

Using Formants to Compare Short and Long Vowels in Modern Standard Arabic

Veton Kepuska, Mohamed Alshaari

Computer Engineering and Sciences, Florida Tech, Melbourne, USA

Email: vkepuska@fit.edu, malshaari2016@my.fit.edu

How to cite this paper: Kepuska, V. and Alshaari, M. (2020) Using Formants to Compare Short and Long Vowels in Modern Standard Arabic. *Journal of Computer and Communications*, 8, 96-106.

<https://doi.org/10.4236/jcc.2020.85006>

Received: April 17, 2020

Accepted: May 25, 2020

Published: May 28, 2020

Copyright © 2020 by author(s) and

Scientific Research Publishing Inc.

This work is licensed under the Creative

Commons Attribution International

License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This study was concerned with the short vowels in modern standard Arabic words with Consonant Vowel-Consonant Vowel-Consonant Vowel (CVCVCV) structure, and the long vowels in words with Consonant Vowel Vowel-Consonant (CVVC). Even though there has been a dispute on the precise number of Arabic vowels that exist between language studies, this study used the opinion that the Arabic language has three vowels; the elongation of each vowel gave the other three because this is the opinion of classical Arabic linguists which is the source of the Modern Standard Arabic (MSA). Studies said that the first and second formant values (F1, F2) can represent the vowels. In this study, the formants were measured using LPC (Linear Predictive Coding), verifying the measurement to see if the measured follows the pattern of formants measurements of the other studies, and the formants were used to investigate the relationship between short and long vowels. Furthermore, the study figured out if the dialect of speakers can affect the values of formants, even if the spoken language is MSA, some statistical measurements were calculated to evaluate the relationship.

Keywords

Arabic Vowels, Formant, MSA, CVCVCV Words, Corpus, Vowel Triangle, Dialect

1. Introduction

The sixth most broadly spoken language in the world is Arabic. Nowadays, there are three different kinds of Arabic: Classical, MSA, and many Arabic Dialects. Classical Arabic is old Arabic, which is used in holy texts and for linguistic studies. Standard Arabic is the formal language for all Arabic countries. It is used for official communications and writing in schools. It is also the language used

in the media. There are many Arabic Dialects. This study only uses MSA.

Arabic has only six vowels: three short vowels and three long vowels. If the short vowels extend for a certain period of time, then the longer vowels are produced. If we look for short vowels in the cardinal vowel chart, we find the nearest English or The International Phonetic Alphabet (IPA) vowels are located at the edges of the chart, which are /a/, /i/, and /u/ and they are shown as follows in **Table 1**.

Some of the characteristics of vowels in terms of articulators of the vocal tract can be measured as acoustic resonance which arises from the voice when it passes through the vocal tract. The range of frequencies of augmented resonance is called Formants [1]. Formants are used to distinguish vowels since every vowel has a distinct value of F1 and F2. Formant is a term coined by Ludimar Hermann for the frequencies. He “observed that the spectrum of the decaying elementary wave of a vowel is peaked at a number of frequencies, characteristic of the vowel.” [2] In other words, a formant is a concentration of acoustic energy around a particular frequency in the speech wave. There are several formants: a formant with lowest frequency is called F1, the second formant is called F2, and the third formant is called F3. “Most often the two first formants, F1 and F2, are enough to disambiguate the vowel.” [3] So, every vowel has a distinct value of F1 and F2.

This study measures the F1 and F2 for Arabic short and long vowels from MSA words with a CVCVCV pattern, and then compares the results with some old studies and investigates if the nationality/dialect can affect the formant results.

2. IPA and Formants of Arabic Vowels

IPA [4] puts a notation for speech sounds used by humans to speak any language. The notation expresses the vowels depending on the position of the tongue, either bottom or top, the figure of lips, and the opening of the mouth. The phoneticians such as Daniel Jones tried to express all vowels using the triangle chart (**Figure 1**).

