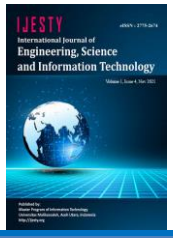
International Journal of Engineering, Science & InformationTechnology (IJESTY)
Volume 1, No. 4 (2021) pp. 25-30
eISSN: 2775-2674
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v1i1.147
Research Paper, Short Communication, Review, Technical Paper



Analysis of Human Voice Spectrum based on Regional Accent in Vowels and Consonants

Puspa Indah Devitasari*, Benazir Amalia Firdausy, Shobhi Al-Ghifari Azhary, Heru Kuswanto

Department of Natural Science Education, Graduate Program, Yogyakarta State University, Yogyakarta, Indonesia
**Corresponding author E-mail: puspaindah.2018@student.uny.ac.id*

Manuscript received 15 August 2021; revised 1 Sept 2021; accepted 15 Sept 2021. Date of publication 4 Nov 2021

Abstract

This study aims to analyze the spectrum of the human voice based on the accents of several regions. This research is included in the type of experimental research. Data were collected from 4 male respondents aged 23 years as research subjects. Respondents come from Central Java, East Java, Jambi, and South Sulawesi or come from Javanese, Malay, and Bugis ethnicities. Each respondent pronounces the vowels /a/ and /e/ as well as the consonants /b/, /d/, and /g/ once in each sound under normal circumstances or does not have throat problems or other things that make the sound abnormal. Voice recording is done in a closed and quiet room without any disturbance by using a microphone connected to a laptop using the cool record edit pro application. The sound spectrum was analyzed using Praat verse 2 software, then the recording was converted to *wav format. Furthermore, the sound spectrum analysis procedure uses software to determine the shape of the spectrum, the maximum or minimum pitch frequency (F0), and the formant frequency of F1-F4 in each sample. The results showed that the response of men from South Sulawesi had the highest tone of 156 Hz in the letter /b/ and 145 Hz in the letters /d/ and /g/ respectively. while the lowest tone of voice owned by men from the Jambi area is 105 Hz, 109 Hz, and 102 Hz for each consonant and the output from F0 (pitch) to F1-F4 (formant) can change and does not describe the high pitch, meaning the formant will always high.

Keywords: *Voice Spectrum, Regional Accent, Consonant, Pitch, Formant*

1. Introduction

Indonesia has diverse tribes and cultures with different trademark. Indonesia had more than 200 ethnic groups and 1,340 ethnic tribes, such as: ethnic Javanese, Minang tribe, Malay, Batak, tribe of Bugis etc. [1]. Those tribes have different accents influenced by the geographical location and environmental conditions. Due to these different accents, the frequency of sound is also difference. According to census conducted by the Indonesia Central Bureau of Statistics (BPS) [2], in 2010, there were 237,641,326 people around Indonesia in 34 provinces [3].
Podesva and Smith carried out study the change in the resonance frequency on the vowels in 1999. They used acoustic vowel at the Toba Batak tribe and Bugis and resulted of the frequency distribution of formant on sound /i/ and /o/ in the second parts overlap in distribution ellipses [4].
In vowels, there are the difference of voice spectrum when an individual speak /a/, /i/, /u/, /e/, /o/ because there is special characteristics taken as identifier [5].
The study analyzes the spectrum of human voice based on regional accent in 23-year-old individuals by speak the vowels and consonants. The sound spectrum is analyzed using software Praat verse 2 it can detect the frequency all the regional accent [6] [7].

2. Literature Review

2.1 Pitch

The vibrating frequency of the vocal cords are also referred to as the fundamental frequency (basic) by notation F_0 is referred to as pitch. Gaol, Erwin, and Ginting [1] mentions that the pitch or fundamental frequency humans is determined by several things such as age, sex, and length of the vocal cords [8].
Rahim, distinguish three categories of age to determine the pitch man [9]. In the male group of great kids pitch ranges on average between 210-270 Hz, the adult group ranges from 120-150 Hz, and in the elderly group ranges from 100-140 Hz. In women groups of children have a voice pitch of 200-290 Hz, the adult group of 200-280 Hz, and in the elderly group ranged between 160-200 Hz.

This pitch analysis can be used to perform voice recognition to the voice of someone that is through statistical analysis of the pitch minimum value, maximum mean pitch, and pitch.

