The Role of Tone Frequency Characters for Identity Recognition

Ronald Maraden Parlindungan Silalahi

Universitas Bunda Mulia

Abstract

Speech sounds contain thousands of acoustical features. The linguistic acoustical features can be importantly used to recognize the identity of the speakers due to the fact that every person produces different sounds. This paper is descriptively aimed to investigate the role of tone frequency characters to recognize a speaker's identity. In order to achieve the objective of this research, the researcher pictures the character of tone frequencies by using PRAAT 4.5.1.4. Supported by mathematical measurement, the writer will formulize the role of these features for identification of personal identity and describe the logical implication of this formulation.

Keywords:

Tone, frequency, acoustic, identity, recognition

1. Introduction

Every person has some characteristics that specifically build his/her personal identity. One of the most unique features of language used as an identity marker is language. Language itself contains some different features that are not the same as another or each other. These linguistic differences reflect the identity of the speaker. This phenomenon is the basic reason why a speaker can recognize the identity of other speakers without having a conversation directly.

Due to the advancement of technology, the linguistic features of speech sounds are specifically used to identify and recognize personal identity. The acoustical features of language such as pitch, intonation, stress, or juncture pattern are computanionally applied to identify the unique characteristics of speech sounds. Every speaker specifically produces different levels of pitch, intonation, or stress when using language. Nowadays, these differences can be computationally detected, pictured, or analyzed further.

2. Purposes

This paper is aimed to investigate the role of tone frequency characters to recognize speaker's identity. There are two supporting purposes that base this research:

1. Finding the characteristic of tone frequency.
2. Identifying and comparing the characteristic of tone frequency.

3. Theoretical Review

This section contains the previous studies and theories that are fundamentally related with this research.

3.1 Previous Studies

a. Demenko (2000)

Grazyna Demenko (2000) conducted a research on the role of supreasegmental features to verify the identity of speakers. Her research is specifically aimed to evaluate or estimate the acoustical features recorded in the text-independent verification system. The result indicates the significant individual differentiation in suprasegmental features.

b. DyVis (2000)

Francis Nolan, Kirsty McDougall, Gea de Jong, and Toby Hudson conducted a research on the role of linguistic signals to characterize speech sounds. In a scientific project named Dynamic Variability in Speech (DyVis), Nolan *et al* (2006) investigated the characteristics of human speech sounds diachronically by comparing two speech sounds produced by the same source but different time. The quality of acoustical frequency is highly influenced by the shape of the articulator.

c. Sarwono (2008)

Identity recognition can be used to identify personal identity in forensic linguistics. It is conducted by comparing two or more speech sounds. Based on

this comparison, we can identify whether the speech sounds originaly come from the same source.

d. Silalahi (2011)

Silalahi (2011) investigates the role of formant frequencies for the identification of personal identity. In his research, Silalahi (2011) finds the tolerant limit that can be used as the identity marker.

3.2 Theoretical Review

3.2.1 Tone Frequencies

Basically, speech sounds are produced by relying on the air mechanism (Silalahi: 2011). The lungs produce the flow of air that moves through the larynx and vocal tract and thus create vibrations. These vibrations are transmitted through thousands of sound waves. They are investigated and pictured by using periodic cycle.

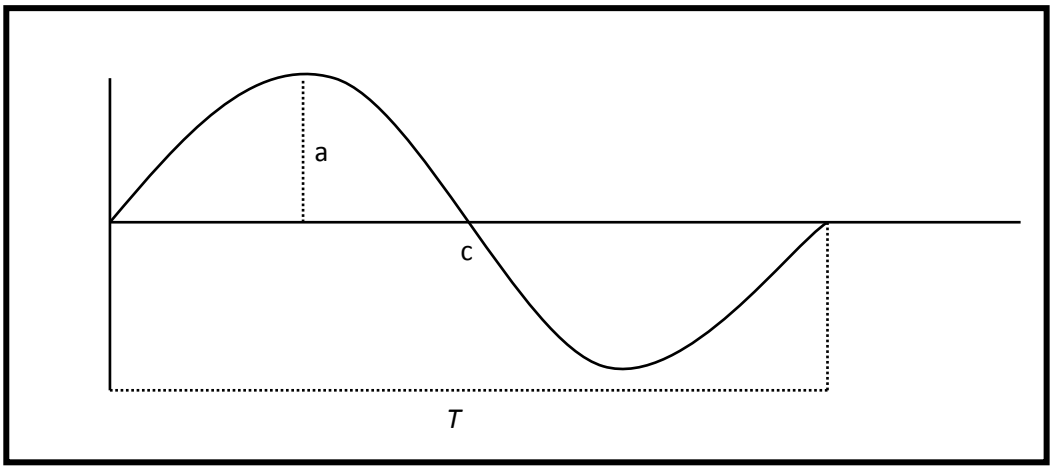

Figure 1. Speech Sounds
(Silalahi: 2011)

Frequency (tone frequency) can be defined as the number of cycles produced per second (*T*) (Kridalaksana, 2001: 60). It is symbolized with *f* and the unit of frequency is Hertz (Hz) (Silalahi: 2011).

