Quranic Letter Pronunciation Analysis based on Spectrogram technique: Case Study on Qalqalah Letters

Tareq Altalmas[1], Salmiah Ahmad[1], Wahju Sediono[1] and Surul Shahbudin Hassan[2]

[1] Department of Mechatronics Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia
{salmiah,wsediono}@iium.edu.my
[2]Centre for Languages and Pre-University Academic Development, International Islamic University Malaysia, Kuala Lumpur, Malaysia
surur@iium.edu.my

Abstract. Recitation of the Holy Quran with Tajweed is an essential activity as a Muslim. Reciting Quran correctly indicates the correct meaning of the words of Allah has been received from this significant resource among Muslims. That is why Muslims stress on the Quranic Education since in the early age. It is important to pronounce the letter correctly based on its characteristics as well as the articulation point of each letter. In this paper, the characteristic based on Qalqalah letters is considered to be analyzed. The audio signal from a person who is very good at Quranic recitation was taken and analyzed. We implement spectral analysis to find the features of Qalqalah letters and extract the correlation between the first formant frequency and the pharyngeal space of the signal. Spectrogram was successfully implemented and proved this relation, and it described the mechanism of Qalqalah correctly, which is unique as compared to other Quranic letters.

Keywords: Spectrogram · Qalqalah · Tajweed · Formant frequency · Vowels · phonetics

1 Introduction

The word TAJWEED means "to improve" or "to make better". It is also the rules and knowledge that help people to recite the Holy Quran as similar as what has been recited by the Prophet Mohammed peace and blessings be upon him. Moreover, an important part of Tajweed is to pronounce the letters from its correct articulations (*Makharij*) and by giving the letters its inherent characteristics (*Sifaat*) and dues in conditional characteristics. The characteristics of Arabic letters helps in differentiate letters that have the same points of articulations. The *Sifaat* of the Arabic letters are divided into two groups, characteristics with opposites and characteristics without opposites, [1].

Qalqalah is one of the most known but very difficult characteristics to be pronounced. It means bouncing or echoing the sound. The letters with this characteristic

show an echoing sound when the phone carries *sukoon* (◦). Qalqalah letters come from the strength category of Quranic letters. Strength means the points of articulation of this group of letters is completely closed when the pronunciations come with *Sukoon*. The consonants letters pronunciation happens by the collusion between the parts of the articulation of these letters. While the pronunciation of vowel letters occur by the parting between the two parts of the articulation points, [2]. Practically we can describe the Qalqalah pronunciation by a bottle that is covered with the air imprisoned inside it. When we open the bottle quickly a strong sound will come out because of the air pressure and the bottle will vibrate. Qalqalah has similar concept whereby when the letter is pronounced, the air is imprisoned for a while behind the point of articulation because of the completely closeness of the points of articulation. Then, the air will be released by the parting of the articulation points without any change of the mouth and the jaw shape. This mechanism will cause a small silent segment in the sound signal, [1].

The research on Tajweed has been conducted previously by some researchers, but still can be considered within a small associated society. The Multilayer Perceptron (MLP) has been used to investigate the process of detecting the right recitation of Qalqalah Kubra of a reader. Mel-frequency Cepstrum Coefficients (MFCC) has been implemented as a feature extraction to get the essential characteristics of the pronunciation signals. Then the MLP was trained to distinguish between the correct and the incorrect pronunciation. It is clear from the results that the classification accuracy was high, and the MLP was claimed to be successful to differentiate between the correct and the incorrect pronunciation, [3]. The author in [4] has developed a filter to remove noises when the recitation was done in a noisy environment. The authors have used Least Mean Square (LMS) based on the adaptive filter. They have only focused on 7 alphabets from the total 30. From the result, it was clear that the designed filter has removed the noise successfully. On the other hand, the researchers in [5] have also focused on the similar approach in developing a filter to remove noises in a noisy environment where the Quranic letters pronunciation was conducted. They have used the normalized least mean square based on the adaptive filter. They have also focused on 7 letters from the 30 Quranic letters. From the result, it was clear the designed filter has canceled the noise successfully.

