Middle-East Journal of Scientific Research 25 (6): 1370-1379, 2017
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2017.1370.1379

Analysis of Formant Frequencies of the Correct Pronunciation of Quranic Alphabets Between Kids and Adults

Syarifah Nuratikah Syd Badaruddin, Salmiah Ahmad, Tareq M.K. Altalmas, Momoh Jimoh E. Salami, Nik Nur Wahidah Nik Hashim, Abdul Halim Embong and Safiah Khairuddin

Department of Mechatronics Engineering,
International Islamic University Malaysia, Kuala Lumpur, Malaysia

Abstract: It is an obligation for a Muslim to become skilled and proficient in reciting Al-Quran considering that Al-Quran is the fundamental source of revelation from Allah SWT. In Al-Quran, there are 28 alphabets where each of them has their own unique sound. The Quranic alphabets produce sound that are characterized from their point of articulation (Makhraj) and their characteristics (Sifaat). Knowing the correct way of pronunciation through engineering perspective may help Muslim in learning Al-Quran, in the sense that the signal of the experts can be used in Quranic teaching and learning as a reference model. Since both adults and children possess different vocal tract, therefore there will be different outcomes of the pronunciation between both experts. The features identification of the pronunciation of both experts is needed to represent the actual and correct pronunciation that will be used as a reference for Quranic teaching and learning at later. In this paper, the focus was on the identification and analysis of the correct pronunciation of the Quranic alphabets on the data obtained from adults and children experts. The first and second formant frequencies (***F1*** and ***F2***) were used as the features where they were used to represent the pronunciation of each alphabet for both adults and children category. The speech analysis software PRAAT was used to accomplish the pre-processing of the data using Spectral Subtraction technique and also used to measure the ***F1*** and ***F2*** values. Linear Discriminant Analysis (LDA) was used for classification of the signals and results shows that some of the alphabets can be identified uniquely using ***F1*** and ***F2*** features of the two categories.

Key words: Formant · Spectrogram · PRAAT · Linear Discriminant Analysis (LDA)

INTRODUCTION

Over the past decades, the implementation of speech recognition systems has been significantly utilized in various applications such as criminal investigations, wireless communication and speech therapy as mentioned by Arshad *et al*. [2]. Speech is one of the most crucial platforms used by human to communicate with each other. Arabic language is one of the main languages to all Muslims around the globe as it is the language of the Holy *Quran*. Most of the Muslims know how to recite *Al-Quran,* but not all Muslims can recite the *Qura*n with the correct *Tajweed* though the learning has started from young age. *Tajweed* is the correctness of diction or proper pronunciation and technique in *Quran* recitation as explained by Arshad *et al*. [3]. In Arabic, the meaning of the word can easily change even with slight differences in pronunciation. Thus, it is compulsory for all Muslims to learn the proper way of reciting *Quran*. Usually, a Muslim starts to learn *Quran* from his early age, where he/she needs to go to the experts (*Ustaz*), or inviting the *Ustaz* to come at home and this process requires years before he/she can read the Quran appropriately. It is *face-to-face* method by which the student is taught directly by the *Ustaz* where by the *Ustaz* will verify the recitation of *Quran* and will correct any mistake done by the student. The correction is made by listening and perhaps by looking at the face of the *Ustaz* so that the student will

Corresponding Author: Salmiah Ahmad, Department of Mechatronics Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia.

repeat accordingly. Commonly in Malaysia, a formal learning of *Quran* starts from as early as 4 years old in kindergarten, where different kindergarten adopt different technique for teaching. In fact, in Malaysia there are many existing methods or modules for *Quranic* education including *Iqra'*, *Al Furqan*, *Al Baghdadi*, *Al Barqi*, *Qiraati*, *Insani*, *Tartila* and many others, which differ in their teaching methodology and module. Among all techniques, there is none that stresses on the signals produced during reciting or pronouncing the *Quran*. Often, if a student can pronounce the alphabet correctly, most of the time he/she can read the Quran well when the times come. But the difficulty is that there is no proof that the student pronounces the alphabet correctly except by listening or looking at the way he reads or pronounces the alphabet as what is expected by the *Ustaz*. In fact it is so difficult to see what is happening inside the speech organs of the student to confirm his pronunciation.

