IJFL (International Journal of Forensic Linguistic)
Vol. 3, No. 1 April 2022, Page 81-89

P-ISSN: 2723-1542
E-ISSN: 2723-5254
Available Online at https://www.ejournal.warmadewa.ac.id/index.php/ijfl/index

UTILIZATION OF PRAAT IN DETERMINING THE AUTHENTICITY OF VOICE

A.A. Istri Wirdayanthi
Universitas Warmadewa
agungistriwrdynthi@gmail.com

How to cite:

Wirdayanthi, A.A. Istri (2022). Utilization of PRAAT in Determining the Authenticity of Voice. .*IJFL (International Journal of Forensic Linguistic)*, Vol. 3 (1), 81-89. Doi: https://doi.org/10.22225/ijfl.3.1.4774.81-89

Abstract- Nowadays, the advances of technology have affected all the aspects in human life, especially the existence of smartphone and computer, it is very helpful in the process of communication, education, and even forensic stuff. In this case, PRAAT program which can change the sound waves that used to be heard only now become visible to the human eye. In addition, the PRAAT program is equipped with tools to see the tone motion, the magnitude of the pause, the length of speech all of which are needed to determine the inaccuracy of an utterance or the error of the utterance. Therefore, this study is aimed to describe and to find put the utilization of PRAAT in identifying the value of pitch and formant appropriately. This study use direct measurement method uses a laptop equipped with PRAAT software. This method is technically based on Forensic audio method. Furthermore, the data was obtained form the audio-recorded which have derived from one native speaker and ten non-native speakers as comparison sounds. This study took ten words of Indonesian language of the utterance that has been determined by women, the results of the analysis found that in the utterance of Indonesian speakers all words consisting of some vowels and sometimes pause, sometimes there is no pause. Nevertheless, it was also found that there were differences in the presence values of pitch and formant although in the same word, vowel, and even same speakers' gender.

Keywords: Audio-Forensic, Formant, PRAAT, Pitch.

I. INTRODUCTION

Generally, human's voice has different characteristics one each other, it can be based on gender, age, race, (ethnicity), emotion, etcetera. However, in identifying human voices without seeing faces, such as in telephone conversations, adult male or female voices are easier to identify than children's voices (Smith and Patterson., 2005). In this case, the aspect of age is easier to be analyzed because most of the children's voice characteristics are commonly not as heavy and loud as an adult voice.

It has been considered that communication directly (conversation by voice of human) is cheap, fast and very natural. However, the organ of speech and hearing sense are contributed mostly to help human to recognize and responses by training the brain to be able in recognizing the communication process continuously.

Therefore, the communication process can be occurred well in human daily life. In addition, the existence of media of technology has create the easier way of communication. Especially, the oral communication which involved the voice or the sound of human. On the contrary, the existence of smartphone, laptop, applications, and software have been used to manipulate and even to do a crime. In one case, almost all fields of daily human activity have used computers. Of course, it would be very good if computers could interact with humans through voice communication.

Thus, the field of forensic linguistics has become popular to solve the problem of law, police investigation, etcetera that sourced from language data such as voice, sound recording, wiretapping, and many more by using technology device as smartphone or computer. Ideally, smartphone technology in recent years has developed so rapidly. Indonesia is one of the countries with the fourth largest smartphone users after China, India and America. According to the digital marketing research institute E-marketer, in 2018 the number of active smartphone users in Indonesia is estimated to reach more than 100 million people. the developments that occur are not only in smartphone technology, but the supporting application system on the smartphone is also growing rapidly. One of them is in voice recording technology.

However, speech sound database in Indonesian is one of the speech sound databases that has not been fully developed and widely published. In fact, Indonesian is one of the most widely used languages in the world today. Voice recordings are often used by someone to capture a conversation in person or by telephone. So, in practice, sound recordings are used as evidence that can strengthen law enforcement charges during the trial process.

The use of voice for identity identification individuals have been widely applied by the police or other law enforcers (Shrivastava, Payal, Rastogi, and Tiwari, 2013)

For using voice recordings as evidence is necessary. In this case, the sound recording data processing was carried out. Identify the sound spectrum based on acoustic measurements for a boy's or girl's voice by measuring the value of the fundamental (pitch) and formant frequencies. Research on the analysis of the human voice spectrum based on gender and age groups using computers has been carried out.