On the other hand, Spectrogram has been implemented as well in the research involving the Arabic letters pronunciation. The authors in for example have investigated the impacts of implementing vowels (*fatha*, *kasra*, and *dhamma*) on the dental consonants. Dental consonants are the letters that the teeth are part of its points of articulations. It is divided into three groups as follows;

- Labiodental, which can occur on the lower lip and the upper teeth.
- Dental, which can occur on the tongue against the upper teeth.
- Interdental, which can occur at the tip of the tongue between the upper and the lower front teeth.

The experiment has been conducted among Malay children (non-native Arabic speakers). Spectrogram and formants have been used to find the influence of the vowels on

the dental consonants. They have found that the effect of *Kasrah* and *dhamma* can be easily extracted from the spectrogram.

The paper is organized as follows; Section II describes the process of sound production, formant frequencies and its relation to the path of the air inside the pharynx and the spectrogram analysis. Section III is described the experiment and the spectrogram results. Section IV explained the conclusion and the recommended future works.

2　Phonetics production and analysis

Phones are produced by three steps. First, the process of providing the energy for the sound from the lungs, followed by converting this energy into a sound by the vibration of the vocal folds. The last step is to translate this sound from the vocal fold into an understandable speech. Fig. 1 shows the diagram of the speech system from the energy generation until the filtering of the speech in the vocal tract, where the final stage of the system is the vocal tract. The vocal tract works as a filter to shape the signal from the vocal folds which contains many frequencies into a new signal with formant frequencies and to be sensible to human. When the air flows from the vocal fold through the oral cavity, the tongue, lips, teeth, and various regions of the mouth configuration and shape help in filtering the sound into its formant frequencies, [7].
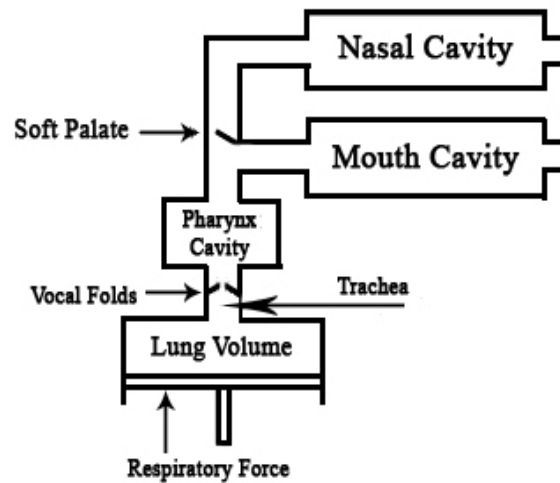

Fig. 1. The diagram of the speech system

The spectrogram is a visual illustration to represent the frequency components of the speech signal. It is considered one of the applications of the Fourier transform. Fig.2. Shows the steps to calculate the spectrogram graphs [8]. Spectrogram is a time-frequency representation where the horizontal axis of the spectrogram represents the

time of the signal while the vertical axis represents the frequency range and the color map on the graph show the level of the frequencies that is representing the signal.
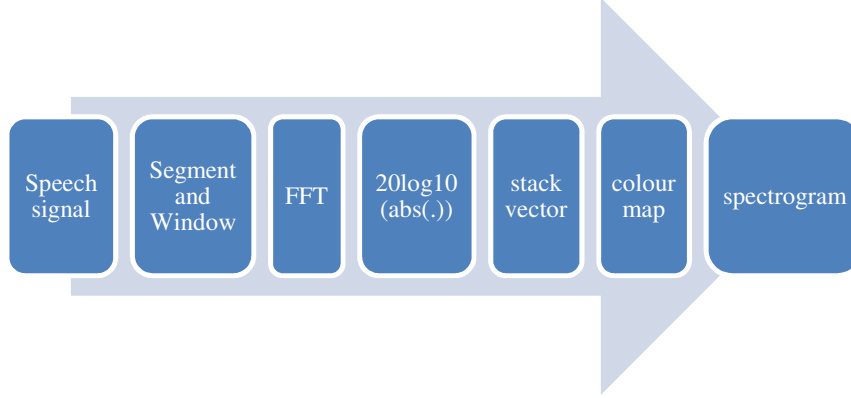

Fig. 2. Flow for calculating the spectrogram of speech signals

3   Experimental Performance Analysis

In order to calculate the spectrogram of Qalqalah letters, the first step was collecting the pronunciation data in terms of audio signals. These data has been recorded from a person who is a specialist in recitation Al-Quran with Tajweed, where the recording process was conducted inside a studio room to reduce the environmental noises. The five letters of Qalqalah which they are gathered in one Arabic sentence "ق - ط- ب ـ ج - د" have been recorded, and the silence segments have been removed by using Matlab software. The data was then normalized such that the amplitude varies from 1 to -1. The short-time Fourier transform (STFT) has been implemented to calculate the spectrogram of the audio signals. Table 1 summarizes the configurations of the STFT that was used in this paper.