In this project, the identification of the *Quranic* alphabets from adult and children experts was made to analyse to suitable features that can represent the correct pronunciation of each *Quranic* alphabet. A *Quranic* alphabet is attributed by *Makhraj* (articulation) and *Sifaat* (characteristics). *Makhraj* is the correct position of the speech organs to produce a particular sound that represents a unique alphabet so that it can be differentiated from others, [2]. *Sifaat* refers to the characteristic(s) of the *Quranic* alphabets in which can be technically defined as the way each alphabet is articulated that differentiates it from others, [4]. The characteristic is important in distinguishing the alphabets that share the same articulation points (*Makhraj*). The audio signals (sound) were recorded from both the adult and children expert to analyse if there is any similarity or difference among the two categories in producing the correct pronunciation of *Quranic* alphabets, since it is known that they have different vocal tract size [4,5]. The speech assessment on *Makhraj* recognition and Arabic phonemes can be found in [3] and [5]. At present, there are no researches that analyse the correct pronunciation produced by experts with different ages, with the aim to produce models that represent the correct pronunciation of the *Quranic* alphabets of an adult and children experts. Therefore, this paper presents the outcomes of the identification of features that represent most of the correct pronunciation of the *Quranic* alphabets. Later, classification was made where Linear Discriminant technique was used to classify a new signal using the selected features.

MATERIALS AND METHODS

Data Collection from Experts: Series of recording for audio signals were obtained from 10 male experts, where among them, 5 adults and 5 children who were selected based on their proficiency in reading *Al-Quran* according to the established rules and principles. Some of the requirements/standards that have been highlighted in choosing so call the experts are; they need to have a certificate of recognition in the field of study of *Al-Quran* and it is more preferable if they have experience in formally teaching *Al-Quran* to others. The audio signals were recorded using a portable high-quality field recorder, TASCAM DR-05, with frequency response of 40 Hz to 20 kHz. The expert needs to recite the 28 combinations of Quranic alphabets at their best level. The audio data were digitized with sampling rate of 99.6 kHz and processed using open source software known as *PRAAT*. Normal class room environment has been used during the recording, at the Kulliyyah of Engineering, International Islamic University Malaysia. Data taken were using the recitation that is based on *Rasm 'Uthmani* narrated by Hafs bin Sulaiman for *Qira'at Asim,* [3]. Table 1 shows the *sukun*(ـْ) combination of *Quranic* alphabets and its pronunciation that we used for the experts to recite. This *sukun* (ـْ) combination dataset is best to describe the *Makhraj* (point of articulations) and *Sifaat* (characteristics) of each alphabet.

Data Pre-Processing: Pre-processing stage has been conducted for noise reduction and signal normalization between -1 and 1. A speech analysis software known as *PRAAT* was used to perform noise removal filter operation for all recorded audio data as discussed by Almisreb *et al.* [6]. Noise removal based on Spectral Subtraction method was used and the details of the default parameter settings are shown in Table 2. To simplify, all stereo data have been converted to mono channel prior to the noise reduction step. Fig. 1 illustrates the waveform for 'أَبْ' pronunciation in time domain prior to the pre-processing and Fig. 2 shows the same waveform after the pre-processing step.

Features Extraction: There are several types of features extractions that are commonly used previously in many speech analyses such as formant frequency, [6, 7, 8], Mel Frequency Cepstral Coefficient (MFCC), [9] and Power Spectral Density (PSD), [10]. However, in this paper, we focused and tested on the formant frequencies (***F1*** and ***F2***) as the features of the *Quranic* alphabets.

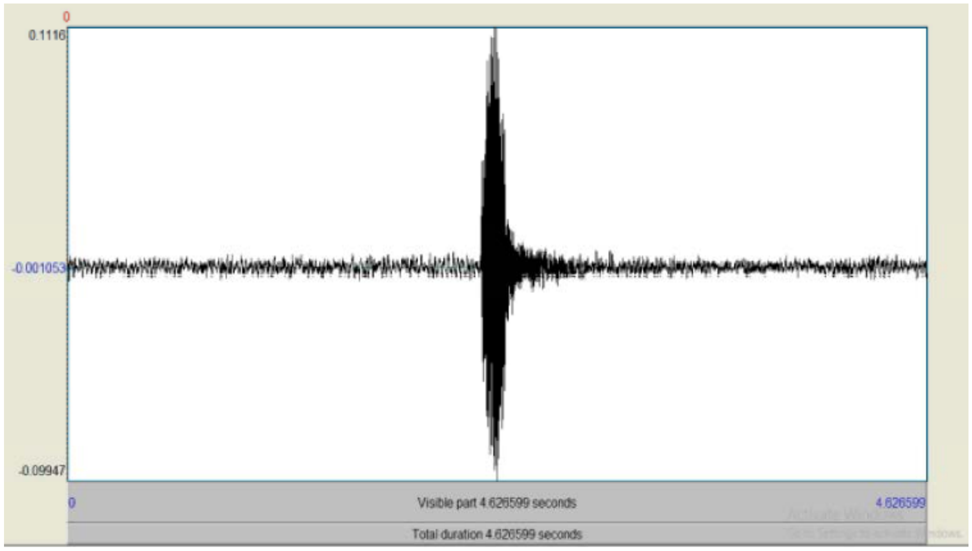

Fig. 1: The time domain waveform for 'أَبْ' pronunciation before pre-processing procedure.

