

THE RELATIONSHIP BETWEEN ELECTROMYOGRAPHY SIGNAL OF NECK MUSCLE AND HUMAN VOICE SIGNAL FOR CONTROLLING LOUDNESS OF ELECTROLARYNX

Fatchul Arifin^{*,†,¶}, Tri Arief Sardjono^{†,‡}
and Mauridhi Hery Purnomo^{†,§}

^{*}*Electronic Department*
Universitas Negeri Yogyakarta, Indonesia

[†]*Electrical Engineering Department*
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia

[‡]*sardjono@elect-eng.its.ac.id*

[§]*hery@ee.its.ac.id*

Accepted 24 October 2013

Published 9 May 2014

ABSTRACT

Human voice intonation is affected by pitch and loudness. Pitch is related to the frequency of human voice, while loudness is related to the magnitude of human voice. Someone who does not have vocal cords, has no ability to produce voice. This problem is suffered by laryngectomy patients. Over half of all laryngectomy patients worldwide use electrolarynx for the rehabilitation of their speech ability. Unfortunately, the electrolarynx voice produces monotonic and flat intonation. Small changes in pitch and loudness of electrolarynx will give a better expression in laryngectomy patients. However, previous researches have focused on utilization of electromyography (EMG) signal of neck muscle for only pitch control. In this research, the relationship between human voice intonation (i.e. frequency and magnitude) and EMG signals of neck muscles was studied by looking for their correlation and their mutual information. Human voice signal and EMG signal of neck muscle were recorded simultaneously while subjects were saying “A” with varying intonation. The EMG signal of neck muscle was processed using amplifying, filtering, rectifying and “moving average” process. On the other hand, the human voice was processed by FFT Algorithm to obtain magnitude and fundamental frequency. The result shows that the correlation coefficient between human voice magnitudes and EMG signal of neck muscle is 0.93, while the correlation coefficient between human voice frequency and EMG signal of neck muscle is 0.88. Moreover, the mutual information between human voice magnitudes and EMG signal of neck muscle is 1.07, while the mutual information between human voice frequency and EMG signal of neck muscle is 0.65. These results show that the relationship between human voice magnitudes and EMG signal of neck muscle is stronger than the relationship between human voice frequencies and EMG signal of neck muscle. Therefore, it is more appropriate to use the EMG signal of neck muscle for controlling loudness of electrolarynx than that of the pitch of electrolarynx.

Keywords: Neck muscle electromyography signal; Loudness of electrolarynx voice; Laryngectomies patients; Correlation; Mutual information.

[¶]Corresponding author: Fatchul Arifin, Electronic Department, Universitas Negeri Yogyakarta, Indonesia, E-mail: fatchul@uny.ac.id; fatchul.ar@gmail.com

INTRODUCTION

Human voice is the sound produced by humans using the lungs and the vocal cord in the larynx. Voice is produced by pumping the air from the lungs, through the vocal cords. When air is pushed through the vocal cords with sufficient pressure, the vocal cords vibrates. Variation in pitch and loudness create intonation of human voice.^{1,2} Pitch corresponds to the frequency of voice. The vibration speed of vocal cord influence the pitch. On the other hand, loudness determines how loud the voice is produced. Loudness corresponds to the magnitude of voice. The tension in the vocal muscle and the amount of air pressure from the lungs influence the loudness. Therefore, the higher the tension of the vocal cord and the greater the pressure of air from lungs, the louder the voice is produced.

However, not all of voices are meaningful and useful to communicate with others. Infants babble, humans laugh, cry, whinny and growl are some examples of voice with little meaning. Some voices are articulated by oral and nasal cavity for speech. Speech is used to express one's feelings and desires to others.