As an example, the vowels of the English language contain more than 24 sounds which are expressed in the cardinal vowels chart. /a/ vowel is expressed at the bottom position of the tongue which is very low. When the tongue is high at the top of the mouth the vowel is /i/. However, when the tongue is far back, very high, and the lips are rounded the vowel is /u/.

Table 1. Arabic vowels and english approximation.

| Symbol | English approximation | Arabic letter/diacritic | Notes |
|--------|-----------------------|-------------------------|-------|
| a | father, but shorter | اَ | Short |
| a : | father | آ | Long |
| i | milk | اِ | Short |
| i : | machine | يِ | Long |
| u | put | اُ | Short |
| u : | rule | وِ | Long |

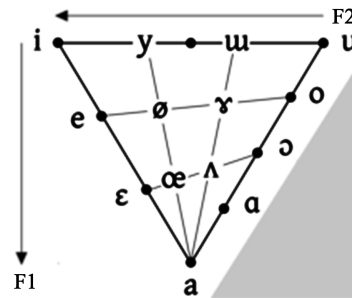


Figure 1. Daniel Jones triangle chart [5].

Most of the time, there are some differences between the two sets of formant measurements, this may be due to dialect and/or methodological differences. Some studies use LPC spectral peaks and others use narrow band spectra produced on an analog spectrograph [6]. “LPC is a technique used to model the vocal tract. The resulting curve is presented in two dimensions like the FFT spectrum. Generally, the LPC spectrum follows the envelope of the standard spectrum and is useful to measure vowel formants which correspond to the peaks.” [7]

If formants are well used, the vocal tract may be characterized and represented with fewer parameters e.g. using the LPC method which is one of the most effective and valuable methods for speech analysis. LPC is a method used mostly in the processing of audio signals and speech, and for encoding voice of good quality at a low bit rate which provides highly accurate estimates of speech parameters. “Several authors have therefore investigated formant frequencies as speech recognition features, using various methods for basic analysis, such as linear prediction.” [8]

However, the approximations for F1 and F2 values of short vowels /a/, /i/, and /u/ in IPA, American English, and Arabic studies are different as Table 2 shows.

In addition, Table 3 is a comparative list of F1 and F2 of Arabic vowels according to a number of studies [1].

3. The Arabic Corpus

Since the Arabic language lacks available corpora which contain isolated Arabic sounds and CVCVCV words necessary to fulfill the purpose of this research, which is investigating the formants of Arabic vowel. A corpus of Standard Arabic was built through recording a multitude of standard Arabic speakers pronouncing a set of words [11]. The corpus can be used in linguistics, Arabic speech recognition, and identification. In addition, the data can also be used to extract certain sound features such as formants and MFCC features.

To build the corpus, the chosen utterances are nineteen male adult Arabic native speakers who studied standard Arabic, they recorded a list of 24 Arabic words as displayed in Table 4 to extract the short vowels and 3 Arabic words to extract the long vowels as shown in Table 5, each one of the speakers has recorded each word three times. The words have CVCVCV and CVVC pattern,

Table 2. Approximation of short vowels formants.

| Short Vowels | F1 [hz] | | | F2 [hz] | | |
|--------------|---------|----------------------|-------------|---------|------------------|--------|
| | IPA | American English [9] | Arabic [10] | IPA | American English | Arabic |
| a | 850 | 768 | 616 | 1610 | 1333 | 1460 |
| i | 240 | 342 | 440 | 2400 | 2322 | 1770 |
| u | 250 | 378 | 480 | 595 | 997 | 1170 |

Table 3. A comparative list of F1 and F2 of Arabic vowels.

| Vowels | i | | u | | a | |
|------------|------------|-------------|------------|-------------|------------|-------------|
| | F1 | F2 | F1 | F2 | F1 | F2 |
| Al-Ani | 290 | 2200 | 290 | 800 | 600 | 1500 |
| Belkaid | 355 | 1830 | 340 | 995 | 400 | 1640 |
| Haidar | 485 | 1750 | 500 | 1120 | 675 | 1585 |
| Newman | 440 | 1770 | 480 | 1170 | 616 | 1460 |
| Min | 290 | 1750 | 290 | 800 | 400 | 1460 |
| Max | 485 | 2200 | 500 | 1170 | 675 | 1640 |