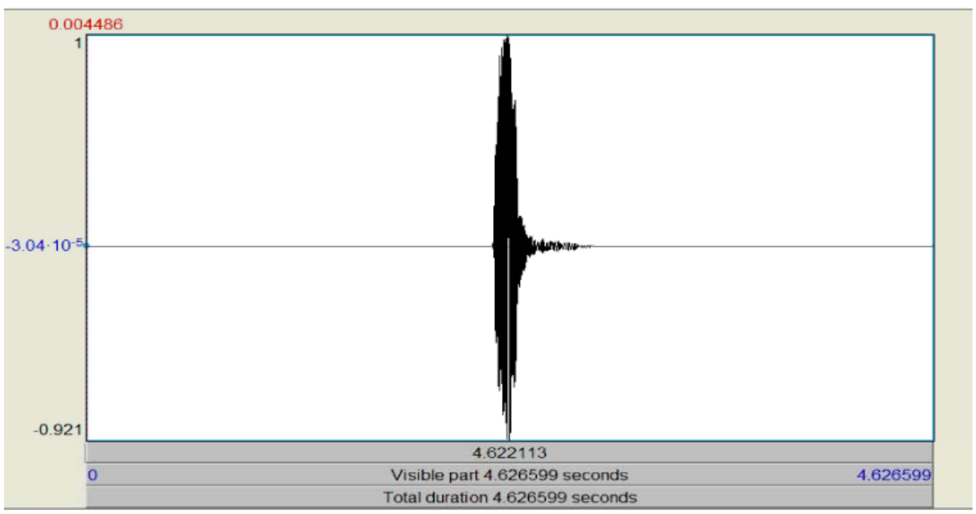

Fig. 2: The time domain waveform for 'أَبْ' pronunciation after noise was reduced and signal was normalized to +/- 1

Table 1: The sukun (ْ) combination of Quranic alphabets and its pronunciation

28 Combination of Quranic letters and its pronunciation			
أءْ	a	أضْ	adh
أبْ	ab	أطْ	athd
أتْ	at	أظْ	azd
أثْ	ath	أعْ	a'
أجْ	aj	أغْ	agh
أحْ	ah	أفْ	af
أخْ	akh	أقْ	ak
أدْ	ad	أكْ	akk
أذْ	az	ألْ	al
أرْ	ar	أمْ	am
أزْ	azz	أنْ	an
أسْ	as	أهْ	ahh
أشْ	ash	أوْ	aww
أصْ	asd	أيْ	aii

Table 2: Noise reduction settings.

Method	Spectral subtraction
Noise time range (s)	0.00 – 0.00
Window length (s)	0.025
Filter Frequency Range (Hz)	80 – 10000
Smoothing (Hz)	40

Table 3: Spectrogram settings.

Window length (s)	0.005
Maximum frequency (Hz)	5000
Time step (s)	0.002
Frequency step (Hz)	20
Window shape	Gaussian
Method	Fourier
Number of time step	1000
Number of frequency step	250

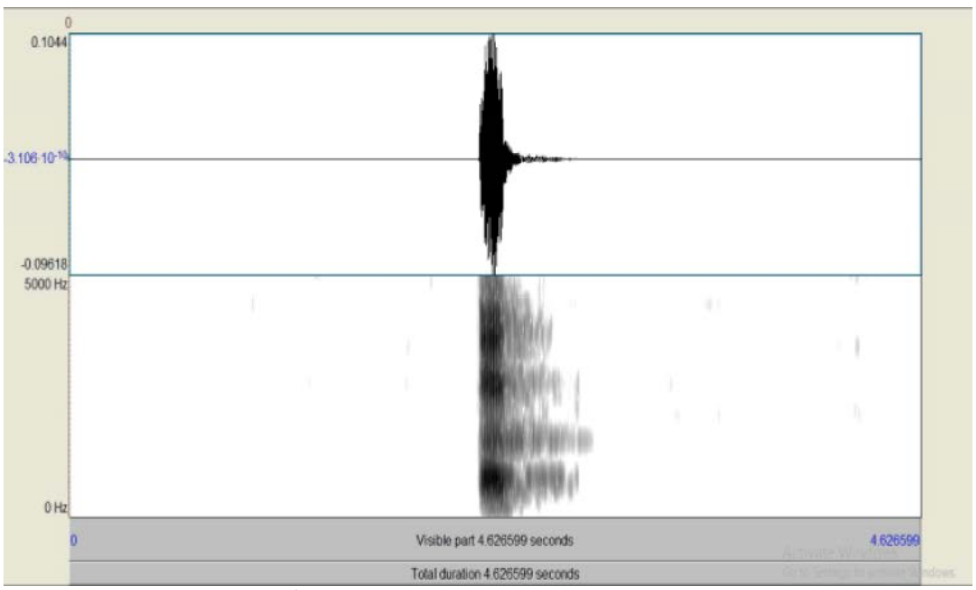

Fig. 3: The enhanced waveforms for 'اًبْ' pronunciation after pre-processing with its corresponding spectrogram graph.