Someone who does not have vocal cords, has no ability to produce voice and speech. This problem is suffered by late-stage laryngeal cancer patients. They are usually treated with total laryngectomy, in which larynx, and tissues around it, including vocal cord, should be removed. By doing surgery, a hole in front of the patient's neck, known as stoma, is made. Then, the trachea is attached to this stoma which is used by the patients to breathe. As the vocal cord of the laryngectomies patients have been removed, they will not be able to speak anymore. They have lost their ability to speak as they did.³

To help laryngectomy patients to be able to speak again, an electronic device called electrolarynx is used. Over half of laryngectomy patients worldwide have used electrolarynx because it can be used easily and no special requirements are required.^{3,4} The basic operation principle of electrolarynx is to generate vibration which is used by placing it at low chin, so that the air inside the mouth vibrates. Then, oral and nasal cavities modify the vibration signal into speech by acting as a resonator for certain frequencies (articulation).⁵

To date, there are three key issues in electrolarynx study; electrolarynx hardware, user friendliness (patient

must use one hand for reaching it from his pocket, applying it to the chin, and pressing the button for on-off electrolarynx) and monotonic-like robot voice. There are number of researches on electrolarynx related to these issues. Riesz patented an activator hardware of electrolarynx. He placed tone generator at low chin.^a Lukocovic patented a model of diaphragm as a sounding head of electrolarynx.^b He also patented solenoid as electrolarynx activator.^c Lennox designed hands-free electrolarynx. First, tone generator was placed at the glasses. Next, there was a tube that goes into the mouth to supply the vibration. To turn the system on or off, the rotation of the arm was used.^d Barney and Burtschi controlled the intonation of electrolarynx by varying resistor (R) and capacitor (C).^{e,f} The setting was performed manually through a button. Uemi *et al.* controlled the intonation of electrolarynx using breath pressure,⁶ while Mitsuo *et al.* developed electrolarynx with microcomputer with some installed melodies to sing a song.⁷ Hoping that these melodies would give improvements in the intonation of electrolarynx.

On the other hand, intonation can also be controlled by using electromyography (EMG) signal of neck muscle. The EMG signal is a signal that shows the electrical activity of muscle. The value of EMG signal is related to the activity of the muscle.⁸ The more muscle activity, the greater the EMG signal produced. EMG signal is detected via a pair of electrodes that are placed at neck strap muscle. Heaton and Stepp utilized EMG signal of neck muscle to control the pitch of electrolarynx.^{9,10} Seven electrodes were used in order to detect EMG signal of neck muscle. Pitch controlling is performed via voltage controlled oscillator (VCO).

Variation in pitch and loudness creates intonation of human voice. Therefore, the intonation of electrolarynx is not only determined by the pitch control but also by the loudness control. This research evaluates the relationship between the two elements that affects intonation of human voice (i.e. frequency and magnitude) and EMG signal of neck muscle. This work is very important because it will reveal which one has a stronger relationship to the EMG signal of neck muscle. Therefore, this research offers the possibility to improve the intonation control of electrolarynx.

^aUS Patent 2041487.

^bUS Patent 20030031325 A1.

^cUS Patent 20030031324 A1.

^dUS Patent 4292472 A.

^eUS Patent 3072745.

^fUS Patent 4039756 A.

MATERIALS AND METHODS

For that purpose, 78 human voice productions were recorded from three normal subjects of different ages. These three normal subjects were males, who were 21, 24 and 41 years old. Each subject was asked to say “A” 26 times repeatedly with a varied intonation. The reason behind this recording design was to focus on the observation of the frequency and magnitude of the human voice from certain productions. If the voice was produced in more various units (such as reading words or sentences), the observation would be biased and unfocused.

The human voices and EMG signals were recorded simultaneously to the subjects. The recording process was conducted using voice recording and EMG instruments unit. The recorded human voice then was processed using Fast Fourier Transform (FFT) algorithm to obtain its frequency and magnitude. FFT was used with 512 data points and Hanning method for the window.

A pair of general-purpose surface-electrodes (positive and negative) were placed at neck muscle (sternocleidomastoid muscle), while a ground surface electrode was placed at the sternum. The shape of the electrodes was circular with a diameter of 2 cm. The raw EMG signals typically have amplitudes in the range of mV, thus a powerful medical instrumentation with thousands of gain was needed. The presence of EMG signal could be found within the frequency of 20–500 Hz, but dominant at 40–200 Hz. Some filters were required for eliminating higher and lower frequency and also power line interference noises. Band pass filter (BPF) was used to eliminate EMG signal with frequency lower than 40 Hz and higher than 200 Hz. Finite Impulse Response (FIR) filter was used with its parameters; frequency stop1 = 38 Hz, frequency pass1 = 42 Hz, frequency pass2 = 198 Hz and frequency stop2 = 202 Hz. On the other hand, notch filter 50 Hz was also used to eliminate power line interference noise.