3.2.2. The Production of Speech Sounds

The production of speech sounds are the result of articulator contraction such as pharynx, velum, hard palate, alveolar ridge, *tongue*, teeth and lips (Roach: 1983). Basically, speech sounds can be classified into three types (vowel, consonant, and semi vowel). In this paper, the writter focuses on vowel because the production of vowel is totally influenced by the tone frequencies produced.

The vowel can be classified in terms of several aspects such as tongue height, tongue backness, lip rounding, and the tenseness of the articulators (Silalahi: 2011).

No	*Dimension*	*Types*	*Sounds*	*Production*
1	Tongue Height	High	[i], [I], [u], and [ʊ]	Narrow space between the tongue and the hard palate.
		Mid	[e], [ɛ], [o], and [ɔ]	Tongue is positioned between the high and low vowels.
		Low	[æ] and [a]	Wide space between the tongue and the hard palate
2	Tongue backness	Front	[i], [I], [e], [ɛ], and [æ]	Forward in the mouth
		Central	[a] and [ə	tongue between the front and back vowels.
		Back	[u], [ʊ], [o], and [ɔ])	Far back in the mouth
3	Lip rounding	Rounded	[u] and [o]	Lip rounding
		Unrounded	[i] and [ɛ]	Without lip rounding
4	Tenseness	Tense	[i] and [e]	high degree of tenseness.
		Tenseless	[i] and [e]	Low degree of tenseness

Table 1. The Production of Speech Sounds

This classification of speech sound production is clearly pictured from the following figure.

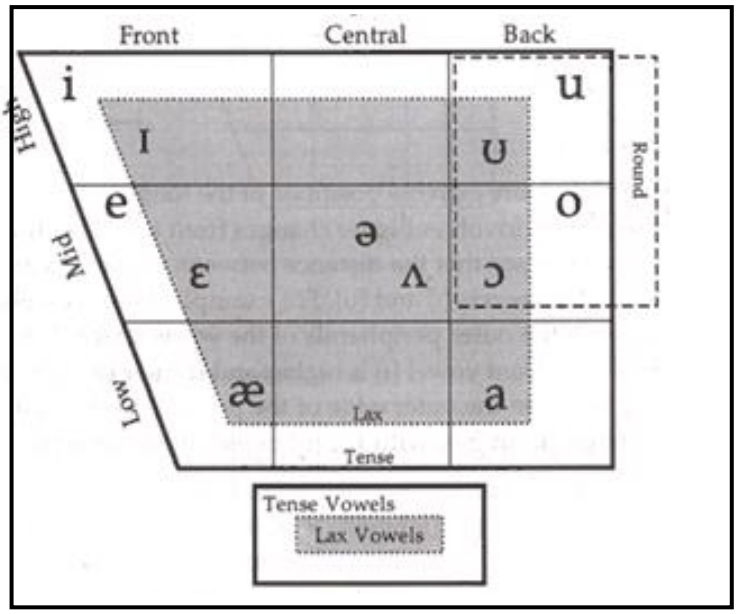

Figure 2. The Production of Speech Sounds

(http://www.ic.arizona.edu/~lsp/IPA/SSAE.html)

3.2.3 Identity Recognition

A speaker's identity can be recognized without having direct communication (face to face). Speakers' capability in identifying identity is based on two processes (1) identification and (2) verification (Silalahi: 2011). Identification relates to the prior knowledge of the speaker about the specific acoustical features of the source sounds. Verification relates to the ability to recall the specific characters of the acoustical features obtained in the identification phase and compare it with the newly acquired sound.

Silalahi (2011) emphasizes that similarities and differences of the acoustical features of this benchmarking process can be used to determine the identity of the speakers. The process of identification and verification in speech recognition is pictured in the following figure.

