A Preliminary Study on Malay Vowel Formant for Young Adult and Aged People

Norezmi Jamal¹, Shahnoor Shanta¹*, Fadhilah Rosdi², Nabilah Ibrahim¹,
¹Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia.
²Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia.

E-mail: shahnoor@uthm.edu.my

Abstract. This paper presents a study on Malay vowel formants of different age group between young adults and aged people. About 12 of the available young adult speech (6 males and 6 females; mean age 23 years) and 4 of the recorded aged people speech (2 males and 2 females; mean 67.3 years) were examined for analysis of formants to study the effect of age and gender. This work is significant in studying the hypothesis that the formant of aged people is lower than young adult population. The findings obtained shows that the formant of aged group is not critically affected by their age and male groups produces lower formant frequencies compared to the female groups either young adult or aged people. Formant frequency values also subject to variation due to the influence of surrounding phonemes.

1. Introduction

In recent years, the numbers of people suffering from stroke have increased gradually. In Malaysia, as many as six people experience a stroke every hour [1] and it is one of the top five leading causes of death [2]. Most of them are aged people. An aged person can be defined as a person whose age is above 60 years old [3]. They might suffer from communication disorders following a stroke. In communication disorder, a stroke survivor may have language impairment and articulation problem, which is known as aphasia [4]. To assess and diagnose this disorder, most clinicians and researchers use acoustic analysis tools [5]. The advantage of this analysis is to differentiate the pathologic and normal voice with the given healthy population.

Formant is one of the essential acoustic properties that will be considered in this study. It is known as the resonating frequency of the vocal tract [6]. In spoken language, vowels play an important role by carrying the linguistic information about any spoken word or sentence. Without vowel, the sound of word might not occur. Although some research has already been carried out on the acoustic property analysis of vowels, especially the formants, the point to note here is that it is normally incomparable if resulting from the cross-language analysis of phonetic units such as English, Swedish, Russian, Turkish, Bangla and Mandarin [5, 7-9]. For example, Kuhl et. al. claimed that the acoustic forms of speech are highly inconstant, changing with few reasons that include speaker gender, speaking rate, and the phonological context of the sound, which makes the task of sorting ambient language sounds into phonetic categories a complex one [7].

Demirhan et. al. in their study on acoustic voice analysis of young Turkish speakers, reported that the fundamental frequency, F0 of vowels (/a/, /i/ and /u/) for female were greater than male [5].
The authors investigated the F0, perturbation parameters, and noise-to-harmonics ratio (NHR) for three sustained vowels in a group of young Turkish speaking adults. The perturbation values were compared in the study, which significantly changed for some pairs of vowels. On the other hand, there was no significant difference between NHR values of three vowels in men. The study is conducted, where the acoustic analysis is only limited to adult Turkish speakers with a small sample data [5]. Moreover, Ghosh et. al. reported the distinct Bangla vowels exhibiting different formant frequencies [8]. While, Liu and Ng studied the formant characteristics associated with the vowels produced by esophageal and laryngeal speakers of Mandarin [9]. The results indicated that the esophageal speakers were associated with significantly higher formant frequencies and a significantly diminished vowel space circumscribed by the three corners vowels (/a/, /i/, and /u/) compared to laryngeal speakers.

A considerable amount of literature has already been published on Malay vowel based acoustic properties [10, 11]. For example, Ting et. al. and Zourmand et. al. claimed that Malay acoustic properties of female is higher than male irrespective of young adult or children, especially considering the fundamental frequency and the formants of vowels. However, there are significant differences between the vowels reported for all perturbation measures in Malaysian Malay males. Rosdi et. al. also investigated the effect of speech rate on formant frequency of Malay vowels [12]. The results showed that there is a small variant, where the changing of second formant, F2 values from vowels spoken at a fast rate as compared to those spoken at a normal rate.

In short, formants can differ depending on the configuration of the vocal tract for articulating different types of voiced sounds. It can be noted that normally vocal tract of female is shorter than male [13]. Due to this, female formant is higher than the male one. Besides, the effect of age also can cause the difference on the formant as discussed in the age-related changes in acoustic characteristics of adult speech article [14]. The focus of the article was the aged population and English language. Still, there is not much study reported on Malay aged people formant analysis. Due to this, the aim of this work is to discuss the comparison of Malay aged people and young adult formant analysis. This paper is organized as follows. In section 2, the previous related works on the formant of vowels are discussed. Section 3 provides the methodology of formant analysis from a given audio file. In section 4, the results and discussion are presented. Finally, our work of this paper is summarized in the last section together with the future direction relevant to this work.

2. Related Works on Analysis of Formants Resulting from Malay Vowels

Basically, the different vowels could be represented with distinct characteristics formants due to the different phonetic sounds. The formant frequencies are the oscillation of frequencies that produces high amplitude of sounds (resonances) along the distinct cross-sectional area of the vocal tract [15]. In simple terms, vowels are the most significant sounds that often get affected. This is because; the vowels are produced from the excitation of fixed vocal tract shape and the vibration of the vocal cord [15]. The vowels sounds are also primarily based on the positions of tongue, jaws, lips and velum [15]. There are three typical formants, namely as the first formant (F1), the second formant (F2), and the third formant (F3).