Formant Frequency: A formant frequency represents the concentration of acoustic energy around a particular frequency in the speech signal. At each different frequency, there will be several formants in which the formants can occurred at roughly 1000 Hz interval and each corresponds to a resonance in the vocal tract, [10] & [11]. It is more convenient to extract the formant frequency of any sound signal from its spectrogram graph as stated by Kadir and Sudirman, [8]. In this paper, the extraction of formants frequencies of the audio data was obtained using *PRAAT* software. By using this software, we can get at most 5 different formants but in this paper the focus was only on the first formant frequency (***F1***) and the second formant frequency (***F2***). The important of ***F1*** and ***F2*** comes from its role in examining vowel categories and mapping acoustic realisation into articulatory expectation. They also stated that the lower the values of ***F1*** resembles that the tongue is closer to the roof of the mouth. Meanwhile, the value of ***F2*** is proportional to the front or the back of the highest part of the tongue during producing the vowel. Table 3 shows the spectrogram settings in PRAAT used to get the spectrogram graph as well as to obtain the respective formants frequencies.

Fig. 3 shows the enhanced waveform for (اًبْ) pronunciation with its corresponding spectrogram after undergoes pre-processing steps and Fig. 4 illustrates the waveforms after manually removed the unwanted segment speech with the indication of the location of the formants on the spectrogram. The location of formant is indicated with red points on the graph and the obtained values of ***F1*** and ***F2*** were summarized in result part.

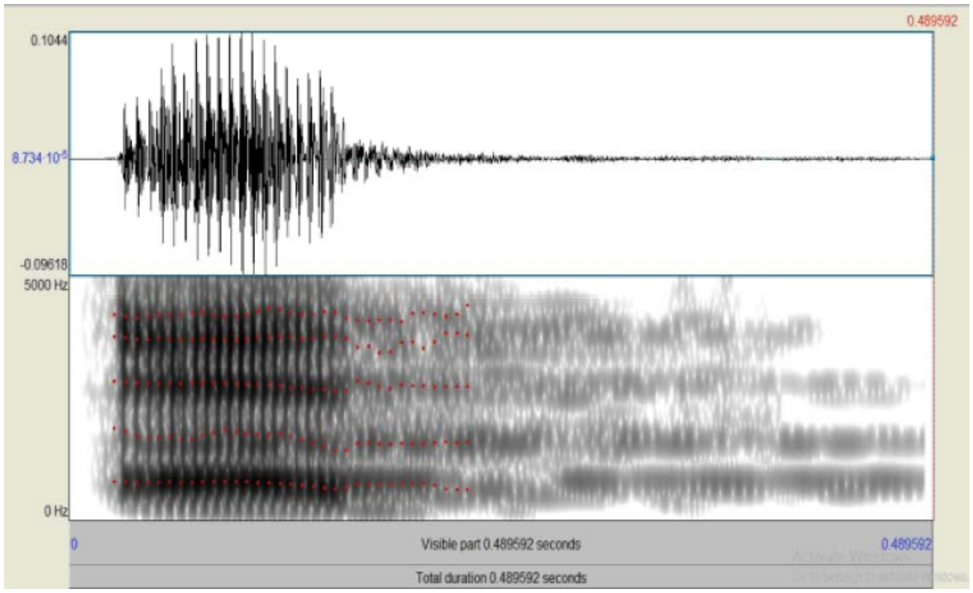

Fig. 4: Waveforms for 'أَبْ' pronunciation after manually remove the unwanted segment with the indication of formants location on the spectrogram graph.

Data Classification and Analysis: The data classification procedure that has been implemented in this study is known as Linear Discriminant Analysis (LDA) or linear classifier. The discriminant analysis performed on the obtained features was done based on pair wise analysis classification consisting of adult/ child. Next, a re-sampling method is used while performing linear classifications since the number of data sets obtained are small. The re-sampling method that was adopted in this project is Jackknife (Leave-One-Out) since it can give an optimum result as mentioned by Hashim, [10].

The "classify" command in MATLAB that was used in this research requires three input parameters which are test, train and class label. The class labels are made of a bunch of zeros and ones that were stacked together in a column vector. The size of the column vector depends on the number of training samples. The training data consists of *n* row by *c* column matrix where the size of *n* varies depending on the training size sample. The two features extracted from formant frequencies are represented as *c.* Each training sample is associated with a class label, where in this case, the class label would be either zero or one. The test data were assumed to have an "unknown" class label. The classify command will output an estimated class label for the test data according to the training data distribution and classification.