2.2 Formant

Formant frequencies is the resonance of the filter, the vocal tract (articulator) forwarding and filtering sound output (output) in the form of words that have meaning [10]. The image above is the formant frequencies generated image of a person. Horizontal lines that make up the tape from the bottom to the top is referred to as Formant 1 (F_1), Formant 2 (F_2), formant 3 (F_3), and so on [5]. In general, the formant frequencies are not limited, but to identify the formant person is usually analyzed namely, Formant 1 (F_1) and Formant 2 (F_2) is usually associated with the position of the tongue. Formant 3 (F_3) and Formant 4 (F_4) associated with the color (timbre) of sound produced [11].

2.3 Spectrograms

Spectrograms is a representation of the spectral (color noise) which varies with time which indicates the level of density (energy intensity) spectral [12]. In other words, spectrograms are the visualization of each formant values that are equipped with the energy level that varies with time [13]. Energy levels, known by the term formant bandwidth. Spectrograms to form a common pattern that is typical in the pronunciations and special patterns of each formant in the pronunciation of syllables, so spectrograms also be used to analyze a person's voice identification. If the unknown recording duration is long.

3. Method

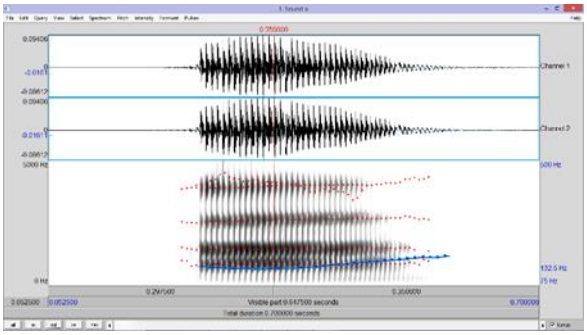
This study included in this type of experimental research. The tools used in this study is a microphone to record sound as well as laptops for analyzing sound. Applications used to record the sound that is cool record edit pro and applications used to analyze the spectrum of sound is praat.

This study uses 4 respondents which are 4 male who came from various regions in Indonesia. They are from Central Java, East Java, Jambi, and South Sulawesi or come from ethnic Javanese, Malay tribe, and the tribe of Bugis. Everyone pronounces the vowel /a/ and /e/ and the consonants /b/, /d/, and /g/ once. Sound is recorded through a microphone that was connected to a laptop using a cool record edit pro applications, here in after respondent pronounce vowels and consonants with predetermined intervals, then the recording converted to the format **wav*.

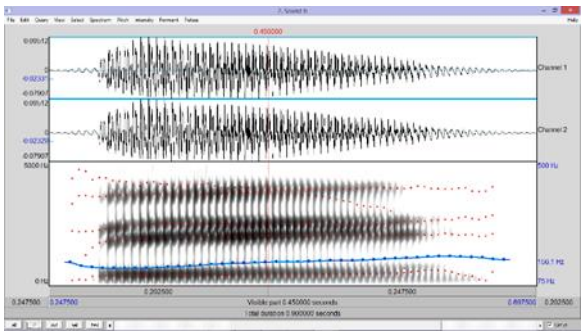
Furthermore, the sound spectrum analysis praat using software to determine the shape of the spectrum, the frequency of maximum or minimum pitch, and the frequency of formant from F_1 until F_4 in each sample [14].

4. Result and Discussion

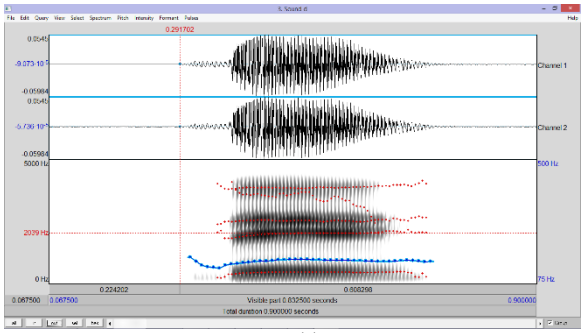
In this section we will discuss the results of the analysis of the spectrum of the human voice when pronouncing the vowel /a/ and /e/ and the consonants /b/, /d/ and /g/.