Table 4. List of MSA CVCVCV vowel carrier words.

| | | | | | | | |
|---|--------|----|--------|----|---------|----|--------|
| 1 | فَعَلَ | 7 | فَعِلَ | 13 | فَعُلَ | 19 | فَعِلَ |
| 2 | رَفَعَ | 8 | بَحَلَ | 14 | بَلَّغَ | 20 | ذَكَرَ |
| 3 | ذَكَرَ | 9 | عَمِلَ | 15 | صَلَحَ | 21 | جُمِعَ |
| 4 | ذَهَبَ | 10 | حَفِظَ | 16 | سَهَّلَ | 22 | خُلِقَ |
| 5 | شَرَحَ | 11 | سَمِعَ | 17 | كَبَّرَ | 23 | كُتِبَ |
| 6 | كَتَبَ | 12 | فَرِحَ | 18 | كَرَّمَ | 24 | حُشِرَ |

Table 5. List of MSA CVVC vowel carrier words.

| | |
|---|-------|
| 1 | بَابُ |
| 2 | حِينُ |
| 3 | سُورُ |

they involve all the Arabic vowels followed by the different types of consonants [11]. In addition, the file type is wav. The sampling frequency rate used in recording these words is 48 kHz and 32-bit resolution mono. The data files were recorded and saved as wav audio file format to retain and keep the accuracy of the measurements, because wav files are a raw and uncompressed file format.

Since the language used in the study is standard Arabic, the nationality of speakers is supposedly will not have a major effect. However, the study will investigate whether that is true. In addition, the speakers' nationalities are: six Libyans, five Egyptians, two Syrians, two Saudis, two Lebanese, one Moroccan, and one Iraqi.

4. Data Preprocessing

Arabic short, and long vowels have been chosen and extracted from the recorded files, each file represents one word from 1368 Arabic words (*i.e.*, nineteen speakers times 24 words times three trials for each speaker). The total of the extracted audio vowels is 513 (V) for Arabic short (a, i, u) and 171 (VV) for long (a:, i:, u:) vowels. Then, every vowel is recorded in a separate wav file.

5. Formants Detection with LPC Analysis

F1 and F2 are important to determine the vowels, and F3 is essential in determining the quality of a sound, and F4 and F5 are important to figure out the quality of the sound. However, only the vowels are concerned in this study. The LPC method was applied to the vowels' recorded files to measure the formants F1 and F2 of Arabic short and long vowels.

6. The Experiment and Results

The formants for short and long vowels were measured, since the formant values were expected to be close to each other, the mean (sum of the values/number of values) was used to know whether the values of the measurements are close to each other. The standard deviation ($SD = \frac{\sum(\text{each values}-\text{mean})^2}{\text{number of values}}$) was used to know how the formants are scattered out from the mean, if the SD is low then that means the formants' values are close to the mean. The coefficient of variation (CVar) was used to express the proportion of the standard deviation to the mean (SD/Mean), it was used to measure the accuracy of the measurements. Those statistical values were also compared to those of the IPA. In addition, the slight affect in measurements between certain nationalities was also looked at and was verified.

6.1. Short Vowel Formant

The F1 and F2 for short vowels have been calculated. **Table 6** shows the mean of F1 and F2 for short vowels in front of each speakers' number. In addition, the mean (see **Figure 2**), minimum and maximum values of all F1's and F2's are also shown at the bottom.