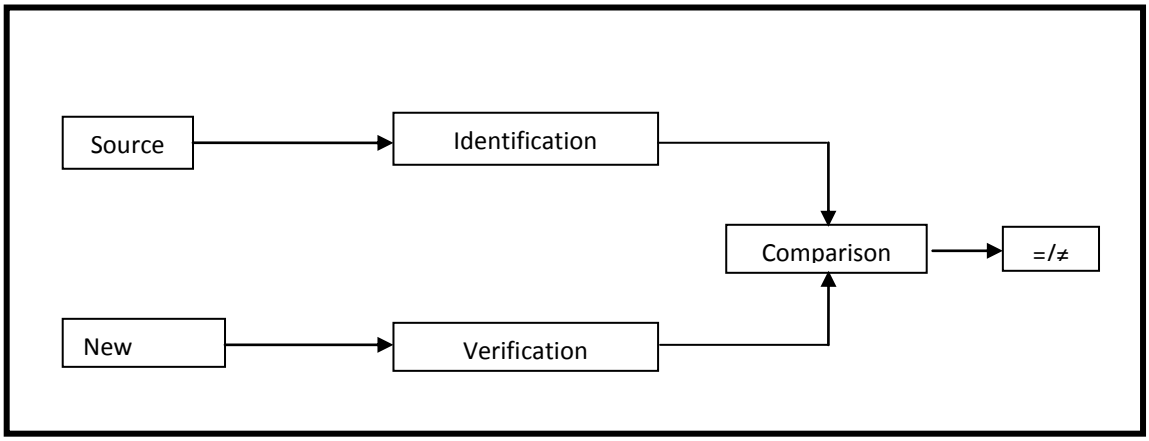

Figure 3. Identity Recognition

The differences in identification and verification process can be influenced by three factors: (1) Different articulators or physical form of articulator, (2) linguistic differences (such as accents, dialects, or registers), (3) and paralinguistic variation (for example, the speaker's emotional state).

4. Method of Research

This article is based on a quantitative research that is descriptively aimed to investigate the role of tone frequency characters to recognize a speaker's identity. The data are collected by providing stimulus to 20 respondents and obtaining the responses from them. There are ten male respondents and ten female respondents.

The respondents were asked to utter 8 words in English that contain eight primary cardinal vowels ([*i*], [*e*], [*ɛ*], [*a*], [*u*], [*o*], [*ɔ*], and [*a*]).

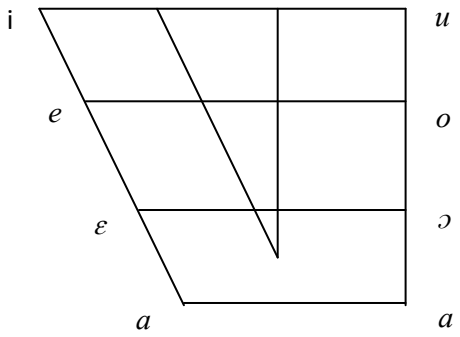

Figure 4. Primary Cardinal Vowels

The vowel sounds are strongly influenced by frequencies. The form of articulators will influence the frequencies produced (Wood, 2006).

Each word contains primary cardinal vowels and must be in the same phonetical environment. For example, vowel [a] and [i] in the words *mud* and *mid* appear between two similar consonants [m] and [d]. Each word will be uttered three times. Therefore, the writer collects 480 words to be analysed in this research. These data are recorded and analyzed by using PRAAT 4.5.14 created by Boersma and Weenink. There are 3 (three) stages processed by using PRAAT. It is clearly illustrated in the following figure.

1. Sound Extraction in WAV
2. Words Segmentation
3. Transferring the data. There are three types of data collected.
    a. The highest and lowest tone
    b. The mean of frequencies
    c. The range of frequency

5. Analysis

After Sound Extraction in WAV and Words Segmentation, each data is measured and tabulated. There are three types of data collected: (a) the highest and lowest tone, (b) the mean of frequencies, and (c) the range of frequency.

5.1 Findings

There are some significant differences between male and female respondents. The statistic equation shows that the highest frequency is produced by the female respondents ranging from 148.25 Hz (*minimum frequency*) to 289.81 Hz (*maximum frequency*). Male respondents produced the lowest frequency ranging from 94.90 (*minimum frequency*) to 165.59 (*maximum frequency*).