The Jackknife method re-samples data without replacement and so the training sample will not be duplicated when performing classification. For this research, the overall sets consist of 10 data for each 28 pronunciation of *Quranic* alphabets from two different classes which are adult experts and children experts. Then, these data sets were classified in according to the respective classes and each pronunciation of *Quranic* alphabet has different number of row vectors.

The implementation of Jackknife in this research is on the basis of leave-one-out pronunciation of *Quranic* alphabet instead of per vectors. Therefore, the tested data set will come from the one-leave pronunciation and the training data set will be the remaining 9 data for each pronunciation. For the purpose of this study, the class label for testing data set is assumed to be unknown. The classifier output will be a vector of all zeros, all ones or mixture of both depending on how well the classifier can classify the data accordingly. The process was repeated by excluding the next pronunciation from the overall sets of data until all pronunciations have been chosen as testing data. This method can give optimum outcome for re-sampling since it uses most amounts of data as training when doing classification.

RESULTS

Data obtained from Features Extraction Stage: All recorded data have undergone a data pre-processing and features extraction stages. This stage was done using *PRAAT* software. Table 3 and Table 4 summarizes the values of ***F1***, ***F2*** and the average of both formants

Table 3: Value of F1, F2 and the average formants from adult experts.

		ADULT					
Letters	Formant	1	2	3	4	5	Avg.
اءْ	F1	814.006	705.162	658.616	652.411	626.027	691.244
	F2	1710.578	1574.647	1350.131	1817.796	1521.547	1594.940
أبْ	F1	736.642	558.829	490.870	608.361	599.347	598.809
	F2	1639.862	1557.231	1289.306	1713.681	1484.476	1536.911
أتْ	F1	759.374	596.076	496.502	571.163	618.807	608.384
	F2	1731.275	1742.506	1584.363	1727.613	1569.121	1670.976
أثْ	F1	740.240	595.935	550.959	563.115	603.365	610.723
	F2	1672.497	1762.684	1506.255	1494.726	1613.933	1610.019
أجْ	F1	716.225	410.980	444.939	530.548	495.553	519.649
	F2	1895.701	1829.884	1551.531	1689.793	1790.580	1751.498
أحْ	F1	898.890	854.112	832.830	942.938	810.374	867.829
	F2	1860.777	1925.647	1638.535	1815.357	1587.258	1765.515
أخْ	F1	811.159	888.665	738.731	648.838	699.881	757.455
	F2	1394.916	2197.480	1626.574	1350.325	1563.359	1626.531
أدْ	F1	689.811	496.915	423.8255	500.226	524.138	526.983
	F2	1781.905	1746.805	1480.368	1697.041	1668.793	1674.982
أذْ	F1	642.980	327.539	438.752	440.616	480.410	466.059
	F2	1779.211	1666.433	1456.702	1668.228	1588.927	1631.900
أرْ	F1	726.148	551.131	654.967	517.677	575.989	605.182
	F2	1372.546	1044.402	1373.807	1089.629	1260.795	1228.236
أزْ	F1	631.892	433.676	438.433	497.001	495.824	499.365
	F2	1798.770	1858.061	1362.011	1642.465	1641.310	1660.523
اسْ	F1	770.282	603.685	520.677	621.425	695.568	642.3274
	F2	1997.148	2116.893	1602.939	2038.547	2209.814	1993.068
اشْ	F1	1149.257	1631.812	1375.727	1918.769	1928.992	1600.911
	F2	1752.955	2473.557	2562.402	2525.033	2525.547	2367.899
اصْ	F1	824.378	1192.466	700.072	808.061	816.592	868.3138
	F2	2035.439	2046.711	1853.863	2468.494	2470.510	2175.003
اضْ	F1	714.194	367.523	528.680	407.467	404.713	484.515
	F2	1482.247	1060.017	1211.592	940.703	938.640	1126.640
اطْ	F1	742.329	653.052	519.906	666.440	669.540	650.253
	F2	1305.086	951.934	1174.669	1051.573	1052.979	1107.248
أظْ	F1	710.321	564.286	561.945	545.645	539.663	584.372
	F2	1173.675	1660.621	1295.237	1081.253	1029.530	1248.063
أعْ	F1	878.413	953.509	895.295	841.727	841.671	882.123
	F2	1392.168	1632.529	1380.494	1288.446	1290.372	1396.802
أغْ	F1	685.845	565.883	539.298	568.204	566.976	585.2412
	F2	1308.502	1155.951	1505.484	1151.965	1149.432	1254.267
أفْ	F1	748.455	930.441	732.762	725.929	753.434	778.204
	F2	1602.485	1821.465	1864.489	2059.041	2071.800	1883.856
أقْ	F1	750.540	715.470	625.836	592.286	593.447	655.516
	F2	1327.537	1046.744	1282.933	1039.236	1012.706	1141.831
أكْ	F1	814.213	788.929	794.900	910.837	889.790	839.734
	F2	1915.783	1898.257	1895.987	1866.142	1837.208	1882.675
ألْ	F1	627.307	341.854	333.391	400.645	408.181	422.276
	F2	1673.184	810.783	1745.690	2066.801	2063.284	1671.948
أمْ	F1	742.947	322.106	401.444	422.570	419.073	461.628
	F2	1663.127	1835.931	1687.277	1692.866	1705.242	1716.889
أنْ	F1	729.445	331.741	422.040	415.619	417.326	463.234
	F2	1715.232	1902.467	1966.892	1770.339	1759.984	1822.983
أهْ	F1	748.916	871.383	722.480	725.467	724.695	758.588
	F2	1727.190	1530.834	1651.726	1750.728	1748.714	1681.838
أوْ	F1	551.535	497.596	531.053	482.302	481.421	508.781
	F2	1032.660	1248.081	1149.295	1094.950	1069.676	1118.932
أيْ	F1	523.439	390.345	427.365	406.777	408.473	431.279
	F2	2158.175	2179.647	1920.896	2187.559	2186.988	2126.653