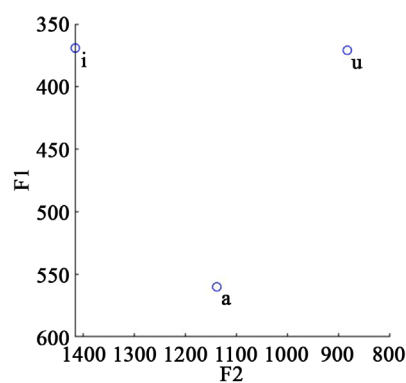


Figure 2. Formants' means of Arabic short vowels.

Table 6. The mean of F1 and F2 for short vowels for all speakers.

| Volunteer No. | a_Vowel | | i_Vowel | | u_Vowel | |
|---------------|------------|-------------|------------|-------------|------------|-------------|
| | F1 | F2 | F1 | F2 | F1 | F2 |
| 1 | 616 | 1195 | 380 | 1449 | 345 | 980 |
| 2 | 652 | 1280 | 400 | 1599 | 429 | 971 |
| 3 | 465 | 1033 | 337 | 1286 | 341 | 810 |
| 4 | 559 | 1025 | 339 | 689 | 390 | 750 |
| 5 | 465 | 1057 | 262 | 1592 | 347 | 1570 |
| 6 | 586 | 1215 | 393 | 1608 | 388 | 1042 |
| 7 | 528 | 1063 | 381 | 1451 | 360 | 732 |
| 8 | 576 | 1027 | 339 | 1240 | 402 | 1005 |
| 9 | 583 | 1176 | 339 | 1382 | 354 | 810 |
| 10 | 626 | 1066 | 404 | 1406 | 392 | 815 |
| 11 | 529 | 1062 | 404 | 1746 | 371 | 852 |
| 12 | 555 | 1272 | 422 | 1805 | 352 | 858 |
| 13 | 591 | 1155 | 394 | 1250 | 363 | 836 |
| 14 | 623 | 1400 | 373 | 1285 | 376 | 815 |
| 15 | 536 | 1113 | 361 | 1562 | 366 | 1018 |
| 16 | 520 | 1035 | 419 | 1448 | 398 | 787 |
| 17 | 502 | 1351 | 394 | 1769 | 375 | 820 |
| 18 | 559 | 1367 | 320 | 1130 | 350 | 613 |
| 19 | 571 | 736 | 354 | 1192 | 349 | 692 |
| Mean | 560 | 1138 | 369 | 1415 | 371 | 883 |
| Min | 465 | 736 | 262 | 689 | 341 | 613 |
| Max | 652 | 1400 | 422 | 1805 | 429 | 1570 |

In order to verify the accuracy of the formants for each vowel and be confident that the extracted formants are accurate and reliable, the SD and CVar were calculated, the results are illustrated in **Table 7**.

All the CVar values are less than 1 which is considered low, which means that the measurements tend to be close to the mean. Therefore, the results are precise and reliable.

6.2. Formants of Long Vowels

From the isolated long vowels, F1 and F2 have been calculated. **Table 8** shows the mean of F1 and F2 for long vowels in front of each speakers' number. In addition, the mean (see **Figure 3**), minimum and maximum of all F1's and F2's are shown at the bottom.

The SD and CVar for the mean values of long vowels' formants were calculated, and the results are illustrated in **Table 9**.

Table 7. The mean, SD and CVar of F1 and F2 for short vowels.

| | all volunteers | | | | | |
|-------------|----------------|----------|---------|----------|---------|----------|
| | a | | I | | u | |
| | F1 | F2 | F1 | F2 | F1 | F2 |
| Mean | 560 | 1138 | 369 | 1415 | 371 | 883 |
| SD | 129.1780 | 352.1000 | 56.3061 | 589.5177 | 42.2882 | 279.0713 |
| CVar | 0.2306 | 0.3093 | 0.1525 | 0.4165 | 0.1140 | 0.3160 |

Table 8. The mean of F1 and F2 for long vowels for all speakers.