In the next process, the EMG signal was rectified using full wave rectification to translate the raw EMG signals into single polarity (positive). The rectified EMG signal of the neck muscles was then extracted to find the envelope signal. One of the various ways in obtaining the envelope signal is by applying the moving average process. Suppose that the sequence of original data $y_1, y_2, y_3, y_4, \dots, y_k$ were initially obtained. The moving average process was calculated using Eq. (1).

$$X_i = \frac{\sum_{k=i}^{(i+n-1)} y_k}{n}, \quad (1)$$

X_i is moving average element, while y_k and n are the original data and the number of members in the subgroup, respectively. The first element of the moving average

process was obtained by forming subgroups consisting of n initial data, and its average was later obtained. The average of this subgroup was the initial value of the moving average X_1 . The next element of the moving average was obtained by shifting forward the subgroups in which the first value of initial data of previous subgroup was left behind and the next value of initial data was taken to be the members of new subgroup. The average of this new subgroup was used to be the second value of the moving average, X_2 .¹¹ This process was then repeated over the entire data. The value of n affected the softness of the produced envelope signal. The greater the value of n , the more coarse signal envelope produced. In this study, the value $n = 10$ was selected.

As mentioned earlier, the purpose of this research is to study which relationship to the EMG signal of neck muscle is stronger, the magnitude of human voice or the frequency of human voice. To get an answer, the existing data must be processed by some tools. Two kinds of tools were used in this study, namely, correlation and mutual information which were used separately.

Correlation is a way of measuring the linear relationship between two random variables. The correlation coefficient r between two variables can be calculated using Eq. (2)¹²:

$$r = \frac{n \cdot \sum(x \cdot y) - \sum x \cdot \sum y}{\sqrt{(n \cdot (\sum x^2) - (\sum x)^2)} \cdot \sqrt{(n \cdot (\sum y^2) - (\sum y)^2)}}, \quad (2)$$

x and y are two variables that are correlated, while n is the number of data. The value of the correlation coefficient shows the strength of linear relationship between two variables. Two variables will have a correlation if one variable influences the other variables, either in the same direction or the opposite directions. Greater value of r indicates a stronger correlation. The ranges of r is 0 to 1. In this study, the calculation of r is performed using Pearson’s method.

Correlation is a valid tool if the measured variables are considered linear. However, if a variable linearity is absorbed, another tool is required to calculate the non-linear dependence. In this present study, the nonlinear dependence is measured by mutual information. Mutual information is a quantitative measurement of how much a random variable (X) gives information about another random variable (Y).¹³ Mutual Information is often used to rank sources of information based on the assumption that the uncertainty of each variable is represented by a probability distribution in which the function can be represented by entropy. Entropy is a statistical measure that summarizes randomness.¹⁴ Figure 1 shows the illustration of mutual information of two variables. The

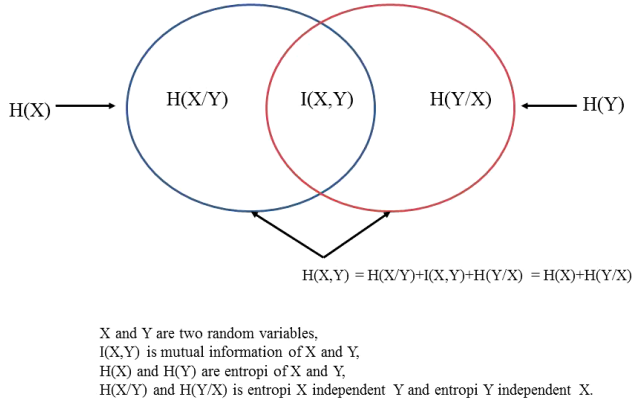


Fig. 1 Mutual information of two variables.

calculation of mutual information is calculated by Eqs. (3)–(7).