Nowadays, the development of forensic linguistics has been applied massively. According to Olsson (2004) an application of linguistics in the context of crime, court proceeding, or arguments in law. Similarly, Coulthard and Johnson (2010) mention that forensic linguistics ranges from courtroom discourse and legal language to plagiarism. However, in the linguistic system, speech is one of the most basic things to determine the occurrence of communication.

Therefore, the use of voice for identity identification individuals have been widely applied by the police or other law enforcers (Shrivastava, Payal, Rastogi, and Tiwari, 2013). However, this process of voice identification needs some tools in acquiring and analyzing the voice scientifically. Ideally, for using voice recordings as evidence is necessary.

According to Malik (2013) First stage is recording data processing scientifically without manipulation so that the voice recording can be accepted as the evidence of law, police investigation, evidence in court, etcetera. Then, signal processing and deeper analyzing by relating to the field of science are needed. Technically, audio forensic is one of related study in conducting the voice recognition. Commonly, in applying audio forensic there are several steps consist of: the process of data collection (acquisition), feature analysis or features, evaluates and presents the final results in the form of reporting to be used as evidence (Maher, 2009).

One of the stages of audio forensics is: data acquisition, this initial process is quite important for attention, this is due to an error data sampling can result in evidence become useless or unusable. The acquisition process also needs to keep the data taken unchanged or updated (Ademu, Imafidon, and Preston, 2011).

Meanwhile, the process of acquisition in audio forensics in addition to evidence

recording as well required acquisition for comparison records. Comparison sound sample acquisition is necessary pay attention to several things including characteristics of the evidence record. Character differences voice recording may result in feature differences even though both voices are taken from individuals the same (Huizen, Jayanti, and Hostiadi, 2015). In order to maintain the retrieval, process the comparison is in accordance with what is needed then there needs to be a framework for the process. The framework used refers to the framework forensic audio. The purpose of the framework is to ensure the acquisition process is in accordance with provisions for identification (Ibrahim, Yavuzcan, and Ozel, 2013).

Moreover, according to Subki et al., (2018) Audio forensics is one of the sciences that juxtaposes between science and scientific methods in the process of sound recording analysis to assist and support the disclosure of a crime required in the trial process. However, The ITE Act No.19 of 2016 states that voice recording is one of the most valid digital instruments and can be used as an indictment. Voice recordings that are digital evidence are extremely easy and prone to be manipulated, either intentionally or unintentionally.

Similarly, AlAzhar Nuh, (2011) mentions in his book Audio Forensics: Theory and Analysis that sound recordings can be analyzed through the parameters of tone, formant, and spectrogram. This component can be used to identify the characteristics of a person's voice for speech recognition purposes by using the fragments of the analyzed voice recording. In addition, one of the digital forensic techniques is Voice Recognition, namely digital forensic techniques for detecting records Voice. People who have conversations can identity is known through audio forensic examination for speech recognition by comparison method, namely, comparing the voices in the recorded evidence (unknown sample) with sound recorded as Comparison (known sample). If the result of voice recognition indicates that the sound of the unknown sample is identical to known sample voice, then the voice in the conversation in the recording Evidence can be obtained from the owner of the vote Comparison. (Septiyansyah, 2015).

Technically, Binyamin Widi Prasetya, et al, (2008) stated that the digital forensic science is by definition a combination from the disciplines of law and computer science in collect and analyze data from computer systems, networks, wireless communications, and storage devices digital data for later use as evidence in problem-solving in the realm of law.

In the speech sound system, there are several things that allow it to be analyzed, such as the existence of different frequencies and pitches in each speaker. In this case, native and non-native speakers will have different pitches in pronouncing the same word. However, in measuring the magnitude of the frequency or decibels issued by the speaker, an accurate method of taking and measuring data samples is needed. The pitch height as well as the intensity in speech is especially easy to measure with words that have double vowels, or diphthongs. And this is what is used in this study to measure the frequency and intensity of sound in the utterance of words. The analysis carried out is in the form of contrastive analysis as an effort to observe differences, deviations or errors in pronouncing a word. According to Durao (Richards &Weber) (2007) The related software used in this study is PRAAT which is a phonetic program to analyze speech sounds, even though the speech is manipulative.