| Volunteer No. | a:_Vowel | | i:_Vowel | | u:_Vowel | |
|---------------|------------|-------------|------------|-------------|------------|-------------|
| | F1 | F2 | F1 | F2 | F1 | F2 |
| 1 | 512 | 1101 | 312 | 2285 | 339 | 697 |
| 2 | 672 | 1613 | 353 | 1588 | 376 | 1040 |
| 3 | 517 | 1801 | 277 | 1725 | 281 | 708 |
| 4 | 220 | 611 | 315 | 2302 | 313 | 634 |
| 5 | 436 | 1244 | 228 | 1166 | 277 | 772 |
| 6 | 485 | 908 | 332 | 1667 | 335 | 783 |
| 7 | 626 | 797 | 266 | 540 | 307 | 611 |
| 8 | 627 | 1401 | 296 | 717 | 383 | 918 |
| 9 | 634 | 1065 | 272 | 560 | 343 | 811 |
| 10 | 544 | 734 | 350 | 578 | 358 | 777 |
| 11 | 403 | 941 | 377 | 2297 | 384 | 835 |
| 12 | 523 | 1369 | 339 | 2212 | 324 | 756 |
| 13 | 445 | 899 | 323 | 1975 | 303 | 713 |
| 14 | 647 | 1479 | 337 | 1638 | 364 | 777 |
| 15 | 445 | 936 | 296 | 1991 | 331 | 843 |
| 16 | 539 | 909 | 354 | 523 | 403 | 791 |
| 17 | 533 | 987 | 247 | 399 | 303 | 576 |
| 18 | 530 | 1315 | 297 | 608 | 303 | 627 |
| 19 | 660 | 1576 | 317 | 2253 | 308 | 648 |
| Mean | 526 | 1141 | 310 | 1422 | 333 | 754 |
| Min | 220 | 611 | 228 | 399 | 277 | 576 |
| Max | 672 | 1801 | 377 | 2302 | 403 | 1040 |

Table 9. The mean, SD, and CVar of F1 and F2 for long vowels.

| | all speakers | | | | | |
|-------------|--------------|----------|---------|----------|---------|----------|
| | a: | | i: | | u: | |
| | F1 | F2 | F1 | F2 | F1 | F2 |
| Mean | 526 | 1141 | 317 | 1782 | 333 | 754 |
| SD | 128.3245 | 429.1378 | 41.2598 | 814.9596 | 36.7849 | 116.1518 |
| Cvar | 0.2439 | 0.3760 | 0.1332 | 0.5730 | 0.1103 | 0.1541 |

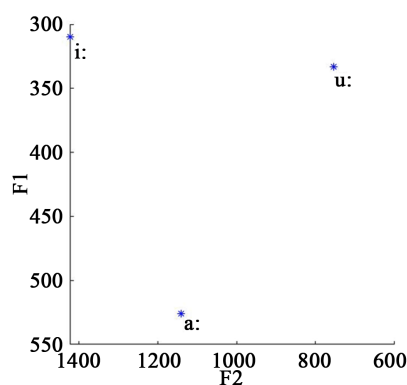


Figure 3. Formants' means of Arabic long vowels.

All the CVar values are smaller than 1 which is considered low, which shows that the measurements of the formants appear to be close to the mean. Therefore, the results are accurate and reliable. In addition, when compare the formants means of long vowels with short ones, see **Table 10**. It can clearly be seen that the values are close to each other which emphasizes that short and long vowels represent the same sound, the only difference is their durations.

6.3. Vowels' Formants versus IPA and Some Studies

In the late nineteenth century, an alphabet, which is a set of internationally recognized phonetic symbols, was created by the IPA based on the idea to assign a symbol to each distinctive sound. **Figure 4** shows the 2005 revised copy of the IPA chart. [4] Vowels with the same symbols and similar sounds of Arabic short vowels (a, i, u) could be seen approximately in the same positions as the IPA ones. They all take a shape of an upside-down triangle.