A few researches have been carried out, examining the effect of different age, gender and ethnicity on the formant of vowels [16-17]. Table 1 shows the reported baseline formant values (F1, F2 and F3) of vowels (/a/, /u/ and /i/) for male and female speakers which were obtained by different authors in their respective research articles. The baseline value is calculated from the average of the research findings on the formant analysis of Malay native speakers considering different age and ethnicity. It can be seen that the formants of each vowel for male speaker is lower than that of female speaker, where F1, F2 and F3 were reported. Ting et. al. investigated the first three formant frequencies for the control Malaysian Indian young adult groups [16] and Malaysian Malay children group [17]. For Malay Indian young adults, the F1, F2, and F3 of average vowels for male speakers were reported at 359 Hz, 1457 Hz and 2799 Hz respectively [16]. While F1, F2 and F3 of average vowels for Malaysian Indian females were reported at 485 Hz, 1707 Hz, and 3081 Hz respectively.
[16]. Malaysian Indian females had significant higher formant frequencies in all vowels than those of Indian male speakers. It is similar for the children speakers, where the results showed that, Malay female children had higher formant frequencies compared to male speakers [17].

Table 1. The baseline formant values of vowels based on the previous related works.

| Authors Details | Vowels | Male Speakers | Female Speakers |
|-----------------|--------|---------------|-----------------|
|                 |        | F1           | F2           | F3           | F1           | F2           | F3           |
| [16] Indian Speakers | /a/    | 445          | 1155         | 2703         | 769          | 1458         | 3045         |
|                  | /u/    | 320          | 853          | 2763         | 369          | 872          | 2853         |
|                  | /i/    | 263          | 2376         | 3193         | 324          | 2763         | 3521         |
| [17] Children Speakers | /a/    | 749          | 1454         | 3225         | 787          | 1530         | 3247         |
|                  | /u/    | 408          | 950          | 3012         | 436          | 981          | 2997         |
|                  | /i/    | 353          | 2621         | 3573         | 361          | 2428         | 3584         |
| Baseline (Average) | /a/    | 597          | 1304         | 2964         | 778          | 1494         | 3146         |
|                  | /u/    | 364          | 902          | 2888         | 403          | 927          | 2925         |
|                  | /i/    | 308          | 2499         | 3383         | 343          | 2596         | 3553         |

3. Methodology
In this study, 16 of the native Malay speakers were categorized into two large groups, which are: (i) young adult (YA) and (ii) aged people (AP). For young adult group, the audio files were used from the available Malay digit speech corpus (databases) [18]. 12 of the recorded speech signals were examined including 6 male (YM) and 6 female speakers (YF). Their mean ages are 23 years. And, for aged people, 4 speakers including 2 males (AM) and 2 females (AF) ranging in age from 61 to 79 years were evaluated. Aged people speech data were recorded using a high performance Keystudio microphone attached with the pop filter at a clean environment. The details of the recorded audio data are; mono channel, 16-bit resolution and a sampling rate of 16.0 kHz respectively. All the speakers were given a task to say Malay digit words (empat, tujuh and sembilan) for three times. The vowels (/a/, /u/ and /i/) of three selected Malay digit words as shown in Figure 1 were analysed for acoustic properties. All the recorded data were imported into Praat speech analysis software. The formants frequencies (F1, F2 and F3) of Malay vowels were analysed by highlighting the desired vowel area. Formants of each Malay spoken digits were identified by averaging the value of formants from three different sets of similar spoken words. The resulting waveform for the word empat is shown in Figure 1 (a). While the word tujuh and sembilan are respectively illustrated in Figure 1 (b) and 1 (c).
4. Results and Discussion

The formant values of each vowels (/a/, /u/, /i/) were compared for young adult (YA) and aged people (AP) speakers. Figure 2 shows the distribution of formants (F1, F2 and F3) for different vowels. For the vowel class /a/, young male (YM), young female (YF), aged male (AM) and aged female (AF) are clearly grouped in the same class. As shown in Figure 2, the AF group that is represented by red star legend are found to have significantly more formant values of /a/ vowel than the other three groups. In contrast to earlier findings in the age-related changes in acoustic characteristics of adult speech for English vowels reported in article [14], however, no evidence for the age of female population, affecting the formant values was detected. The authors reported that young adult female produced higher formants of /a/ vowel among other groups [14].