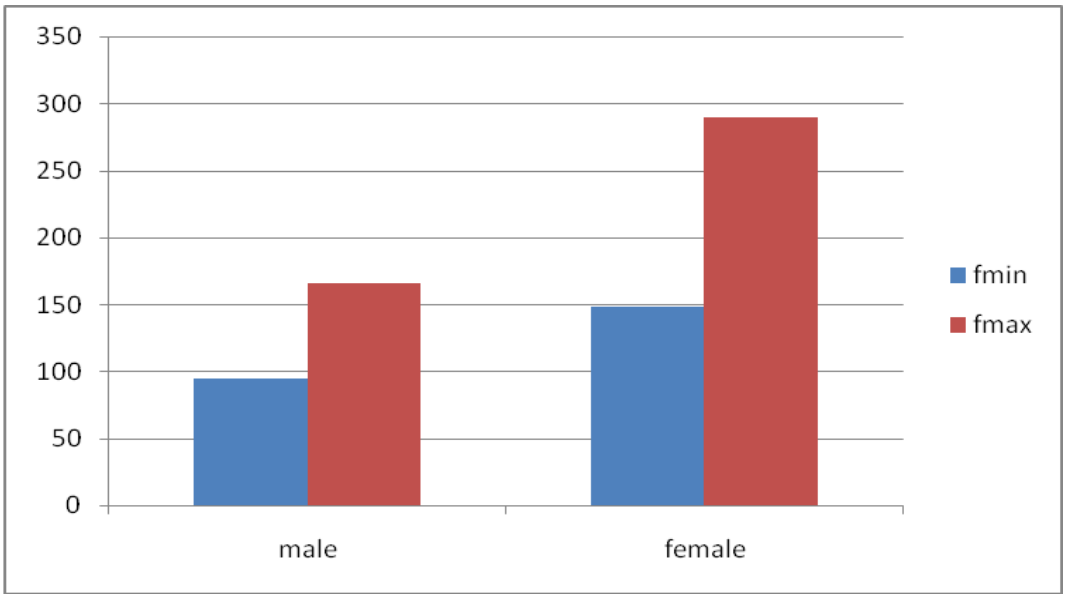

Figure 5. Maximum and Minimum Frequency

The statistic equation shows a significant difference of the range of frequency between female and male respondents. The range of frequency of female respondents is 141.56 Hz and the range of frequency of male respondents is 70, 69 Hz.

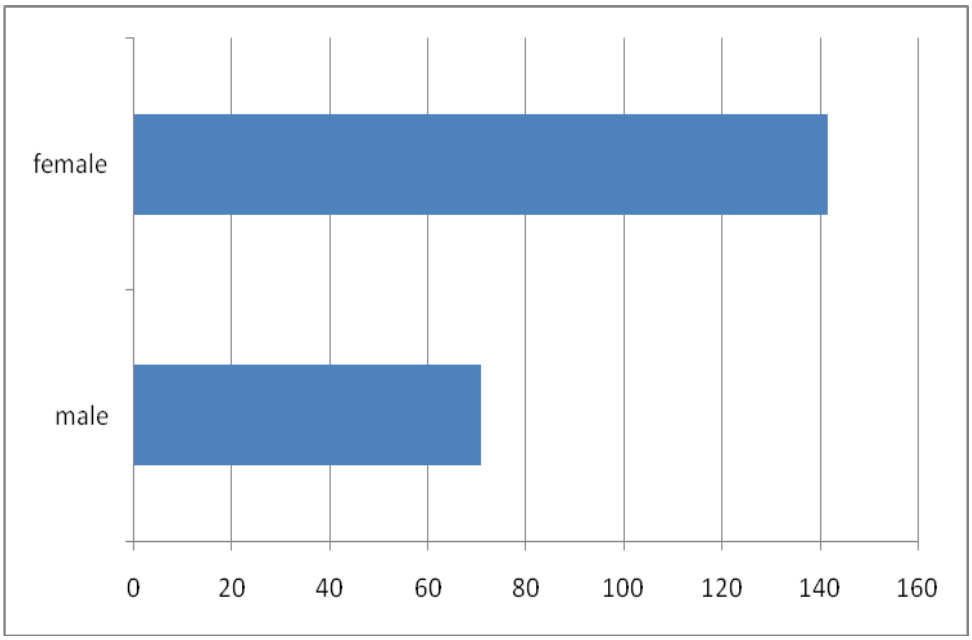

Figure 6. The Range of Frequencies

Statistic equation shows that every Respondent produces different frequency. It can be seen from the following 10 samples of female respondents compared.