In addition, the researches that were performed at the IMMII Laboratory at Hassan First University, Settat, Morocco, [12] extracted the formants for Arabic vowels and the results are as follows:

- F1 for /a/ vowel is between 500 - 800. F2 is between 1000 - 1500
- F1 for /i/ vowel is between 100 - 400. F2 is between 2000 - 3000
- F1 for /u/ vowel is between 300 - 600. F2 is between 600 - 1100.

The other study was done in the Department of Computing Science, Stirling University, UK, [3] investigated the formants for Arabic vowels, the results and the means are as follows:

- F1 for /a/ vowel is 400. F2 is 800
- F1 for /i/ vowel is 400. F2 is 2100
- F1 for /u/ vowel is 700. F2 is 1100

It can be seen that both studies represent the same pattern which is a shape of an upside-down triangle, which means that the extracted values in this research are accurate and reliable.

6.4. The Speakers' Nationalities and the Formants of the Vowels

Since this research is concerning only the MSA, the nationality of the speaker

should not affect the measured results. However, the pronunciation of a small number of speakers might be affected from their dialects. So, it is good to see if every formant that is measured for nationality follows the pattern of Arabic vowels. Libyans and Egyptians were considered as a case study to investigate that. The following equation calculates the percentage difference or percentage error:

$$\text{percentage difference} = \frac{|\text{Libyan value} - \text{Egyptian}|}{((\text{Libyan value} + \text{Egyptian})/2)} * 100$$

The percentage difference is calculated between Libyan and Egyptian F1 and F2. See **Table 11**.

Most of the results are less than 11, except the F2 in vowels /i/ and /u/ are a little bit high, this may be indicating the impact of nationality. However, the research is concerning only the MSA.

On the other hand, F1 and F2 measurements for Libyans and Egyptians are precise, and close to each other. Furthermore, they both reflect the pattern of short Arabic vowels, **Figure 5** illustrates that.

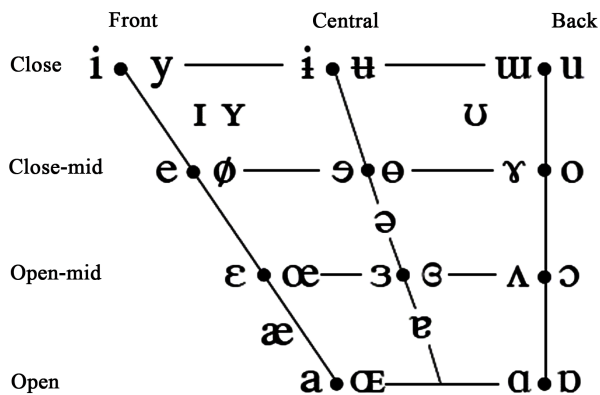


Figure 4. The international phonetic alphabet triangle [4].

Table 10. Comparing the short and long vowels.

| | F1 | F2 | F1 | F2 | F1 | F2 |
|-----------------|-----|------|-----|------|-----|-----|
| Mean a, i, u | 560 | 1138 | 369 | 1415 | 371 | 883 |
| Mean a:, i:, u: | 526 | 1141 | 310 | 1422 | 333 | 754 |

Table 11. The percentage difference of f1 and f2 for libyan vs egyptian speakers.

| Vowel | Formant | Libyan | Egyptian | Percentage Difference |
|-------|---------|--------|----------|-----------------------|
| a | F1 | 601 | 540 | 10.7 |
| | F2 | 1230 | 1129 | 8.6 |
| i | F1 | 390 | 354 | 9.7 |
| | F2 | 1537 | 1161 | 27.9 |
| u | F1 | 382 | 361 | 5.7 |
| | F2 | 932 | 748 | 21.9 |

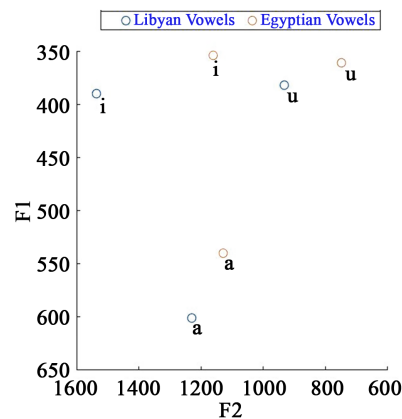


Figure 5. The mean of F1 and F2 for Libyan vs Egyptian vowels.