Figure 1. Waveform and its spectrogram of Malay spoken digits

Figure 2. The distribution of formant values for different vowels from 16 Malay native speakers
The results obtained, also showed the outliers for /i/ and /u/ of the vowel class. The formant value of /i/ vowel for aged male was categorized into the /u/ vowel class. The overlapping as seen in the results obtained might be related to the gender and age differences and also due to the co-articulation [19]. Another reason might be that the aged male group might have slower speaking rates than the younger one. This unexpected result could be attributed to the fact that the formant values of /i/ and /u/ vowels keep continuously changing due to the influence of the neighbouring phonetic sounds as shown in the spectrogram of Figure 1 (b) and Figure 1(c) respectively. Table 2 shows the mean of the first (F1), second (F2), and third (F3) formant frequencies for each vowel (/a/, /i/, /u/) produced by young adult male (YA-M), young adult female (YA-F), aged people male (AP-M), and aged people female (AP-F) obtained from the preliminary analysis of data in Figure 2. Obviously, the mean formant value (F1, F2 and F3) of each vowel for male is lower than those of female groups. For example, for /a/ vowel, young adult and aged people male speakers respectively produced 683 and 680 Hz of F1 which are lower compared to the young adult and aged people female speakers. So, the results obtained demonstrated the similarity with other findings as discussed in section 2. The reason behind this is that the gender significantly affects the formant values; mainly because of the mean size of male adult vocal tract is normally about 17 cm, longer than that of the female and children population [13].

Table 2. Mean first (F1), second (F2), and third (F3) formant frequencies of vowels for four different groups of speakers

| Groups | Vowels | F1 (Hz) | F2 (Hz) | F3 (Hz) |
|--------|--------|---------|---------|---------|
|        |        | M       | F       | M       | F       |
| YA     | /a/    | 683     | 793     | 1387    | 1663    | 2528    | 2845    |
|        | /u/    | 331     | 366     | 1382    | 1765    | 2503    | 2792    |
|        | /i/    | 338     | 444     | 2063    | 2231    | 2967    | 3226    |
| AP     | /a/    | 680     | 868     | 1254    | 1753    | 2245    | 3033    |
|        | /u/    | 329     | 373     | 1093    | 1162    | 2609    | 2808    |
|        | /i/    | 342     | 470     | 1930    | 2432    | 2937    | 3301    |

5. Conclusion
The study investigated the first three formant frequencies of /a/, /u/ and /i/ vowels for normal Malay young adult and aged people speakers. The results found, demonstrates that, generally, Malay female groups have higher formant frequencies compared to male groups. Moreover, the aged people group have lower formant values due to their slower speaking rate. Besides, it is also found that, there is formant pattern ambiguity of vowel sounds for /i/ and /u/. This is because of the formant values of /i/ and /u/ vowels keep continuously changing due to the influence of the neighbouring phonetic sounds. As a result, the findings show that formants are very useful features for speech recognition or early diagnosis for any speech disorder, but they have not been widely used because of the difficulty in estimating them. So, Mel Frequency Cepstrum Coefficient (MFCC) analysis or other feature extraction techniques could be a possible avenue for future research.

Acknowledgement
This work is financially supported by the Short Term Grant (STG), vot number 530 under Universiti Tun Hussein Onn Malaysia. Also, highly appreciated to Microelectronics & Nanotechnology - Shamsuddin Research Centre (MiNT-SRC), Universiti Tun Hussein Onn Malaysia for providing research work space.
References

[1] O W Mohamad, Z Taib M and Jailani R 2014 Systems, Process and Control (ICSPC) 2014 157
[2] Loo K W and Gan S H 2012 International Journal of Stroke 7 165
[3] Ismail N N N N, Lokman A M and Redzuan F 2016 e-Learning, e-Management and e-Services (IC3e) 133
[4] Jamal N, Shanta S, Mahmud F and Sha’abani M 2017 AIP Conference Proceedings 020028.
[5] Demirhan E, Unsal E M, Yilmaz C, and Ertan E 2016 Journal of Voice 30 378
[6] Rabiner L R and Schafer R W 2011 Pearson Upper Saddle River (NJ) 64
[7] Kuhl P K, Andruski J E, Chistovich I A, Chistovich L A, Kozhevnikova E V, Ryskina V L, Stolyarova EI, Sundberg U and Lacerda F 1997 Science 277 684
[8] Ghosh T, Saha S and Ferdous A I 2016 An International Journal (SIPIJ) 7
[9] Liu H and Ng M L 2009 Journal of voice 23 255
[10] Ting H N, Chia S Y, Hamid B A and Mukari S Z M S 2011 Journal of Voice 25 305
[11] Zourmand A, Mahmoudi A, Marimuthu J and Ting H N 2012 Intelligent Systems, Modelling and Simulation (ISMS) 379
[12] Rosdi F, Salim S and Mustafa M 2014 Advanced Science, Engineering and Medicine 6 127
[13] Huang X, Acero A, Hon H W and Reddy R 2001 Spoken language processing: A guide to theory, algorithm, and system development (Prentice hall PTR)
[14] Torre P and Barlow J A 2009 Journal of communication disorders 42 324
[15] Rabiner L R and Juang B H 1993 Fundamentals of speech recognition
[16] Ting H and Lee C 2013 World Congress on Medical Physics and Biomedical Engineering 375.
[17] Ting HN, Zourmand A, Chia S Y, Yong B F and Hamid B A 2012 Journal of Voice 26 664
[18] Al-Haddad S A R, Samad S A, Hussain A and Ishak K A 2008 American Journal of Applied Sciences 5 714
[19] Maurer D, D’heureuse C, and Landis T 2000 International Journal of Neuroscience 39