The PRAAT application provides a feature to record voice directly using an installed device, but it is still a computer-based program, so it is not yet available for mobile. After getting the sound recording, PRAAT can then dissect the sound into a detailed spectrogram and produce separate, more detailed parts of the sound. Spectrogram consists of low and high frequencies in the form of sound signals. In addition, it is also possible to measure the pitch, intensity and formant of the sound as a whole and display a graph to illustrate the results. On the other hand, the intonation of speech is varied among speakers, even in pronouncing the same word, there is always a tendency to have different intonations. Some people always end up with a descending intonation when pronouncing a word, but others choose a rising intonation in ending the pronunciation of a word.

Thus, this study is aimed to provide an explanation of analyzing process and the scenario the application of digital forensic techniques for voice recognizes compare the voices and the suspects' voices on mobile phone media and analyzed through PRAAT application with a voice of women. Moreover, a PRAAT application was used to assist the process of analyzing and comparison of the audio from native speaker and non-native Sample. Briefly, PRAAT is a computer program that is used to sound analysis, synthesis and manipulation. This app developed since 1992 by Paul Boersma and David Weenink at the University's Institute of Phoenix

Sciences Amsterdam. There are several versions released with customization for some common operating systems used are Mac, Windows and Linux. Since 2001, it has been 5000 registered users in 99 countries are using PRAAT.

According to Septiyansyah (2015) Praat app is able to record sound from microphone or other audio devices, besides that this application can also reads sound from an input file or disk. By PRAAT, then the user is able to see into the audio and this process can be recognized as forensic-digital process.

Based on the background above, the problem of this study can be formulated as followed: what are the utilizations of PRAAT program in forensic-linguistics? And how does PRAAT can be applied in identify the authenticity of human voices (women voices) through identification pitch and formant value?

Thus, based on those formulation of the problem, the aims of this study are consisted of to fine out the utilization of PRAAT application in forensic-linguistics and to describe the implementation of PRAAT in identify the authenticity of human voices through identification pitch and formant value.

II. METHODS

This study used the direct measurement method uses a laptop equipped with PRAAT software. This method is technically based on Forensic audio method. In addition, Data collection on the sample sound is done by recording sound using a smartphone's recording application that has been connected to a laptop equipped with praat software. The source of the data are the voices of women which are consisted of one native speaker (women) and ten women non-native speaker. Technically, this study uses software, such as PRAAT application and Microsoft excel. This app PRAAT is used to search information from the comparison between records of native speaker's voice and recorded comparison voice. Microsoft Excel to use to measure the formant, pitch of each word spoken original and comparison. However, this study will focus on identify and described the Analysis of Pitch Statistical, Forman and Bandwidth Statistical Analysis, and Spectrogram Analysis.

III. RESULT

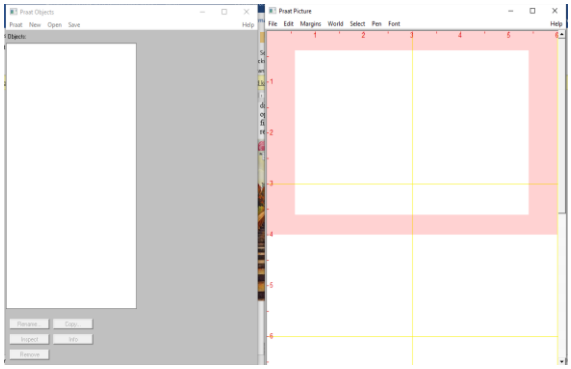
3.1. Analysis Techniques Using PRAAT Application

The first stage in using PRAAT application is recording the audio and transferred to the record program on the laptop which is equipped with the program of sound analyzer. This program is called as Gold-wave sound that contained in the record file is displayed waveform. After that, the part of each words/sentence which were displayed in the waveform are needed to cut of each part as a data to be analyzed.

After discovering the waveform of the words by boundary on praat, the value of pitch and formant can be analyzed appropriately.

Technically, users can directly apply sound data to this program without having to extract or convert the sound file as long as the audio format is in the form of wav to facilitate the analysis process. There are two blank displays when the PRAAT program is first opened; The first display is for inputting audio files and the second display is for the analysis result graph area. The image below is the initial preview of PRAAT.