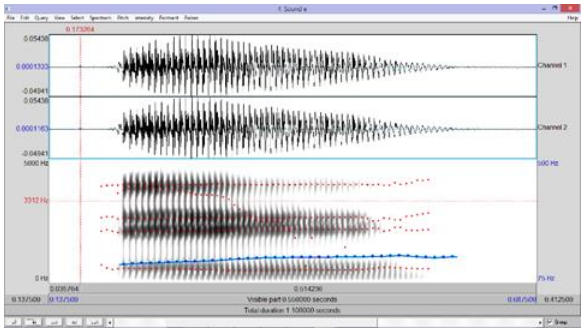
(a)



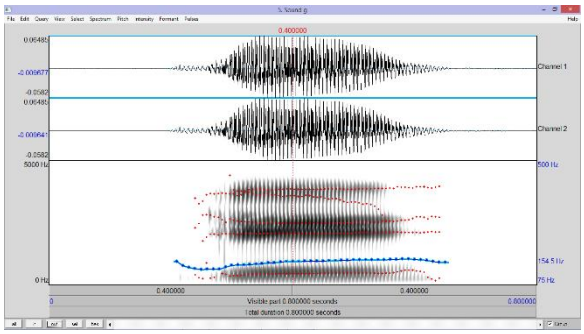
(b)



(c)



(d)



(e)

Fig 1. The spectrum of the sample male voice said (a) the vowel /a/, (b) the consonants /b/, (c) the consonant /d/, (d) the vowel /e/, and (e) consonant /g/.
Source: from primary data

The figure 1 is a picture of one of the sound spectrum research samples that is a man's voice. There was the spectrum of sample male voice said the vowel /a/ and /e/ and consonant /b/, /d/, and /g/.

Once known spectrum and frequency analysis to determine the pitch of the fundamental frequency (basic) or F_0 of each sample by regional accent. Furthermore, based on the image of the spectrum, will obtain data from the analysis as follows.

4.1 Pitch Sounds Great Relationship with the Regional Accent

Everyone had great pitch frequency sound different based regional accent. The pitch sounds great in East Java, Central Java, Jambi, South Sulawesi, and it can be seen in the following graph.

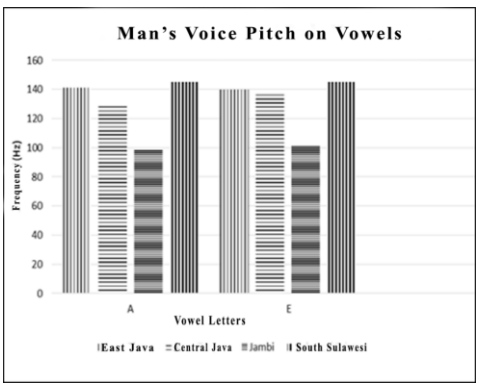


Fig 2. Graph value of the fundamental frequency (F_0) of pitch male voice on the vowels.
Source: from primary data

Based on figure 2, the highest pitch owned by a man from South Sulawesi is equal to 145 Hz on the pronunciation of the vowels. While that has lowest pitch frequency by men of Jambi is equal to 99 Hz on the letter /a/ and 101 on the letter /e/. Men from East Java and Central Java have great pitch frequency on the vowels are nearly equal, namely around 130 Hz-141 Hz.

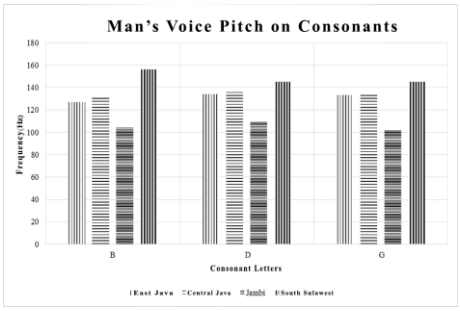


Fig 3. Graph value of the fundamental frequency (F_0) of male voice pitch on consonants.
Source: from primary data

Based on figure 3 the man from South Sulawesi has the highest voice pitch that is equal to 156 Hz on the letter /b/ and respectively 145 Hz on the letter /d/ and /g/. while the lowest voice pitch owned by men from Jambi area that is equal to 105 Hz, 109 Hz, and 102 Hz for each consonant mentioned sequentially.
Based on the data, we found that the man from Bugis culture, it comes from South Sulawesi, has the highest voice pitch of all and Malay culture who come from Jambi, has the lowest pitch. According Mulyani, et al [15] the cultural habits of each region to be factors that lead to the fundamental frequency of each person is different. Harmaini [16] showed that Malay ethnic have more neutral characteristic rather than a form of emotional expression between Javanese culture that is not expressive with the Batak and Minang cultures are more expressive.