No	respondents	The Range of Frequencies
1	R1	229,2249
2	R2	210,2341
3	R3	190,2167
4	R4	180,4163
5	R5	199,6463
6	R6	220,1543
7	R7	201,4512
8	R8	244,4944
9	R9	204,6658
10	R10	219,2098

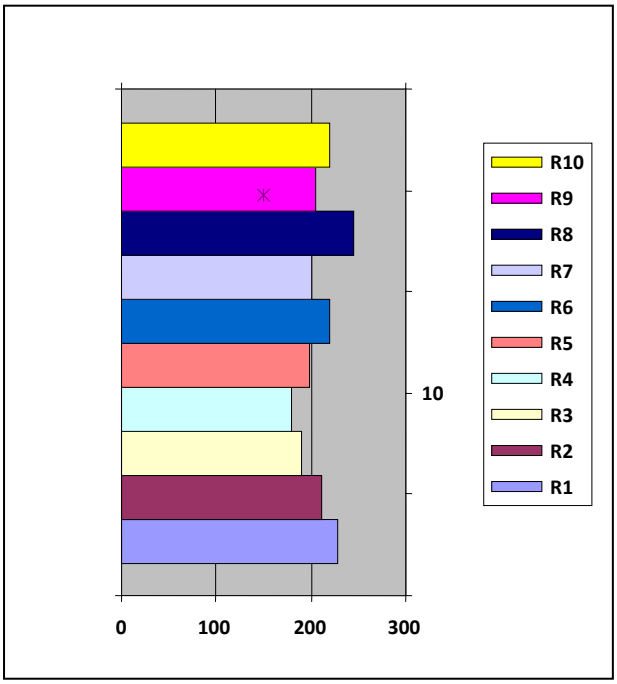

Figure 7. The Level of Frequencies (Female)

Statistic equation shows that the range of frequencies produced by respondent 8 is higher than the others.

From the 10 (ten) male respondents, we found that the average of frequencies produced are different from one another. The average of frequencies produced by respondent 3 (male) is higher than the others.

No	respondents	The Range of Frequencies
1	R1	116,0139
2	R2	119,1298
3	R3	153,0739
4	R4	137,0045
5	R5	134,4408
6	R6	128,8907
7	R7	120,3245
8	R8	127,3476
9	R9	132,56789
10	R10	134,5114

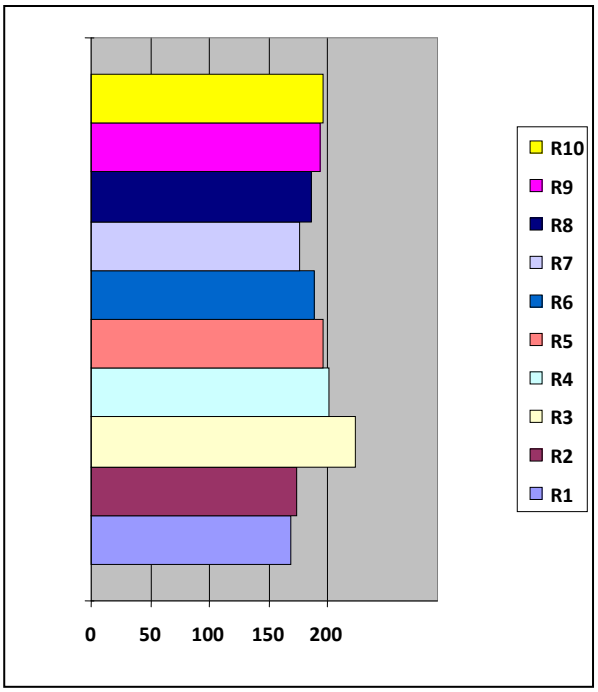

Figure 8. The Level of Frequencies (Male)

The standard deviation of each sound is different from one another. The level of standard deviation between male and female respondents are different. The highest standard deviation is 26.05421 Hz for female respondents and 22.7346 Hz for male respondents.

6. Conclusion

The speech sounds produced is different from one another. The ranges of frequencies produced by female speakers are higher than the male speakers. This clearly portrays that there are some significant differences between male and female speakers. Therefore, the differences between the genders is clearly reflected from the sound produced.

In this reasearch, standard deviation is used to calculate how the frequencies produced by speakers are spread. The logical implication of this statement is the farther the frequency from the center point, the more different the source. The highest standard deviation is 26.05421 Hz (for female respondents) and 22.7346 Hz (for male respondents).

Similarities and differences of the acoustical features can be used to determine the identity of the speakers. We can conclude that two similar sounds are originaly produced by the same speaker if those sounds have some similarities in identification and verification. Therefore, the writer develops three parameters in recognizing personal identity: (1) the highest and lowest tone, (2) the mean of frequencies, (3) the range of frequency.