Figure 3.1
Initial view of the PRAAT user interface



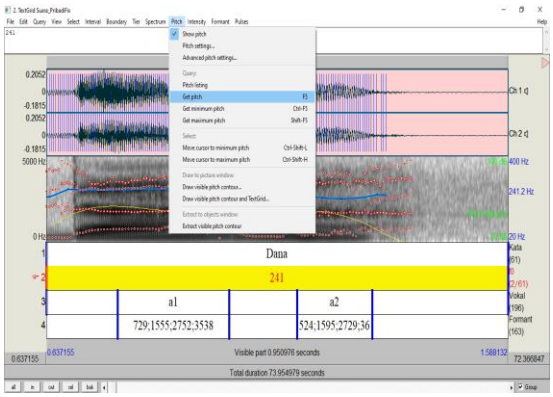
3.1.1 Pitch

Basically, speech sounds consist of several components, namely formant, pitch, intensity which can be used to identify a person's characteristics for the sake of voice recognition. In relation to the frequency of sound, pitch plays a major role in measuring how much the speaker's voice frequency is in pronouncing the word.

Every human being has a pitch that varies and is different so that it becomes unique. This is influenced by the physiological aspects of the larynx which have differences. Under normal conditions, the habitual pitch level ranges from 50-250 Hz for males and 120-500 Hz for females. This can be used linguistically to analyze a person's condition in terms of intonation, related to emotions.

In PRAAT, there is a setting that allows researchers to calculate the pitch of a person's voice that has been entered as a data source for analysis in the program. The pitch mark in PRAAT is blue with the values displayed on the right. To get information about the amount of pitch numbers, users can click on get pitch by previously activating (checking) show pitch.

Figure 3.2
The step for checking pitch's value



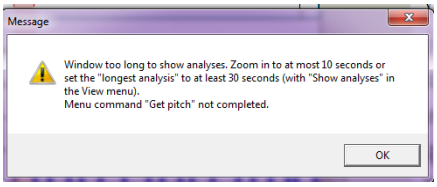
Since the analysis carried out is word includes the vowels of the recorded sound, then listen carefully to the words you want to analyze, then do the blocks on the graph. To get the right block, click the enter button in the corner left to zoom in. here's a brief explanation button located in the left corner.

Figure 3.3
Zoom button

All	All selection (to see the whole chart)
in	Zoom in (to enlarge the graphic view)
out	Zoom Out (to zoom out the graphic)
sel	Selection (to view block graphs only)
back	Back (back to the previous initial view)

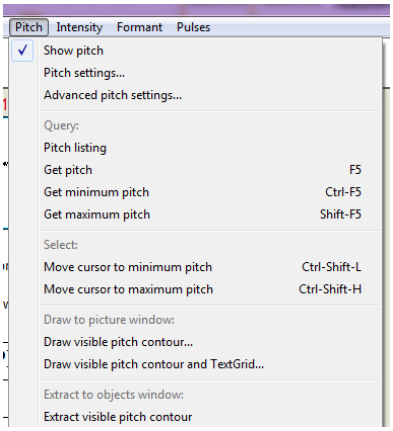
After the one-word graphic block then click File – Save the selected sound as a WAV file. In other words, that way cut or partition each word on each record. If a record contains 30 words, then there will be 30 save file as wav. After snipping each word, open the saved file snippet of the word to know its pitch value. Do not forget to turn on the pitch to see pitch value by making sure the Show pitch option is checked. The method to find out if the option is checked or not is to click the pitch menu tab, if you haven't already checked just click Show pitch. Because if not enable or check show pitch on the menu tab Pitch then the pitch value will not come out.

Figure 3.4
Pitch Value Warning Window



To view the minimum, maximum and mean, select the menu tab Pitch.

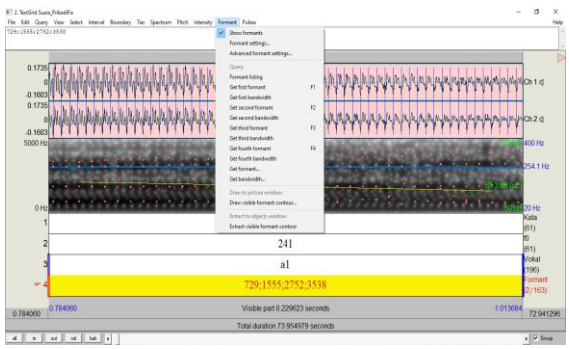
Figure 3.5
Menu minimum, maximum and mean pitch



3.1.2 Formant

In phonetics, formant is a sound wave measured in dB (decibels) and as an indicator of how strong the sound is. In PRAAT, intensity is used to extract sound into decibels. Intensity has a yellow color with a solid line and the value/measurement of the intensity can be obtained by clicking the Get formant List menu on the formant menu.

