of rulings along with textual examples as well as sound ones. The third option is very important as it provides a full dynamic interaction with the system to practice tajweed rulings.

At our best knowledge, both TajE and ETaj systems are the first of their kinds developed for automatic processing of Tajweed in a full interactive manner. They are now ready to be used and can help mastering this important field of Quranic Sciences.

11. Future works

In near future, we plan to improve ETaj system by organizing tajweed rulings in levels with incremental degree of complexity as it is followed in the real teaching of tajweed. Learners have to go through levels one by one. At the end of each level a self examination or testing will be initiated by the system and the learner will not be allowed to go further until he masters the current level. This means that a kind of authentication needs to be added to the system to follow students in their process of learning.

As another future work, we plan to expand our TajE engine to cover other Quranic narrations as we currently covered just "Hafs" narration.

12. Acknowledgements

This work was supported by King Abdulaziz City for Sciences and Technology (www.kacst.edu.sa), as part of the CTHQ Project, under the grant number AT-25-113, Saudi Arabia. We thank all the other project team members, which actively participated in the elaboration of the whole project.

13. References


Adaptive E-learning was proposed to be suitable for students with unique profiles, particular interests, and from different domains of knowledge, so profiles may consider specific goals of the students, as well as different preferences, knowledge level, learning style, rendering psychological profile, and more. Another approach to be taken into account today is the self-directed learning. Unlike the adaptive E-learning, the Self-directed learning is related to independence or autonomy in learning; it is a logical link for readiness for E-learning, where students pace their classes according to their own needs. This book provides information on the On-Job Training and Interactive Teaching for E-learning and is divided into four sections. The first section covers motivations to be considered for E-learning while the second section presents challenges concerning E-learning in areas like Engineering, Medical education and Biological Studies. New approaches to E-learning are introduced in the third section, and the last section describes the implementation of E-learning Environments.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.