
**ANALYSIS OF VOICE LEVEL COMPARISON THROUGH
PITCH AND FORMANT VALUE IDENTIFICATION**

Yonatan E. Laia

. Universitas Warmadewa
jonathanxclusivelaiya@gmail.com

How to cite (in APA style):

Laia, Y. E. (2022). Analysis of Voice Level Comparison Through Pitch and Formant Value Identification. *IJFL (International Journal of Forensic Linguistic)*, Vol. 3 (1), 90-97. Doi: <https://doi.org/10.22225/ijfl.3.1.4614.90-97>

Abstract- This research is aimed to analyze the comparison of the voice level through the pitch and formant identification on the PRAAT software. In addition, this research focuses on a simple case study that involves ten speakers of men, native and non-native, in pronouncing ten words containing vocals with the help of a PRAAT application to calculate the Pitch and Formant of speech. Technically, the research method used in obtaining a comparison of the pitch and formant of the ten speakers is descriptive-analysis by presenting data in the form of tables of numbers and graphs. This research involved one native speaker and ten non-native speakers to get comparable results. The data retrieval method uses PRAAT with audio files as the main data source. The results of this study indicate that the accuracy and the validity of comparison speakers are very low. Based on the sampling of voice data from ten speakers, it can be concluded that the value of pitch and formant are not consistent and can be applied to prove the native and non-native speaker.

Keywords: Forensic Linguistic, Pitch, Formant.

I. INTRODUCTION

The development of technology and information has contributed the tremendous impact on humans' life in carrying out their daily activities. However, this situation is also become the trigger of improving the quality of human life in every aspect such as in economics, language implementation, politics, culture, laws, and even in religion. Thus, the humans' knowledge and the implementation of technology changed the paradigm of human that all the difficulties things will become the easier through tech-knowledge.

One aspect that changed drastically is linguistics and its interdisciplinary, such as forensic linguistics, computational linguistics, etcetera. Moreover, it has been admitted that the aspect of linguistics can be applied for others aspect such as forensic-linguistics specifically.

In this case, the implementation of forensic-linguistics is assisted by some tools or media that refers to the technology which able to analyze the language component such as utterance (voice material) and even the human's expressions. Specifically, forensic linguistics aims to express and prove the truth of the law through language. In this last decades, forensic linguistics has developed rapidly especially in law, judicial procedures, police investigation process. According to Coulthard and Johnson (2010) mention that forensic linguistics ranges from courtroom discourse and legal language to plagiarism.

Therefore, it can be stated that forensic linguistic analysis of conversational implicature from the perspective of symbolic meaning. On the other hand, the public assumption that actions often occur violence to reveal true confessions in police interrogation. Thus, the implicature conversation is one strategy to avoid acts of violence as an effort to prevent violations of the law in carrying out their duties in the police.

Technically, the forensic linguistics has been evolved in to the digital processing data. It means that the process of analyzing data of language in forensic is assisted by the digital devices such as smartphone, computer, software and applications.

One of the digital forensic techniques is Voice Recognition, namely digital forensic techniques to identify records voice. People who have conversations can identity is known through audio forensic examination for voice recognition with a comparative method, namely compare the voices in the recorded evidence (unknown samples) with sound recorded as comparison (known samples). If the results of voice recognition indicates that the sound of unknown samples is identical to the voice of known samples, then the voice of the conversation in the recording the evidence can

be concluded from the owner of the vote comparison.

It is similarly to the ITE Law No. 19 of 2016 which mentions that voice recordings are one of the pieces of evidence, as described in Article 1. It's just that sound recordings cannot be used as evidence without going through a fairly long analysis process, which is carried out by an expert in the field of audio forensics.

Moreover, 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. In some situation for 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 study, the voices of each participant have been recorded through the smartphone's voice record application and use the praat software to analyze the data in audio form. However, 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).

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).

In this case, the comparison of voice level is needed to identify and prove the truth of the original voice and manipulation. It is very necessary for judges or police to get the valid data from the cases which involve language material.