Table 4: Value of F1, F2 and the average formants from children experts.

		CHILDREN					
PRO	FORMANT	1	2	3	4	5	AVG
أءْ	F1	615.212	865.738	730.079	646.007	734.783	718.364
	F2	1399.890	1851.400	1407.390	1666.797	1480.296	1561.155
أبْ	F1	589.279	652.308	547.215	775.823	595.081	631.941
	F2	1494.721	1795.623	1533.465	1755.558	1456.236	1607.121
أتْ	F1	694.345	775.159	659.585	555.43	685.361	673.976
	F2	1807.460	1978.779	1637.034	1361.636	1785.503	1714.082
أثْ	F1	653.610	803.133	801.695	984.031	561.535	760.801
	F2	1863.356	2032.689	2031.043	1935.077	1978.594	1968.152
أجْ	F1	628.311	672.006	511.473	790.269	572.062	634.824
	F2	1881.868	2140.073	1638.592	1968.528	1775.001	1880.812
أحْ	F1	819.982	962.216	991.220	1111.602	815.080	940.020
	F2	1603.599	2140.840	1996.509	2207.974	1886.345	1967.053
أخْ	F1	639.776	842.742	912.694	1144.024	677.341	843.3154
	F2	1038.889	1401.054	1454.046	1550.153	1279.847	1344.798
أدْ	F1	605.0170	677.706	532.580	553.774	620.067	597.829
	F2	1871.714	2014.758	1782.500	1672.307	1919.982	1852.252
أذْ	F1	562.176	669.562	565.211	527.810	681.058	601.163
	F2	1974.972	1959.204	1862.189	1634.358	2067.533	1899.651
أرْ	F1	622.623	741.135	619.257	671.758	751.153	681.185
	F2	1326.842	1258.597	998.788	1253.035	1398.169	1247.086
أزْ	F1	796.757	660.424	557.329	517.244	627.461	631.843
	F2	2217.005	2019.814	2070.286	2093.635	2033.438	2086.836
أسْ	F1	888.583	758.590	688.684	1026.325	805.120	833.460
	F2	2260.914	2168.913	2360.147	2407.258	2043.392	2248.125
أشْ	F1	1669.899	863.152	1279.295	1519.548	762.756	1218.930
	F2	2743.962	2406.837	2614.218	2586.279	2514.843	2573.228
أصْ	F1	837.739	892.435	1365.139	1142.702	778.736	1003.350
	F2	2332.541	2023.779	2586.629	2469.553	1917.816	2266.064
أضْ	F1	687.468	704.213	466.596	457.444	498.316	562.807
	F2	1559.828	1223.557	966.444	1089.319	1280.254	1223.880
أطْ	F1	558.859	753.177	610.613	919.823	697.471	707.989
	F2	1084.005	1441.361	1189.663	1404.554	1293.640	1282.645
أظْ	F1	524.912	775.392	587.646	606.415	613.322	621.537
	F2	1134.814	1302.855	1349.674	1329.579	1738.230	1371.03
أعْ	F1	840.783	1053.187	791.624	1007.053	726.399	883.809
	F2	1518.029	1824.696	1427.026	1413.947	1513.194	1539.378
أغْ	F1	649.775	824.848	719.573	647.763	663.563	701.104
	F2	1153.163	1343.355	1181.607	1358.034	1293.967	1266.025
أفْ	F1	1206.616	817.876	1002.001	994.286	746.780	953.512
	F2	2357.001	2164.608	2147.990	1949.756	1816.204	2087.112
أقْ	F1	650.612	805.427	637.636	773.715	741.100	721.698
	F2	1227.396	1467.019	1090.692	1297.182	1128.304	1242.119
أكْ	F1	1108.834	834.442	698.357	634.457	829.169	821.052
	F2	1926.106	1733.428	1812.647	1512.664	1533.244	1703.618
ألْ	F1	456.607	558.659	521.300	486.871	551.317	514.951
	F2	1639.984	1762.544	1790.974	2090.808	2044.769	1865.816
أمْ	F1	431.997	565.751	422.908	452.561	564.035	487.450
	F2	1671.959	1684.118	1894.572	1530.439	1408.921	1638.002
أنْ	F1	423.053	511.222	432.051	491.153	548.479	481.192
	F2	1924.498	1698.683	2045.638	1879.449	1524.855	1814.625
أهْ	F1	1011.866	879.364	732.821	805.378	880.271	861.940
	F2	1765.115	1839.083	1676.836	1924.192	1687.034	1778.452
أوْ	F1	521.458	604.030	478.212	562.847	628.273	558.964
	F2	1104.107	1358.710	878.338	1371.460	1408.849	1224.293
أيْ	F1	409.541	567.214	407.948	517.519	581.602	496.765
	F2	2396.663	2090.388	2319.759	2720.576	2165.316	2338.540