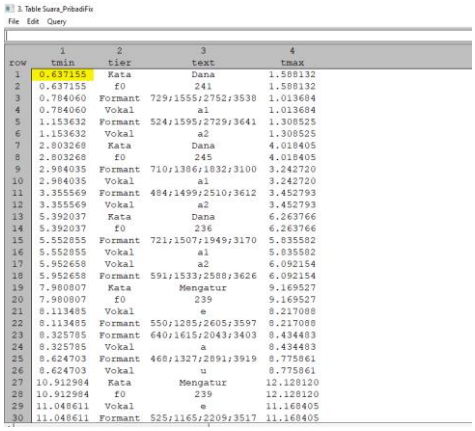
Figure 3.6
The step to get the value of formant



After that, the file will change its format as shown In the following image. After the file changes the format, what is done is to find the value of the formant numerically in the form Month so that it can be compared statistically later.

To find the value, click on the file that has been formant, then on the right menu, select tabulate – list. Then, the values of each formant (F1-F4) from word-1 until w0rd-10 will displayed as shown on the figure below:

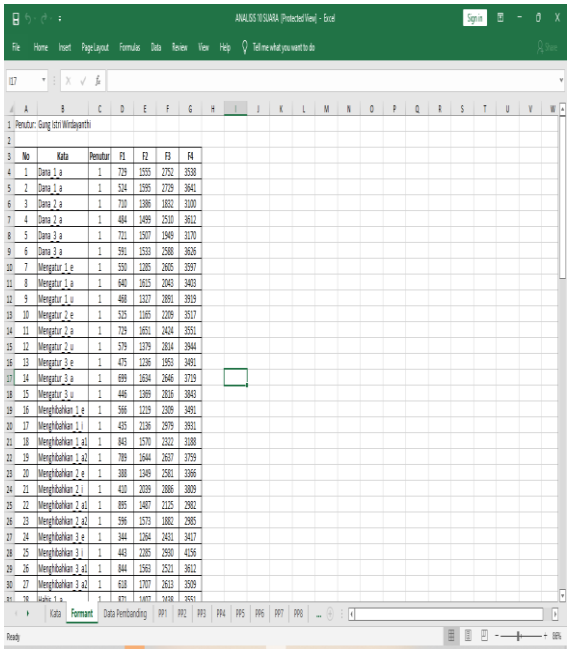
Figure 3.7
The tabulation values of formant on praat



row	1 tmin	2 tier	3 text	4 tmax
1	0.637155	Kata	Dana	1.588132
2	0.637155	f0	241	1.588132
3	0.784060	Formant	729;1555;2752;3538	1.013684
4	0.784060	Vokal	a1	1.013684
5	1.153632	Formant	524;1595;2729;3641	1.308525
6	1.153632	Vokal	a2	1.308525
7	2.803268	Kata	Dana	4.018405
8	2.803268	f0	245	4.018405
9	2.984035	Formant	710;1386;1832;3100	3.242720
10	2.984035	Vokal	a1	3.242720
11	3.355569	Formant	484;1499;2510;3612	3.452793
12	3.355569	Vokal	a2	3.452793
13	5.392037	Kata	Dana	6.263766
14	5.392037	f0	236	6.263766
15	5.552855	Formant	721;1507;1949;3170	5.835582
16	5.552855	Vokal	a1	5.835582
17	5.952658	Vokal	a2	6.092154
18	5.952658	Formant	591;1533;2588;3626	6.092154
19	7.980807	Kata	Mengatur	9.169527
20	7.980807	f0	239	9.169527
21	8.113485	Vokal	e	8.217088
22	8.113485	Formant	550;1285;2605;3597	8.217088
23	8.325785	Formant	640;1615;2043;3403	8.434403
24	8.325785	Vokal	a	8.434403
25	8.624703	Formant	468;1327;2891;3919	8.775861
26	8.624703	Vokal	u	8.775861
27	10.912984	Kata	Mengatur	12.128120
28	10.912984	f0	239	12.128120
29	11.048611	Vokal	e	11.168405
30	11.048611	Formant	525;1165;2209;3517	11.168405

After that, the results must be copied and pasted into the Microsoft Excel, so that the tabular formant results can be processed in the Microsoft Excel as shown below:

Figure 3.8
The display of data on Microsoft Excel



After all the data is summarized, the next step is to compare which suspect voices are has a formant value close to that of a native speaker.