Moreover, the data in form of voices are not limited found in direct conversation, but also through the online media such as whatsapp voice or video call, facebook, youtube video, etcetera which is not only personal but can be taken by several parties such as prosecutors, police, and the corruption eradication commission. however, it is not as easy and fast

as it is decided that the validity of the vote is correct. rather it requires a comparative voice to corroborate the forensic evidence in the law.

Therefore, to express the truth of case's evidence needs the tool in analyzing the data. Commonly, PRAAT is one of software in computer that can be used to solve this problem.

Specifically, PRAAT (Dutch means 'sound') is a freeware developed by Paul Boersma & David Weenink of the University of Amsterdam's Phonetic Sciences Department. The PRAAT program is a program containing tutorials on the analysis and reconstruction of acoustic speech signals. PRAAT software can be downloaded on the internet at the address: <http://www.fon.hum.uva.nl/praat/>.

PRAAT software can be used on different operating systems which can be seen on the web above. PRAAT is a very flexible program for analyzing human speech. In addition, this program has a wide range of standard procedures such as spectrographic analysis, articulatory synthesis and neural networks. In this praat program several topics such as:

- a) Search manually
- b) Creating a sound object
- c) Processing signals
- d) Waveform label
- e) General analysis (wave, intensity, sonogram, pitch, duration)
- f) Spectrographic analysis
- g) Intensity analysis
- h) Pitch analysis
- i) Using long sound files

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.

Therefore, based on the background above, this study formulated the research problem as followed:

- a. How do the voices level can be compared through pitch and formant value identification technique through PRAAT software?
- b. How does PRAAT operate in identifying the original speaker's voices and the comparison voice?

Based on the formulation of the problem above, this study is aimed to analyze the comparison of voices level through pitch and formant value identification technique through PRAAT software and to describe the operational steps of PRAAT that use in identifying original speaker's voices and the comparison voice.

II. METHODS

In this study, descriptive-analysis are used simultaneously because the data and results are displayed and described on number and graphs. However, technique of data analysis uses the Forensic audio technique.

The direct measurement method uses a laptop equipped with PRAAT software. Data collection on the sample sound is done by recording sound using a microphone that has been available on smartphone and sent to a laptop equipped with PRAAT software.

Technically, this study uses software, such as PRAAT application and Microsoft excel. This app 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, the researcher only focuses to find out the value of pitch (F0) and Formant (F1-F4).

The sources of the voices are derived from 10 men. The basic frequency values obtained from the voice recordings were compared between the voices of men, the formant values obtained.

III. RESULT

From the results of the analysis using the PRAAT program, from the ten speakers, this study shows a comparison of the difference in pitch and intensity of the two voices. Both native and non-native respondents are male respondents with almost the same voice range. This study uses ten Indonesian words spoken by the ten respondents by recording.

3.1. The Data Analysis

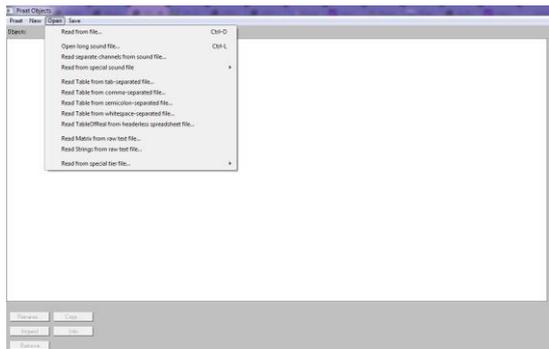
The analysis was carried out on the voice output of respondents in uttering ten words that have different frequency and intensity characteristics. The results of the data in the form of an analysis of the frequency and intensity of sound are displayed in the form of tables and graphs. Broadly speaking, the results of the analysis using PRAAT produce the following details:

3.1.1 Pitch

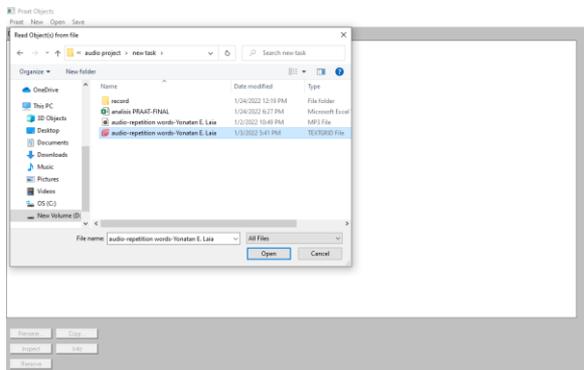
In analyzing the value of pitch on the PRAAT. firstly, sound files that have been noise filtered into praat app are imported. Here are the steps of identifying the value of pitch:

- a. Open the Praat application, then select the Open menu.
- b. select Read from file, then select the sound recording file you want to download analysis.