Table 1. Parameters of the short-time Fourier transform (STFT)

Sampling frequency	44100 Hz
Windowing	Hamming
Window size	2048 samples
HOP SIZE	512 samples

Fig.3 to Fig.7 illustrates the spectrograms of the Qalqalah letters when it is "un-voweled" with *Sukoon*. From the graphs it is clear that the letters have three segments as follows;

- The voiced sound which is clear and in a dark color and it has formants.
- The silent segments followed the main sound which represents the period when the articulation is completely closed and the air is imprisoned inside it.

- The last period is representing the burst pattern that should happen to make the correct pronunciation of Qalqalah.

Formant frequencies are one of the key features that can be extracted from the spectrograms. The first formant frequency is related to the tongue position, and the tongue position will affect the pharyngeal space. The more pharyngeal space leads to less first formant value. Moreover, the degree of mouth opening is related to the first formant frequency, which more opening leads to a higher formant frequency and vice-versa [9]. More investigations will be conducted for the second and third formant frequencies (F2 and F3) in order to establish strong features to be used in the classifier design at a later stage. The magnitude of the spectrum and the bandwidth can be used along with the formants frequencies to support and increase the system accuracy. Praat is open source software for the speech analysis. Formant frequencies can be established easily by using Praat [10]. Table 2 lists the values of the first formant frequency by using the Praat software. The pharyngeal space of Qalqalah letters can be divided into two groups; the first group is ق and ط which it is less space value. While the second group is ب, ج and د which it is more space. It is obvious that the value of the first group is higher than the value of the second group.

Table 2. First Formant frequency value for each Qalqalah letter based on Praat.

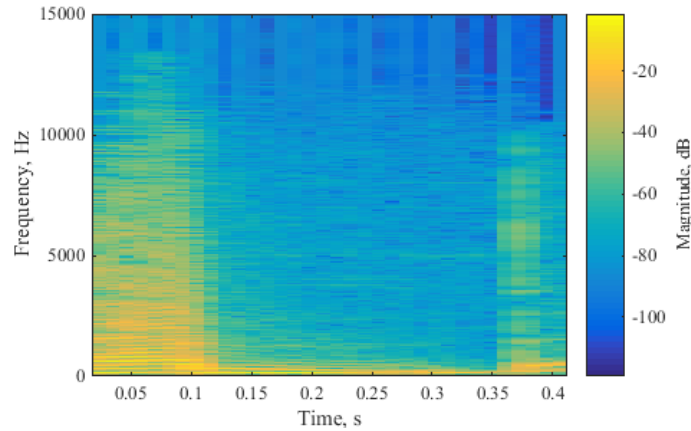| Qalqalah letter	First formant(Hz)
ب	187.457
ج	164.369
د	177.521
ط	687.939
ق	730.963