3.1.3 Spectogram

Open the praat application, then open the file you want to view the spectrogram (file that has been partitioned word by word from original recording). Select the View & Edit menu then select the tab menu Spectrum and make sure the Show spectogram option is checked.

3.2 The result of data analysis

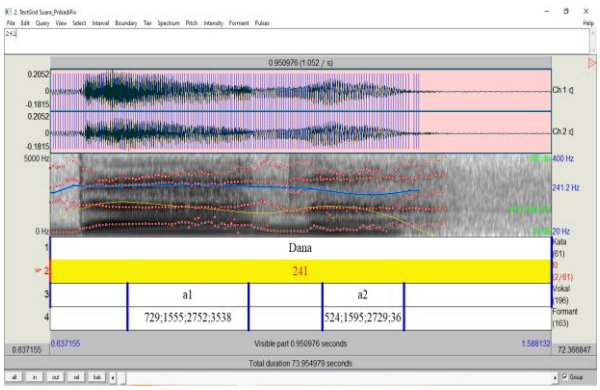
From the results of the analysis using the PRAAT program, from the eleven speakers, this study shows the identification values of pitch and formant of the ten comparison voices. Both native and non-native respondents are female respondents with almost the same voice range. This study uses ten Indonesian words spoken by the two respondents by recording.

Sample data:
Data 1
Words: Dana
Participants: One native speaker and 2 non-native speakers (women)

The analysis was carried out on the voice output of respondents in uttering ten words containing vowels that have different frequency and intensity characteristics. The results of the data in the form of an analysis of the frequency and intensity of sound are displayed in the form of tables and graphs.

Table 3.1
The values of pitch and formant
(Native Speaker)

No.	word	pitch	vowel	formant			
				F1	F2	F3	F4
1.	Dana	241	a-1	729	1555	2752	3538
			a-2	524	1595	2729	3641

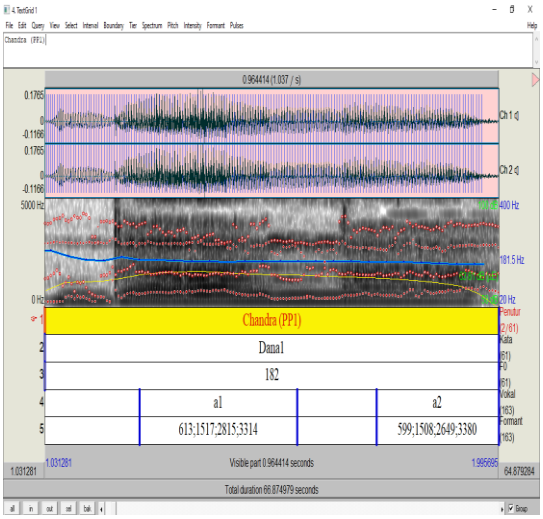


The pitch character of each voice is compared to the minimum pitch value, maximum pitch value and mean (average) pitch value. From the difference in the value of this pitch statistic, which can later help assess the level of similarity of the recording. Here is a sample data

Word: Dana
Participant 1 (PP1)

Table 3.1
The values of pitch and formant
(Non-Native Speaker)

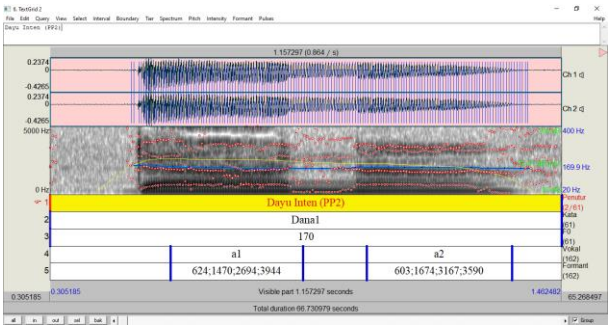
No.	word	pitch	vowel	formant			
				F1	F2	F3	F4
1.	Dana	182	a-1	613	1517	2815	3314
			a-2	599	1508	2649	3380



Word: Dana
Participant 2 (PP2)

Table 3.1
The values of pitch and formant
(Non-Native Speaker)

No.	word	pitch	vowel	formant			
				F1	F2	F3	F4
1.	Dana	170	a-1	624	1470	2694	3944
			a-2	603	1674	3167	3590



Based on the tables and the graph above, it is obvious that there is not similarity from each speaker both of native and non-native speakers although they are all the women. All graphs and spectrograms displayed in this study were obtained by processing sound data through extraction and processing voice data using the PRAAT program. Each graph shows the frequency and intensity of the speaker's voice; while the spectrogram shows all analysis data variables and their supporting elements, complete with pitch and intensity displays for each spectrogram.