$$I(X, Y) = H(X) - H(X/Y) = H(Y) - H(Y/X), \quad (3)$$

$$H(X, Y) = H(X) + H(Y/X) = H(Y) + H(X/Y), \quad (4)$$

$$H(X) = \sum P(X) \cdot \log_2 P(X), \quad (5)$$

$$H(Y) = \sum P(Y) \cdot \log_2 P(Y), \quad (6)$$

$$H(X, Y) = \sum \sum P(X, Y) \cdot \log_2 P(X, Y), \quad (7)$$

$I(X, Y)$ is mutual information of X and Y , $H(X)$ and $H(Y)$ are entropy of X and Y , $H(X/Y)$ and $H(Y/X)$ are entropy of X independent of Y and entropy of Y independent of X . $P(X)$ and $P(Y)$ are the probability of X and Y , while $P(X, Y)$ is the joint probability of X and Y . The mutual information measures the information which is shared by Y and X . If Y and X are independent i.e. Y does not give any information about X , so that their mutual information is zero. In extreme case, if Y and X are identical then all of information about Y is shared with X . Thus, in case of identity, the mutual information is the same as the uncertainty contained in Y (or X) alone namely the entropy of Y (or X). Thus, the mutual information value range is 0 to entropy of Y (or X). In this research, mutual information was calculated by R Software with Minet Library.¹⁵

RESULTS

When the subjects said “A”, EMG signal of neck muscle can be seen in Fig. 2. Figure 2(A) is the raw data of EMG signal of neck muscle. X axes is time in second (s), while Y axes is magnitude of EMG signal of neck muscle in millivolt (mV). Figure 2(B) is the rectified EMG signal of neck muscle. The overall data of this figure are over the baseline. Figure 2(C) is the output of moving average process EMG signal of neck muscle. Peak value of

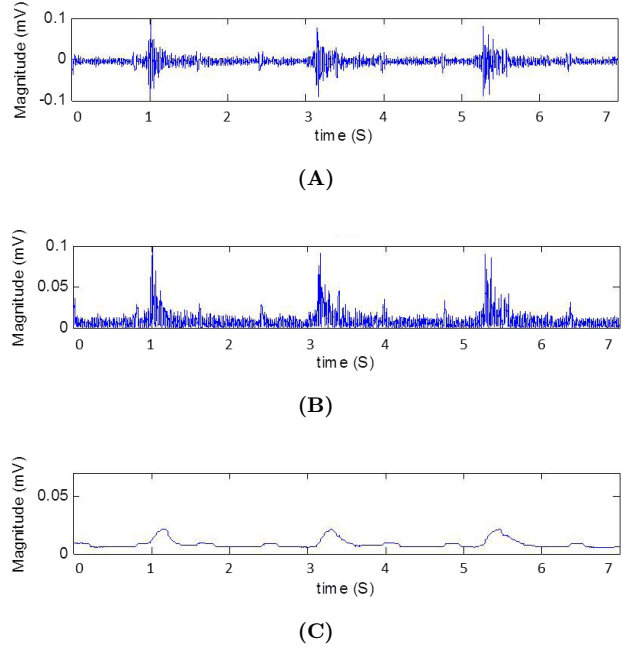


Fig. 2 The raw, rectified and moving average process output of EMG signal of neck muscle.

envelope signal is lower than peak value of rectified EMG signal. It is due to the effect of moving average process.

On the other hand, human voice is processed by FFT method. From the FFT process, frequency spectrum of human voice i.e. frequency and magnitude is obtained. Some of frequency spectrum are displayed in Table 1,

Table 1. Some of the Output Data of FFT Process (Frequency and Magnitude) of Human Voice and the Output Data of Moving Average Process of EMG Signal of Neck Muscle.