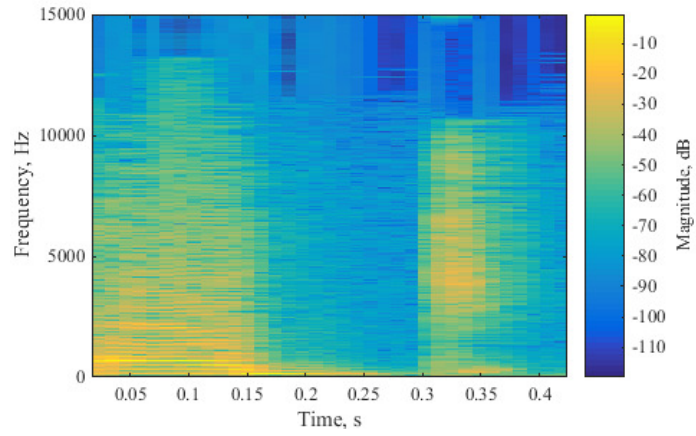**Fig. 3.** Spectrogram of the letter ب

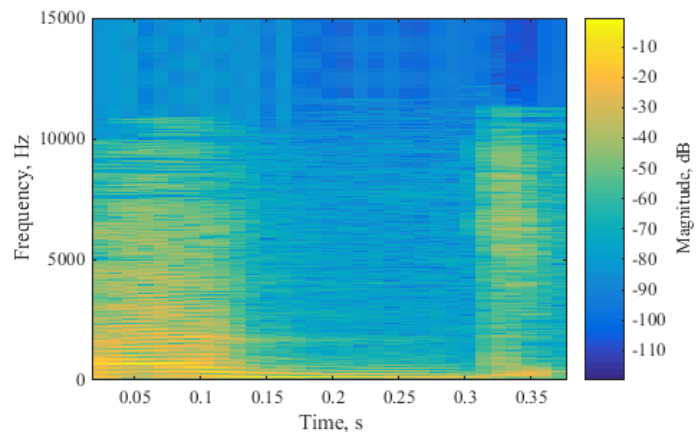**Fig. 4.** Spectrogram of the letter ح
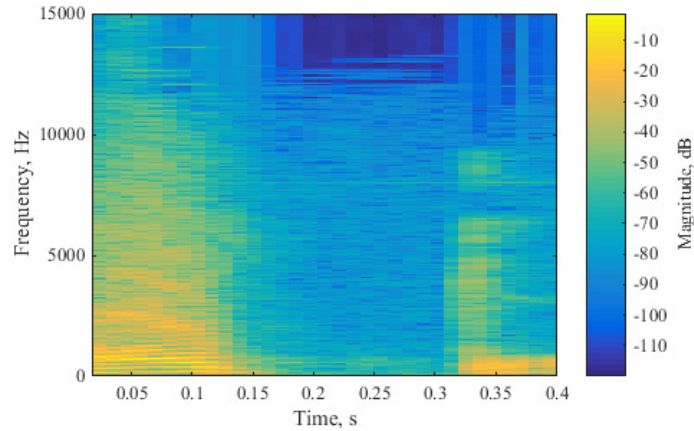