4.2 Large Formant Sound

Formant is the resonance frequency of the filters that are forwarding and filtering sound output. In praat software that is used to know the level of the human voice formant starting from F_1 to F_4.

Table 1. Values Formant Sound Man

Subject	Regional Accent	Formant	Vocal Word		Consonant Word		
			A	E	B	D	G
A	East Java	F_1	871.9	484.4	469.9	409.7	465.7
		F_2	1360.9	2174.4	2136.5	2016	2131.1
		F_3	2907.8	2817.3	2744.8	2580.1	2671.2
		F_4	3339.5	3806.1	3658.3	3458.1	3810.3
H	Central Java	F_1	810.1	598.1	412.7	390.5	377
		F_2	1415	1919.5	2088.7	2084.3	2093.8
		F_3	2742.6	2500.9	2693.7	2766.2	2699.6
		F_4	4179.1	3807.8	3573.4	3745.6	3619.4
G	Jambi	F_1	927.8	533.6	467.9	470.3	502.6
		F_2	1429.6	2079	2037.3	2134.4	2177.1
		F_3	2640	2772	2579.8	2621.7	2776.1
		F_4	3402.7	3587.4	3520.4	3515.5	3862.3
F	South Sulawesi	F_1	851.8	535.8	432.2	405.6	462.7
		F_2	1505.2	2027.2	1954.4	2004.9	2099.9
		F_3	2590	2554.2	2552.1	2548.2	2646
		F_4	3963	3496.2	3646.6	3634.1	3631

Source: from primary data

Based on table 1, first, we can focus on the vowel words. The Formant F_1 on the vowel /a/ has the higher frequency than the vowel /e/ equally distributed in all regions. The highest F_1 in word /a/ come from Jambi that is 927,8 Hz and the lowest come from Central Java that is 810,1 Hz. The highest F_1 in word /e/ come from Central Java that is 598,1 Hz and the lowest come from East Java that is 484,4 Hz.
In F_2 on the vowel words, the Formant F_2 on the vowel /a/ has the lower frequency than the vowel /e/ equally distributed in all regions. The highest F_2 in word /a/ come from South Sulawesi that is 1505,2 Hz and the lowest come from East Java that is 1360,9 Hz. The highest F_2 in word /e/ come from East Java that is 2174,4 Hz and the lowest come from Central Java that is 1919,5 Hz.
In F_3 on the vowel words, man from Jambi has the lower frequency in vowel /a/ that is 2640 Hz than vowel /e/ that is 2772 Hz. The highest F_3 in word /a/ come from East Java that is 2907,8 Hz and the lowest come from South Sulawesi that is 2590 Hz. The highest F_2 in word /e/ come from East Java that is 2817,3 Hz and the lowest come from Central Java that is 2500,9 Hz.

In F_4 on the vowel words, men from Central Java and South Sulawesi have the higher frequency in word /a/ than /e/. The highest F_4 in word /a/ come from Central Java that is 4179,1 Hz and the lowest come from East Java that is 3339,5 Hz. The highest F_4 in word /e/ come from Central Java that is 3807,8 Hz and the lowest come from South Sulawesi that is 3496,2 Hz.
In the consonants, formant F_1 on the words /b/, /d/, and /g/ has the highest frequency in man from Jambi which are 467,9 Hz; 470,3 Hz; and 502,6 Hz and the lowest come from Central Java which are 412,7 Hz; 390,5 Hz; and 377 Hz.
In F_2, South Sulawesi has the lowest frequency on words /b/, /d/, and /g/ which are 1954,4 Hz; 2004,9 Hz; and 2099,9 Hz and fluctuative frequency come from other regions.
In F_3, South Sulawesi has the lowest frequency on words /b/, /d/, and /g/ which are 2552,1 Hz; 2548,2 Hz; and 2646 Hz and fluctuative frequency come from other regions.
In F_4 there are fluctuative frequency but man from South Sulawesi has the most stable frequency in range 3600 Hz than the other regions.
From the data, we found that formant doesn't affect the high of pitch. South Sulawesi, which on average has the lowest formant, has the highest pitch of votes compared to other regions.
Kuang [17] found that both tense voice and creaky voice share the similar spectral balance, with more energy in the high-frequency domain, but they have distinctive frequency and pitch profiles.
According to Azhar [12], because the pitch is the basic tone while the formant is the resonant frequencies of the filter so the output from F_0 (pitch) to F_4 can change and does not illustrate that high pitch means the formant will always be high.