| Volunteer | Repetition | Spectrum Frequency of Human Voice | | Output of MA of EMG Signal (mV) |
|-----------|------------|-----------------------------------|----------|---------------------------------|
| | | Freq (Hz) | Mag (db) | |
| Subject 1 | 1 | 115 | -63 | 0.011 |
| | 2 | 143 | -64 | 0.009 |
| | 3 | 126 | -66 | 0.009 |
| | 4 | 114 | -62 | 0.010 |
| | 5 | 103 | -64 | 0.011 |
| Subject 2 | 26 | 116 | -62 | 0.010 |
| | 1 | 166 | -52 | 0.014 |
| | 2 | 167 | -49 | 0.016 |
| | 3 | 170 | -55 | 0.015 |
| | 4 | 167 | -50 | 0.019 |
| Subject 3 | 26 | 166 | -52 | 0.017 |
| | 1 | 117 | -57 | 0.015 |
| | 2 | 115 | -59 | 0.011 |
| | 3 | 164 | -53 | 0.015 |
| | 4 | 153 | -58 | 0.011 |
| | 5 | 130 | -59 | 0.012 |
| | 26 | 154 | -55 | 0.015 |

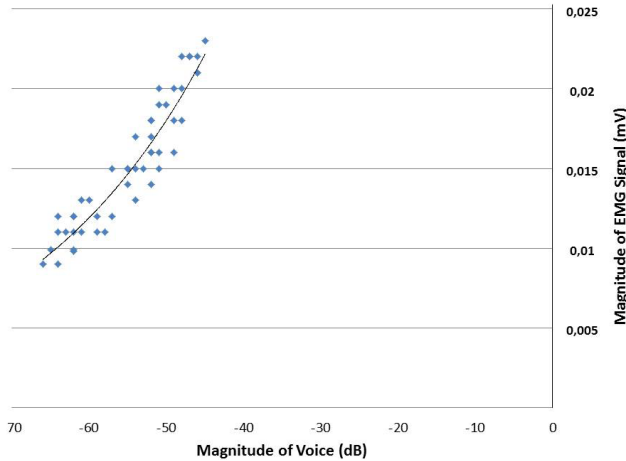


Fig. 3 Plot of magnitude of human voice and EMG signal of neck muscle.

while the peak value of EMG envelope signal which is recorded simultaneously to human voice is also displayed.

Using Eq. (2), r between human voice spectrum (frequency and magnitude) and EMG signal of neck muscle can be calculated. The result shows that the r value between the magnitude of human voice and EMG signal of neck muscle was 0.93, while r value between the human voice frequency and EMG signal of neck muscle was 0.88. Figure 3 shows the plot of magnitude of human voice and EMG signal of neck muscle. X axis is magnitude of human voice in dB, while Y axis is magnitude of EMG signal of neck muscle in mV. Figure 3 shows that the data have a good linearity trend with uniform data spread. Figure 4 shows the plot of human voice frequency and EMG signal of neck muscle. X axis is frequency of human voice in dB, while Y axis is magnitude of EMG signal of neck muscle in mV. It is due to the

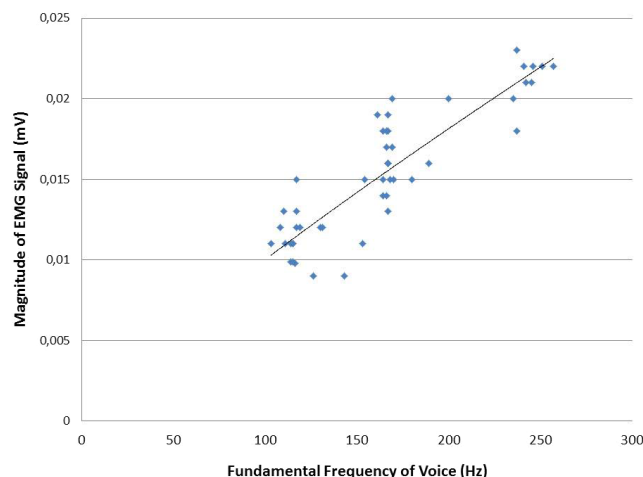


Fig. 4 Plot of human voice frequency and EMG signal of neck muscle.

Table 2. Result of Mutual Information between Human Voice (Frequency and Magnitude) and EMG Signal of Neck Muscle.

| Variable 1 | Variable 2 | Result of MI |
|--------------------|------------|--------------|
| Magnitude of Voice | EMG Signal | 1.074 |
| Frequency of Voice | EMG Signal | 0.650 |

data accumulation in some places. Therefore, trend linearity of data is lower than in Fig. 3.