Table 5: Percentage of Jackknife classification for each pronunciation from both adult experts group and children experts group as well as the overall percentage of both groups.

JACKKNIFE	LINEAR (%)		
PRONUNCIATION	OVERALL	ADULT	CHILD
أَءْ	40	40	40
أَبْ	20	40	0
أَتْ	70	60	80
أَثْ	90	80	100
أَجْ	70	80	60
أَحْ	50	40	60
أَخْ	80	80	80
أَدْ	70	60	80
أَذْ	80	80	80
أَرْ	80	80	80
أَزْ	80	60	100
أَسْ	60	60	60
أَشْ	80	80	80
أَصْ	40	40	40
أَضْ	30	60	0
أَطْ	50	60	40
أَظْ	50	60	40
أَعْ	70	80	60
أَغْ	70	80	60
أَفْ	70	80	60
أَقْ	40	40	40
أَكْ	80	100	60
أَلْ	70	60	80
أَمْ	40	40	40
أَنْ	0	0	0
أَهْ	80	80	80
أَوْ	50	60	40
أَيْ	60	80	40

Pronunciation that has equal to or more than 80% of overall percentage of classification.

Table 6: Average value of F1 and F2 for the pronunciations that has equal to or more than 80% of overall percentage of correct classification.

	ADULT		CHILDREN	
Letter	*F1*	*F2*	*F1*	*F2*
أَثْ	610.723	1610.019	760.801	1968.152
أَخْ	757.455	1626.531	843.315	1344.798
أَذْ	466.059	1631.900	601.163	1899.651
أَرْ	605.182	1228.236	681.185	1247.086
أَزْ	499.365	1660.523	631.843	2086.836
أَشْ	1600.911	2367.899	1218.930	2573.228
أَكْ	839.734	1882.675	821.052	1703.618
أَهْ	758.588	1681.838	861.940	1778.452

from both adult and children experts respectively. The results shown only for the first 4 pronunciations and the number of (1, 2, 3, 4 and 5) in both tables represent the numbers of expert reciters involved with the recording sessions.

By summarizing the average value of ***F1*** and ***F2*** into the table form, it is easier to analyse the difference between the adult and child pronunciations in terms of formant frequencies that have been extracted from the filtered data. However, a classification process needs to be performed for all of the extracted features in order to know whether the obtained features can be used to represent each *sukun* combination alphabets accordingly for both adults and children samples.

Result from Data Classification: The first two formant frequencies ***F1*** and ***F2*** were calculated from all expert reciters and classified using linear classifier. The re-sampling method that was implemented in this research was Jackknife method and the percentage of correctness for each pronunciation was recorded in Table 5. It indicates the classification result for adults and children group respectively as well as the overall percentage for both categories that pronounce each *sukun* combination alphabets. Table 6 summarized the average value of ***F1*** and ***F2*** for all pronunciation that has the overall percentage of classification which is equal to or greater than 80%. In this research, 80% was assumed to be the threshold value in which indicates that the selected

Table 7: Adults and Children features classification using all reciter’s data.

Percentage (%)	All data	Adults	Children	Feature combination	Classifier
Jackknife	57.14	62.14	52.14	*F1, F2*	LDA

features combination will be used to create a reference model if the classifier managed to correctly classified 4 out of 5 expert’s pronunciations from both adult and child categories.

Next, the percentage of correctness for all *sukun* combination of Quranic alphabets or for all pronunciations was recorded in Table 7. The percentage indicates the features vectors that are correctly classified over all classified group. The adults and children percentages represent the percentage of feature vectors that are correctly classified within the adult expert group and children experts group respectively. These optimum results obtained through the combination of formant frequencies (***F1*** and ***F2***) and linear classifier.