Picture 3.1
Step 1, Pitch value identification

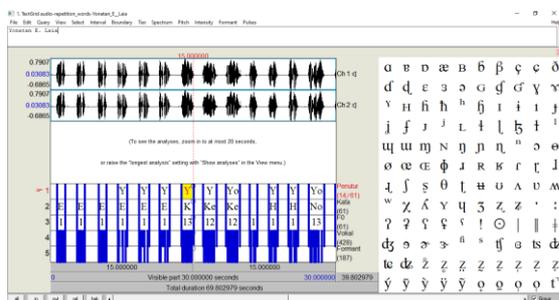


Picture 3.2
Step 2, Pitch value identification



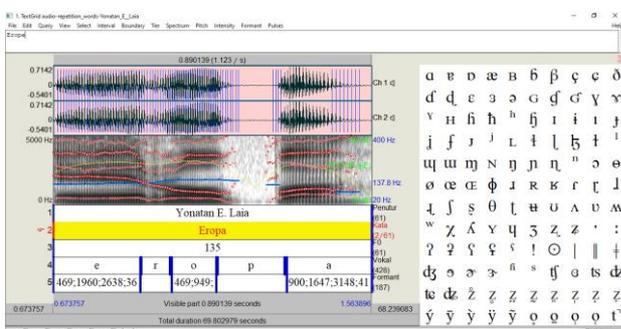
Then, select *view and edit*. The window will be as the picture below:

Picture 3.2
Step 3, Pitch value identification



After that, listen carefully to the words you want to analyze, then do the blocks on the graph.

Picture 3.4
Step 4, Pitch value identification



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.

Picture 3.4

Zoom Button on Praat

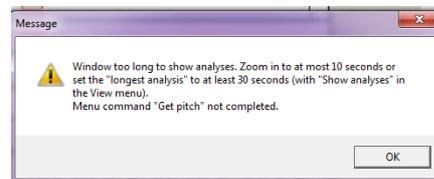


| | |
|------|--|
| All | All selection (to see the whole chart) |
| In | Zoom in (to enlarge the graphic view) |
| out | Zoom Out (to zoom out the graphic) |
| Sel | Selection (to view block graphs only) |
| back | Back (back to the previous initial view) |

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.

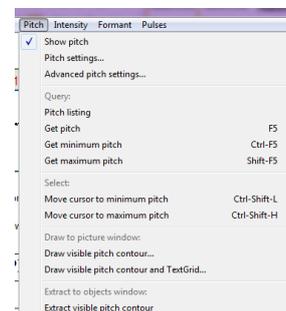
Remind to turn on the pitch to see pitch value by making sure the Show pitch option is checked. The method to find out if the option is checked or not is to click the pitch menu tab, if you haven't already checked just click Show pitch. Because if not enable or check show pitch on the menu tab Pitch then the pitch value will not come out.

Picture 3.5
Pitch Value Warning Window



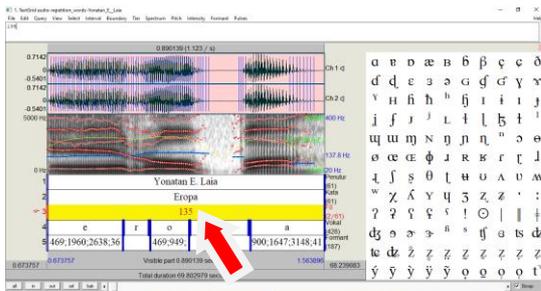
To get the value of minimum, maximum and mean, select the menu tab Pitch.