The relationship between the human voice and EMG signal of neck muscle was also calculated by mutual information. The result of mutual information was obtained by applying R software with MINET library which is displayed in Table 2. The mutual information between the magnitude of human voice and the EMG signal of neck muscle was 1.07. This result is higher than the mutual information between the frequency of the human voice and EMG signals of neck muscle, which is only 0.65.

DISCUSSION

Goldstein classified human voice by grouping “High, Normal, Soft, Loud and Low”. He related human voice frequency (pitch) to EMG signals of neck muscle.³ This classification is less accurate, however. In our study, the magnitude and frequency of human voice are classified accurately by measured value. The results show that both of coefficient correlation, r , are high. It means there are strong relationship between human voice magnitude with EMG signal of neck muscle, and strong relationship between human voice frequency with EMG signal of neck muscle. However, r of human voice magnitude and EMG signal is higher than r of human voice frequency and EMG signal. Moreover, the mutual information of human voice magnitude and EMG signals of neck muscle are also higher than the mutual information of human voice frequency and EMG signals of neck muscle. Figures 3 and 4, and Table 2 provide the evidence on the issue. This indicates that the magnitude of the human voice gives more information on EMG signals of neck muscle compared with the frequency of human voice. As a result, the EMG signals of neck muscle is more appropriate for controlling the loudness of electrolarynx rather than for controlling the pitch of electrolarynx, and the electrolarynx speech will be more natural.

However, to the extent of our knowledge, no previous research has employed the EMG signal for controlling the loudness. A proposed model of controlling the loudness of electrolarynx is shown in Fig. 5. A laryngectomy patient speaks using electrolarynx device. A microphone converts

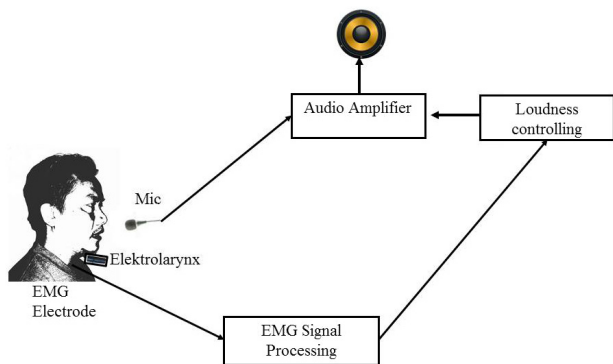


Fig. 5 The proposed model of loudness control of electrolarynx speech.

electrolarynx speech to be electrical signal which is then amplified or attenuated by an amplifier. The magnitude of amplification changes dynamically depending upon the preprocessed EMG signals of neck muscle.

EMG signals of neck muscle has been used by several researchers to improve the performance of electrolarynx, but none of them use EMG signals of neck muscle to control the dynamic loudness of electrolarynx. Goldstein *et al.* used the EMG signals of neck muscle to control the pitch (fundamental frequency) of electrolarynx.³ Pitch setting was done by controlling the frequency of vibration of electrolarynx. Before using, the volunteers were trained how to use it. The results are reported that 60–70% of spoken sentence have correct intonation. Goldstein *et al.* did not mention anything about the dynamic loudness control of electrolarynx, whereas in our study the relationship between loudness with EMG signal of neck muscle is stronger than the relationship between pitch with EMG signals of neck muscle.

While the loudness is controlled by EMG signals, the electrolarynx pitch can be controlled by other alternative mechanisms, one of which is the use of pressure of respiration at stoma, as demonstrated by Uemi *et al.*⁶ Another option is by combining the control of pitch (which is developed by Goldstein) as well as the loudness (which is proposed in this study) simultaneously using EMG signal. One side of EMG signal of neck muscle is applied for pitch control, while another side of EMG signal is utilized for amplification scheme, which results in dynamic loudness variation of electrolarynx. Therefore, by enabling the control of both pitch and loudness, it is expected that the electrolarynx produces better intonation of artificial speech.

CONCLUSION

The intonation of electrolarynx is not only determined by the pitch control but also by the loudness control.