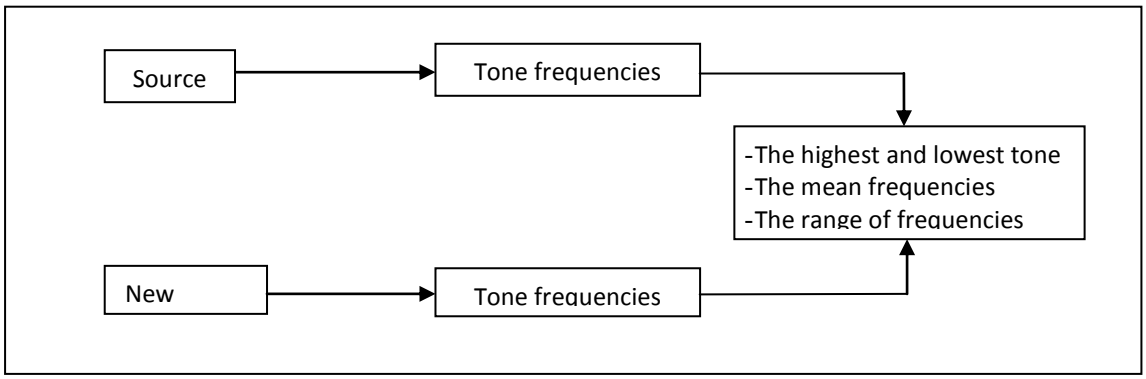

Figure 9. Implication for Identity Recognition

In figure 9, two sounds (source and new sound) are compared. The value of (a) the highest and lowest tone, (b) the mean frequencies, and (c) the range of frequencies are used as the parameters to determine whetheher the source of sounds and the new sound originally come from a similar source.

References

Arikunto, Suharsimi. 2000. *Prosedur Penelitian*. Jakarta: Rineka Cipta.

Crystal, David. 1997. *A Dictionary of Linguistics and Phonetics: Fourth Edition.* Cambridge, MA: Blackwell.

Dellwo, Volker. *Forensic Phonetics*. 2003. http://www.phonetiklabor.de/phonetiklabor/Inhalt/Unterrichtsmaterial/PDFs/Forensic Phonetics.pdf (July, 25, 2011)

Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton and Co.

Fromkin, V., R. Rodman, dan N. Hyams. 2003. *An Introduction to Language: Edisi Tujuh*. Boston: Thomson Heinle.

Kridalaksana, Harimurti. 2001. *Kamus Linguistik*. Jakarta: Gramedia Pustaka Utama.

Matthews, Peter. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.

Nolan, *et. al.* 2006. *A Forensic Phonetic Study of 'Dynamic' Sources of Variability in Speech: The DyVis Project.* Hal: 13-18.

Rahyono, F.X. 2005. *Aspek Fisiologi Bahasa dalam Pesona Bahasa: Langkah Awal Memahami Linguistik*. Jakarta: Gramedia Pustaka Utama.

Riyanto, Yatim. 1996. *Metodologi Penelitian Pendidikan Suatu Tindakan Dasar*. Surabaya: SIC Surabaya.

Sarwono, Joko. *Suara Manusia: Unik kah?*. 2009. http://jokosarwono.wordpress.com/2009/03/17/suara_manusia_unik_kah/ (July, 25 2011)

Sarwono, Joko. *Tak berkutik Karena Akustik*. 2008. http://dosen.tf.itb.ac.id/jsarwono/2008/08/23/tak-berkutik-karena-akustik/ (July, 25, 2011)

Silalahi, Ronald. 2011. *Peran Frekuensi Formant sebagai penanda Identitas: Sebuah Formulasi Fitur Fonetis Eksperimental untuk Kepentingan Forensic Acoustic*. Kongres Internasional Masyarakat Linguistik Indonesia (KIMLI) 2011.

Silalahi, Ronald. 2011. *The Role of Acoustical Features for the Identification of Personal Identity*. Persidangan Linguistik ASEAN 2011.

Verhaar, J.W.M. 1983. *Pengantar Linguistik Indonesia*. Jakarta: Gramedia.

Webb, David. 2007. *All About Forensic Science. com*

http://www.all-about-forensic-science.com/all-about-forensic-science-website.html (July, 25, 2011)

Wood, Sidney. 2005. *Praat for Beginners*. http://person2.sol.lu.se/SidneyWood/praate/whatform.html (July, 25, 2011)