IV. CONCLUSION

As conclusion, the process of data analysis by using the PRAAT program from record sound insert into program cut waves through the Gold-wave program inputting sound waves word by word is processed with the PRAAT program. So, the graph can be shown. Then, the word earlier chopped through the Gold-wave program into word for word to bring up the number of dB/Hz/second that is prepared to analyze the problem formulation that submitted. The PRAAT program can also calculate the amount of pause in seconds for knowing the accuracy of speech, the amount of time the speech is to see fluency rather than the utterance of a sentence, the movement of the tone to see the accuracy of the tone of the Indonesian words which uttered by women. As a final word with the existence of digital-forensic audio especially the PRAAT program is very helpful for the teaching-learning process foreign language or second language. Furthermore, it can be implemented to analyze the evidences in court by using this technique. However, it should be realized that the voice of human is not consistent which means that all people have different intonation, frequency, and intensity value.

REFERENCES

Al-Azhar Nuh, M. 2011. AUDIO FORENSIC :Theory and Analysis, 1–38.

Coulthard, M., & Johnson, A. (2010). Handbook of forensic linguistics. London: Routledge.

R. R. Huizen, N. K. D. A. Jayanti, and D. P. Hostiadi, “Analisis Pengaruh Sampling Rate Dalam Melakukan Identifikasi Pembicara Pada Rekaman Audio,” *Konf. Nas. Sist. Inform.*, pp. 9–10, 2015.

H. Ibrahim, H. G. Yavuzcan, and M. Ozel, “Digital forensics: An Analytical Crime Scene Procedure Model (ACSPM),” *Forensic Sci. Int.*, vol. 233, no. 1–3, pp. 244–256, 2013.

I. O. Ademu, C. O. Imafidon, and D. S. Preston, “A New Approach of Digital Forensic Model for Digital Forensic Investigation,” *Int. J. Adv. Comput. Sci. Appl.*,vol. 2, no. 12, pp. 175–178, 2011.

Smith, R. R., and Patterson, R. D. 2005. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *J. Acoust. Soc. Am.*, 118, 3177-3186.

Binyamin Widi Prasetya, Budi Santoso, and Joko Purwadi, "Identifikasi Suara Pria Dan Wanita Berdasarkan Frekuensi Suara," Informatics Department, Universitas Kristen Duta Wacana, Yogyakarta, 2008.

Subki, A., Sugiantoro, B., & Prayudi, Y. (2018). Membandingkan Tingkat Kemiripan Rekaman Voice Changer Menggunakan Analisis Pitch, Formant dan Spectogram. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *5*(1), 17. https://doi.org/10.25126/jtiik.201851500

Septiyansyah, Hasbi. 2015. Implementasi Metode Forensik Suara Pria

Menggunakan Teknik Voice Recognize Untuk Analisis Kemiripan Suara Pada Media Alat Rekam Telepon Selular. Jurusan Sistem Informasi, Universitas Teknologi sepuluh November, Surabaya.

Olsson, J. (2004). Forensic linguistics: An introduction to language, crime and the law. London and New York: Continuum.

Durao (Richards &Weber). (2007). *Contrastive approach in phonological disorders therapy: theoretical considerations*. Rio de Janeiro Brasil.

A. K. Shrivastava, N. Payal, A. Rastogi, and A. Tiwari, "Digital Forensic Investigation Development Model," *Int. Conf. Comput. Intell. Commun. Networks*, 2013.

R. C. Maher, "Audio Forensic Examination," *Ieee Signal Processing Magazine*, no. March, pp. 84–94, 2009.

H. Malik, "Acoustic Environment Identification and Its Applications to Audio Forensics," *IEEE Trans. Inf. Forensics Secur.*, vol. 8, no. 11, pp. 1827–1837, Nov. 2013