From the result obtained in Table 5, only 8 out of 28 pronunciations can be represented using the combination of ***F1*** and ***F2*** values. Those alphabets are أَهْ, أَكْ, أَشْ, أَزْ, أَرْ, أَدْ, أَخْ, أَثْ and the point of articulation (*Makhraj*) for these pronunciation are tongue and throat as discussed by Afree, [12]. This indicates that the combination of ***F1*** and ***F2*** values alone are not sufficient in representing pronunciation generated through mouth (أَوْ, أَمْ, أَفْ, أَبْ) and chest area (أَيْ, أَوْ, اءْ).

Using Jackknife classifier, 62.14% of the adult’s expert pronunciations and 52.14% of the children’s expert pronunciation data can be correctly classified with a linear discriminator classifier using ***F1*** and ***F2***.

CONCLUSIONS

In this paper the correct pronunciation of the *Quranic* alphabets among adults and children who are expert in reciting *Quran* has been identified and analysed. Adults and children have different vocal tract size which affects the formant frequencies of the pronounced sounds. First two formants have been used as the extracted feature. The results indicate that the extracted formant ***F1*** and ***F2*** features are able to represent 8 alphabets, which are أَهْ, أَكْ, أَشْ, أَزْ, أَرْ, أَدْ, أَخْ, أَثْ uniquely for both adults and children experts’ pronunciation. It is concluded that the extracted formant frequencies ***F1*** and ***F2*** can only represent some of the Quranic letters uniquely and not for the rest of the letters, where the rest need other forms of features and will be addressed at later.

ACKNOWLEDGEMENTS

The author would like to acknowledge to Ministry of Higher Education Malaysia for supporting the research under the project grant RIGS15-146-0146 and also would like to express sincere appreciation to Ustaz Dr. Surul Shahbudin Hassan from the Centre For Languages and Pre-University Academic Development, International Islamic University Malaysia, Kuala Lumpur, Malaysia for his effort in making the experiment successfully conducted by providing the experts. We also thank to all Ustaz and Ustazah who have kindly agreed and spent time in recording the pronunciation and recitation.

REFERENCES

1. Serizel, R. and D. Giuliani, 2014. Vocal tract length normalization approches to DNN-based children's and adult's speech recognition. pp: 135-140.
2. Arshad, N., S.A. Aziz, R. Hamid, R.A. Karim, F. Naim and N. Zakaria, 2011. Speech Processing for Makhraj Recognition - The Design of Adaptive Filter for Noise Removal. International Conferene on Electrical, Control and Computer Engineering, pp: 323-327.
3. Arshad, N.W., S.M. Sukri, L.N. Muhammad, H. Ahmad, R. Hamid, R. Naim and N.Z. Naharuddin, 2013. Makhraj Recognition for Al-Quran Recitation using MFCC. International Journal of Intelligent Information Processing (IJIIP), pp: 45-53.
4. Read With Tajweed Rules. Retrieved from http://www.readwithtajweed.com/tajweed_Sifaat.htm
5. Almisreb, A.A., A.F. Abidin and N.M. Tahir, 2016. Acoustic Analysis of Arabic Phonemes Based-/a:/ Vowel. International Colloquium on Signal Processing & its Application (CSPA2016), pp: 280-284.
6. Zahra N.S. and Z. Zainon, 2012. Speech Quality based on Arabic Pronounciation using MFCC and LDA. International Conference on Control System, Computing and Engineering, pp; 398-403.
7. Altamas, T., S. Ahmad, W. Sediono and S.S. Hassan, 2015. Quranic Letter Pronunciation Analysis based on Spectogram technique: Case Study on Qalqalah Letters, pp: 14-22.

8. Kadir, N.A. and R. Sudirman, 2011. Vowel Effects towards Dental Arabic Consonants based on Spectogram. Second International Conference on Intelligent Systems, Modelling and Simulation, pp: 183-188.
9. Ibrahim, N.J., Z. Razak, E.M. Tamil, M.Y. Idris and Z.M. Yusoff, 2008. Quranic verse recitation feature extraction using Mel-Frequency Cepstral Coefficient (MFCC). 4th International Colloquium on Signal Processing and its Applications, pp: 13-17.
10. Hashim, N.N., 2011. Analysis of Power Spectral Density of male speech as indicators for high risk and depressed decision.
11. Wood, S., 2005, January 15. Praat for beginners. Retrieved from What are formant: http://person2.sol.lu.se/SidneyWood/praate/whatform.html
12. HK, H., W.S, C. MK, S. EM, D. RB, K. RD, . . .G. LR, 2009. March. Retrieved from Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study1: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2669667/.