5. Conclusion

Based on the research that has been done, it was concluded that:
1. Pitch for male voice on the vowels /a/ and /e/ and the consonants /b/, /d/ and /g/ obtained the highest is from South Sulawesi and the lowest is from Jambi.
2. The output from F_0 (pitch) to F_1, F_2, F_3, and F_4 can change and does not illustrate that high pitch means the formant will always be high.

References

[1] N. L. Gaol and M. Ginting, "ANALISIS SPEKTRUM SUARA MANUSIA BERDASARKAN SUKU PADA KELOMPOK USIA ANAK-ANAK DENGAN MENGGUNAKAN SOFTWARE PRAAT."
[2] Badan Pusat Statistik, "Hasil Sensus Penduduk 2010," 2010.
[3] R. Dewi, N. Sylvia, and M. Riza, "The Effect of Rice Husk and Saw Dusk Filler on Mechanical Property of Bio Composite from Sago Starch," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 3, 2021, doi: 10.52088/ijesty.v1i3.113.
[4] R. J. Podesva and N. Adisasmito-Smith, "Acoustic investigation of the vowel systems of Buginese and Toba Batak," in *Proceedings of the 14th International Congress of Phonetic Sciences*, 1999, pp. 535–538.
[5] H. Susilo, W. Wahyudi, and A. Hidayatno, "Analisis Ucapan Vokal Bahasa Indonesia dengan Alihragam PAket Gelombang Singkat." Jurusan Teknik Elektro Fakultas Teknik, 2011.
[6] D. Maisa Putra, O. Oktamianiza, M. Yuniar, and W. Fadhila, "Study Literature Review On Returning Medical Record Documents Using HOT-FIT Method," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 1, 2021, doi: 10.52088/ijesty.v1i1.102.
[7] L. P. Tuti Ariani, "The Effect Of Repetition Sprint Training Method Combined With The Level Of Physical Fitness Toward The Speed Of 100 Meter Run," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 3, 2021, doi: 10.52088/ijesty.v1i3.89.
[8] D. Kumar, "Power System Restoration Using Multilayer Perceptron," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 1, 2021, doi: 10.52088/ijesty.v1i1.35.
[9] W. Rahim, E. Erwin, and U. Malik, "Analisis Spektrum Suara Manusia Berdasarkan Jenis Kelamin (Gender) Dan Kelompok Umur Menggunakan Komputer." Riau University, 2014.
[10] Johan Sundberg, *The Psychology of Music*, 3rd ed. San Diego: Academic Press, 2013.
[11] M. Subali, M. Andriansyah, and C. Sinambela, "Analisis Frekuensi Dasar dan Frekuensi Formant dari Fonem Huruh Hijaiyah untuk Pengucapan Makhraj dengan Metode Dtw," *Pros. PESAT*, vol. 6, 2015.
[12] M. N. Al-Azhar, "Audio forensic: Theory and analysis," *Pus. Lab. Forensik Polri Bid. Fis. Dan Komput. Forensik*, pp. 3–6, 2011.
[13] A. Parlys, A. A. Zahra, and A. Hidayatno, "PENGGOLONGAN LAGU BERDASARKAN SPEKTOGRAM DENGAN CONVOLUTION NEURAL NETWORK," *Transient J. Ilm. Tek. Elektro*, vol. 7, no. 1, pp. 28–33, 2018.
[14] D. Maisa Putra and A. Alfauzain, "Design Of Tracer Using Microsoft Access Unit In Medical Record Primary Health In Padang," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 3, 2021, doi: 10.52088/ijesty.v1i3.90.
[15] E. Mulyani, Erwin, and Salomo, "Penentuan Frekuensi Fundamental dan Formant Suara Manusia Dewasa Berdasarkan Perbedaan Suku dan Gender menggunakan Software Praat," *Komun. Fis. Indones.*, vol. 12, no. 10, pp. 679–685, 2015, doi: http://dx.doi.org/10.31258/jkfi.12.10.679-685.
[16] Harmaini, *Psikologi Lintas Budaya*. Pekanbaru: Almujtahadah Press, 2011.
[17] J. Kuang and M. Liberman, "Pitch-Range Perception: The Dynamic Interaction Between Voice Quality and Fundamental Frequency.," in *InterSpeech*, 2016, pp. 1350–1354.