Fig. 5. Spectrogram of the letter د

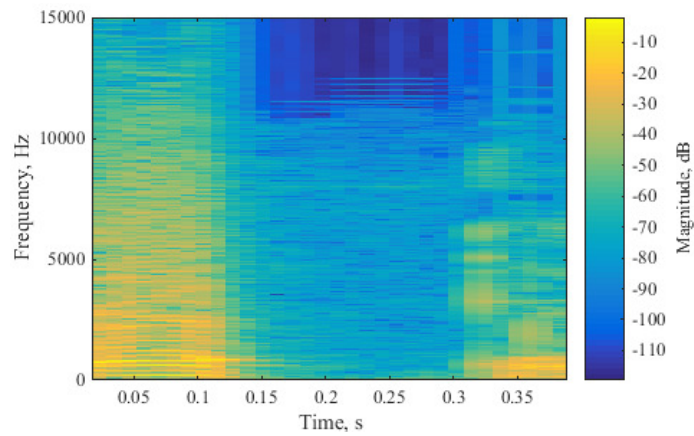**Fig. 6.** Spectrogram of the letter ط

Fig. 7. Spectrogram of the letter ق

Fig. 8 and Fig. 9 are illustrating the spectrogram of another two letters, which are ص and ل. These letters are from different groups of letters and do not have the characteristic of Qalqalah. It is obvious from the spectrogram that the three segments that are clear in Qalqalah letters are not available in other colors. From Fig.8, it can be seen that the sound ص which is produced by the collusion between the tip of the tongue and the top edge of the two front lower incisors. Moreover, the air will flow out in between the incisors and there is no imprisoning for the air. Fig.9, on the other hand, illustrates the sound ل which is produced by the collusion between the tip of the tongue and what lies opposite to it from the gum of the two front top incisors. In addition to that, the air will flow out without imprisoning.
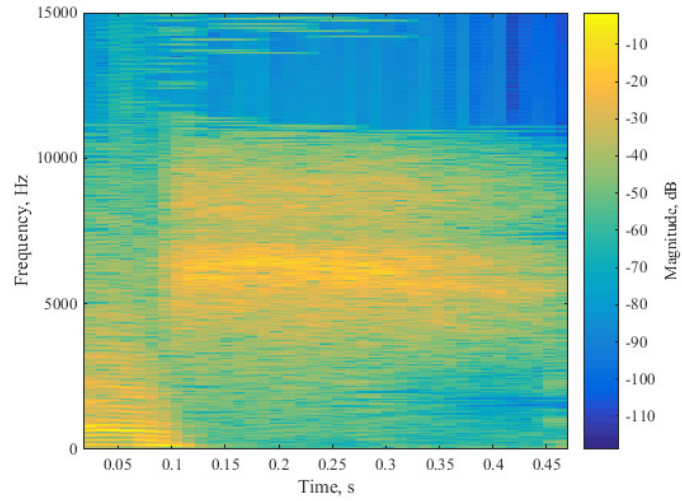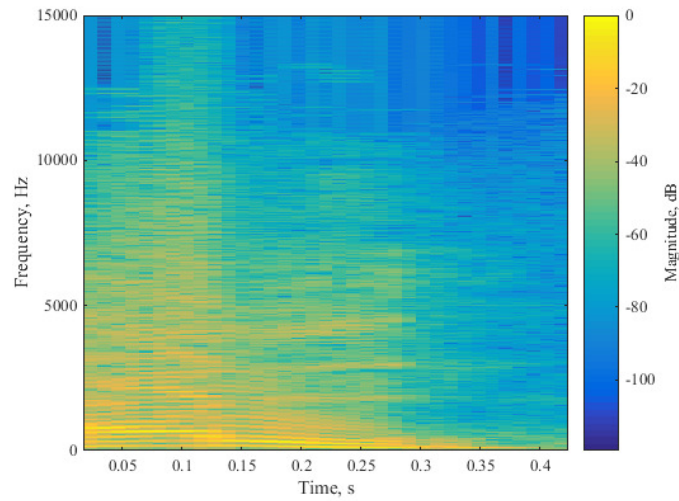
Fig. 8. Spectrogram of the letter ص

Fig. 9. Spectrogram of the letter ل

4   Conclusion and Future Works

The spectrogram has been successfully implemented to investigate the characteristics of Qalqalah sounds. This technique was an initial attempt in characterizing one of the

important *sifaat* of Quranic letters in Tajweed. The results obtained showed the relation between the pharyngeal space and the first formant frequency, which are inversely proportional. Moreover, this technique was able to illustrate the mechanism of the pronunciation of the Qalqalah sounds, where the voiced segment of the sound, silent segment and the burst segment at the end of Qalqalah phones were shown. These features will be used to classify and distinguish the correct recitation of Qalqalah and other sifaat so that the correct pronunciation of Quranic letters will benefit the reciters more.

Reference

1. About Tajweed, http://www.abouttajweed.com/
2. The Qalqalah Mechanism, http://www.abouttajweed.com/qalqalah_mechanism.htm
3. Hassan, H. A., Nasrudin, N. H., Khalid, M. N. M., Zabidi, A., & Yassin, A. I.: Pattern classification in recognizing Qalqalah Kubra pronunciation using multilayer perceptrons. In: 2012 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE). pp. 209-212. IEEE, 2012
4. Arshad, N. W., Aziz, S. A., Naim, F., Karim, R. A., Hamid, R., & Zakaria, N. F.: Speech processing for makhraj recognition: The design of adaptive filter for noise canceller. In: 2011 7th International Conference on Information Technology in Asia (CITA 11). pp. 1-5. IEEE, 2011
5. Arshad, N. W., Aziz, S. A., Naim, F., Karim, R. A., Hamid, R., & Zakaria, N. F.: Speech processing for makhraj recognition: The design of adaptive filter for noise canceller. In: 2011 7th International Conference on Information Technology in Asia (CITA 11), pp. 1-5. IEEE, 2011
6. Abdul-Kadir, N. A., & Sudirman, R.: Vowel effects towards dental Arabic consonants based on spectrogram. In: 2011 Second International Conference on Intelligent Systems, Modelling and Simulation (ISMS), pp. 183-188. IEEE, 2011
7. Phonetics and phonology, http://clas.mq.edu.au/speech/phonetics/phonetics/index.html
8. Saeed V. Vaseghi: Multimedia Signal Processing: Theory and Applications in Speech, Music and Communications, (2007)
9. F1 F2 Handout, http://www2.muw.edu/~mharmon/501F1F2.html
10. Praat: doing phonetics by computer, http://www.fon.hum.uva.nl/praat/