Previous researches have focused on the utilization of EMG signal of neck muscle for pitch control only. In this research, the relationship between EMG signals of neck muscles and human voice intonation (frequency and magnitude) was studied from the correlation coefficient r and the mutual information. The results show that the relationship between the human voice magnitude and the EMG signal of neck muscle (correlation coefficient r and mutual information are 0.93 and 1.07, respectively) is stronger than the relationship between the human voice frequency and the EMG signal of neck muscle (correlation coefficient r and mutual information are 0.88 and 0.65 respectively). Therefore, the EMG signals of neck muscle is more appropriate for controlling the loudness of electrolarynx than for controlling the pitch of electrolarynx. This opens the possibility to improve the control of intonation, not only by pitch control, but also by manipulating the dynamic loudness of electrolarynx.

ACKNOWLEDGMENTS

We would like to express our gratitude to Ministry of National Education and Culture of Republic, Indonesia which has given the financial support to this research. We also would like to thank to Medical Ethical Committee of Moewardi Hospital Surakarta who had approved this research.

REFERENCES

- Halliday M, *A Course in Spoken English: Intonation*, Oxford University Press, Oxford, 1970.
- Daley M, Vocal performance and speech intonation: Bob Dylan's 'like a rolling stone', *Oral Tradition* **22**:84, 2007.
- Goldstein EA, Heaton JT, Kobler JB, Stanley GB, Hillman RE, Design and implementation of a hands-free electrolarynx device controlled by neck strap muscle electromyographic activity, *IEEE Trans Biomed Eng* **51**:325, 2004.
- Kubert HL, Stepp CE, Zeitels SM, Gooley JE, Walsh MJ, Prakash SR, Hillman RE, Heaton JT, Electromyographic control of a hands-free electrolarynx using neck strap muscles, *J Commun Disorders* **42**:211, 2009.
- Sardjono TA, Hidayati R, Purnami N, Noortjahja A, Verkerke, Mauridhy HP, A preliminary result of voice spectrum analysis from laryngectomised patients with and without electro larynx: A case study in Indonesian laryngectomised patients, *Int Conf Instrum, Commun, Inf Technol, and Biomed Eng (ICICI-BME)*, Jakarta, pp. 1–5, 2009.
- Uemi N, Ifukube T, Takahashi M, Matsushima J, Design of a new electrolarynx having a pitch control function,

- Proc Robot Hum Commun IEEE Int Workshop*, Nagoya, pp. 198–203, 1994.
7. Mitsuo H, Uemi N, Oikawa MY, Yoshinori , Sugai Y, Ifukube T, Industrialization of the electrolarynx with a pitch control function and its evaluation, *IEICE Inst Electron Inf Commun Eng* **6**:1240, 2001.
 8. Betts BJ, Binsted K, Jorgensen C, Small-vocabulary speech recognition using surface electromyography, *Interact Compt J* **18**:1242, 2006.
 9. Heaton JT, Robertson M, Griffin C, Development of a wireless electromyographically controlled electrolarynx voice prosthesis, *Eng Med Biol Soc EMBC Ann Int Conf IEEE*, pp. 5352–5355, 2011.
 10. Stepp CE, Heaton JT, Rolland RG, Hillman RE, Neck and face surface electromyography for prosthetic voice control after total laryngectomy, *IEEE Trans Neural Syst Rehabil Eng* **17**:146, 2009.
 11. Blake L, The stability of moving average technical trading rules on the Dow Jones index, *Trading Regul* **5**:12, 1999.
 12. Egghe LL, The relation between Pearson's correlation coefficient r and Salton's cosine measure, *J Am Soc Inform Sci Technol* **60**:102, 2009.
 13. Lilian N, Daniel W, Daniel MC, Jeremy G, *Correlation and Mutual Information, Chemical Engineering Process Dynamics and Controls Open Text Book*, the University of Michigan, 2007.
 14. Subianto M, Measuring the economic potential of fruit plant using mutual information, *J Nat* **10**:25, 2010.
 15. Meyer PE, Lafitte F, Bontempi G, Minet: Mutual information network inference, R package version 2, 2009.