7. Conclusion

This paper concerned only MSA. The CVCVCV Arabic word structure was used to extract short vowels (a, i, u) and the Arabic words with the CVVC pattern were also used to extract long vowels (a:, i:, u:). Since the F1 and F2 measurements express the vowels, the LPC method was used to measure them. The SD and CVar were calculated to evaluate the measurements. From the results, it can be concluded that the data is accurate and reliable. In addition, the formants follow the pattern of the other studies which formulate an upside-down triangle. The formants of short vowels were compared to the formants of long vowels. From that comparison, it is found that the values are close which emphasize that the Arabic long vowels are just an elongation for the short ones. The differences between formants of two nationalities were also compared, and a percentage difference was calculated. The results showed that the nationality does not affect the formant results, because the study only concerned MSA, except the F2 in vowels /i/ and /u/ which indicate that the dialect from each nationality might have a slight effect on the measurements. Finally, the outcomes of this study are useful in speech processing tasks, such as vowel recognition and voice classification. In addition, the study may be developed and extended using corpora with female and children speakers.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Thomas, M. (1992) Improved Modeling of Out-Of-Vocabulary Words Using Morphological Classes. No. 1.
- [2] Chen, C.J. (2015) Elements of Human Voice. World Scientific. <https://doi.org/10.1142/9891>
- [3] Gokulan, M., Gandhi, M., Joshi, S. and Karamchandani, S. (2013) Objective Speech Analysis and Vowel Detection. *International Journal of Computer Applications*, No.

- 1, 22-26.
- [4] The International Phonetic Alphabet, "IPA: Vowels" (2005) <https://www.internationalphoneticassociation.org/content/ipa-vowels>
 - [5] Jones, D. and Wells, J. (2015) Vowel Triangle, Cardinal Vowels. https://en.wikipedia.org/wiki/File:Vowel_triangle,_cardinal_vowels.png
 - [6] Hillenbrand, J., Houde, R.A., Drive, T.P., Society, A. and Louis, S. (1995) Vowel Recognition: Formants, Spectral Peaks, and Spectral Shape. *The Journal of the Acoustical Society of America*, **98**, 2949. <https://doi.org/10.1121/1.414088>
 - [7] Cherif, A., Bouafif, L. and Dabbabi, T. (2001) Pitch Detection and Formant Analysis of Arabic Speech Processing. *Applied Acoustics*, **62**, 1129-1140. [https://doi.org/10.1016/S0003-682X\(01\)00007-X](https://doi.org/10.1016/S0003-682X(01)00007-X)
 - [8] Holmes, J.N., Holmer, W.J. and Garner, P.N. (1997) Using Formant Frequencies in Speech Recognition. In: *EUROSPEECH97 5th European Conference on Speech Communication and Technology*.
 - [9] Newman, D. and Verhoeven, J. (2002) Frequency Analysis of Arabic Vowels in Connected Speech. *Antwerp Papers in Linguistics*, **44**, 77-87.
 - [10] Hillenbrand, J. and Getty, L. (1995) Acoustic Characteristics of American English Vowels. *The Journal of the Acoustical Society of America*, **49008**, 3099-3111. <https://doi.org/10.1121/1.411872>
 - [11] Elharati, H.A., Alshaari, M. and Kępuska, V.Z. (2020) Arabic Speech Recognition System Based on MFCC and HMMs. *Journal of Computer and Communications*, **8**, 28-34. <https://doi.org/10.4236/jcc.2020.83003>
 - [12] Farchi, M., *et al.* (2019) Arabic Vowels Acoustic Characterization to Cite this Version: HAL Id: hal-02296812.