Picture 3.6
Step 5, Pitch value identification
Menu minimum, maximum and mean pitch



Picture 3.7

Step 5, Pitch value identification
The sample of pitch value



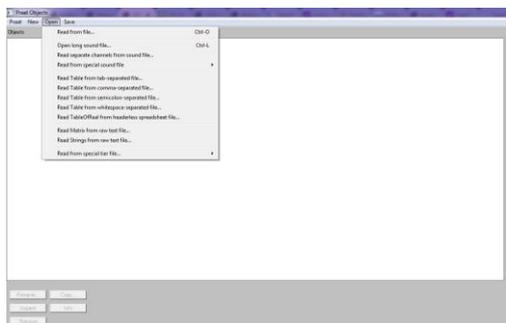
From the pitch value above, it can be seen that the pitch of word “EROPA” is 135 Hz.

3.1.2 Formant

In identifying the value of formant, there are several steps as shown on the pictures below:

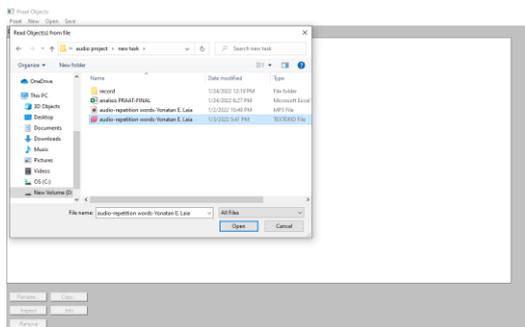
1. open praat app, import file audio that has been per-worded by clicking the open menu and select read from file

Picture 3.8
Step 1, Formant value identification



2. So a dialog box will appear to select a file as following:

Picture 3.9
Step 2, Formant value identification
Import file dialog box

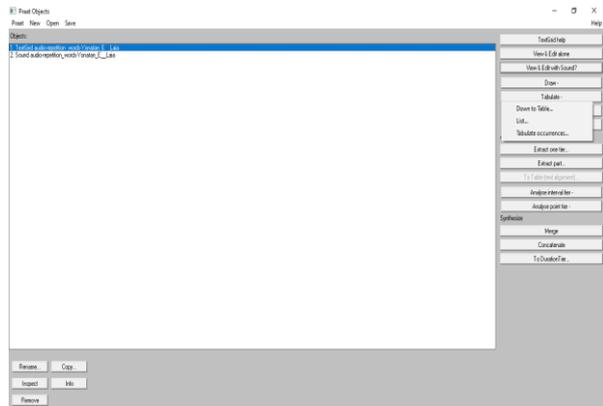


After selecting a file, the file will be listed in the left-field but still in audio/sound format.

To change it, click on the right menu of Analyze Spectrum and select formant.

Next step is changing file 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.

Picture 3.10
Step 3, Formant value identification
Import file dialog box



After selecting the tabulate-list menu, the windows will show the result of formant (F1-F4) as shown on the table below:

Picture 3.11
Step 4, Formant value identification
The value of formant

| Praat Info | | | | |
|------------|---------|--------------------|----------|------|
| File | Edit | Search | Convert | Font |
| 0.673757 | tier | text | tmax | |
| 0.673757 | Formant | 469;1960;2638;3684 | 0.917275 | |
| 0.673757 | Vokal | e | 0.917275 | |
| 0.673757 | Penutur | Yonatan E. Laia | 1.563896 | |
| 0.673757 | F0 | 135 | 1.563896 | |
| 0.673757 | Kata | EROPA | 1.563896 | |
| 0.919609 | Vokal | r | 0.999744 | |
| 1.004412 | Vokal | o | 1.128894 | |
| 1.004412 | Formant | 469;949;2894;3476 | 1.128894 | |
| 1.133562 | Vokal | p | 1.320319 | |
| 1.326483 | Vokal | a | 1.563896 | |
| 1.326483 | Formant | 900;1647;3148;4120 | 1.563896 | |
| 2.808097 | Vokal | e | 3.009456 | |
| 2.808097 | Formant | 501;1893;2797;3657 | 3.009456 | |
| 2.808097 | Penutur | Yonatan E. Laia | 3.618276 | |
| 2.808097 | F0 | 133 | 3.618276 | |
| 2.808097 | Kata | EROPA | 3.618276 | |
| 3.010859 | Vokal | r | 3.111847 | |
| 3.114652 | Formant | 502;969;3027;3426 | 3.208628 | |
| 3.114652 | Vokal | o | 3.208628 | |
| 3.211433 | Vokal | p | 3.429540 | |
| 3.433748 | Formant | 795;1457;2958;3922 | 3.618276 | |
| 3.433748 | Vokal | a | 3.618276 | |
| 5.042746 | Formant | 516;1958;2835;3778 | 5.188244 | |
| 5.042746 | Vokal | e | 5.188244 | |
| 5.042746 | Kata | EROPA | 5.895454 | |
| 5.042746 | F0 | 137 | 5.895454 | |
| 5.042746 | Penutur | Yonatan E. Laia | 5.895454 | |
| 5.191197 | Vokal | r | 5.332916 | |
| 5.335131 | Vokal | o | 5.470207 | |
| 5.335131 | Formant | 487;949;2926;3557 | 5.470207 | |
| 5.471693 | Vokal | p | 5.699763 | |
| 5.704930 | Formant | 784;1388;2896;3856 | 5.895454 | |
| 5.704930 | Vokal | a | 5.895454 | |
| 7.088019 | Formant | 517;2030;2766;3733 | 7.183703 | |
| 7.088019 | Vokal | e | 7.183703 | |
| 7.088019 | F0 | 128 | 8.065198 | |
| 7.088019 | Kata | Ekspor | 8.065198 | |
| 7.088019 | Penutur | Yonatan E. Laia | 8.065198 | |
| 7.187933 | Vokal | x | 7.553347 | |
| 7.557577 | Vokal | p | 7.587182 | |

To analyze the values of formant, the data will be copied and pasted Copy and paste into the Microsoft Excel, so that the tabular formant results can be processed in the Microsoft Excel as shown below:

Picture 3.11
Step 5, Formant value identification
The value of formant

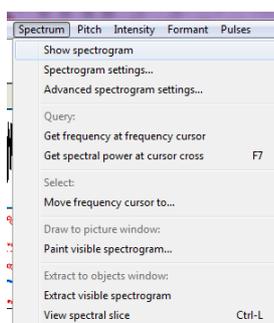
| No. | Kata | Penutur | F1 | F2 | F3 | F4 |
|-----|----------------|---------|-----|------|------|------|
| 1 | Eropa_1_e | 2 | 469 | 1960 | 2638 | 3684 |
| 2 | Eropa_1_o | 2 | 469 | 949 | 2894 | 3476 |
| 3 | Eropa_1_a | 2 | 900 | 1647 | 3148 | 4120 |
| 4 | Eropa_2_e | 2 | 501 | 1893 | 2797 | 3657 |
| 5 | Eropa_2_o | 2 | 502 | 969 | 3027 | 3426 |
| 6 | Eropa_2_a | 2 | 795 | 1457 | 2958 | 3922 |
| 7 | Eropa_3_e | 2 | 516 | 1958 | 2835 | 3778 |
| 8 | Eropa_3_o | 2 | 487 | 949 | 2926 | 3557 |
| 9 | Eropa_3_a | 2 | 784 | 1388 | 2896 | 3956 |
| 10 | Ekspor_1_e | 2 | 517 | 2030 | 2766 | 3733 |
| 11 | Ekspor_1_o | 2 | 440 | 825 | 2989 | 3607 |
| 12 | Ekspor_2_e | 2 | 501 | 2070 | 2718 | 3768 |
| 13 | Ekspor_2_o | 2 | 457 | 799 | 2983 | 3603 |
| 14 | Ekspor_3_e | 2 | 549 | 1960 | 2741 | 3711 |
| 15 | Ekspor_3_o | 2 | 454 | 830 | 2945 | 3533 |
| 16 | Kewalahan_1_e | 2 | 376 | 1208 | 2843 | 3582 |
| 17 | Kewalahan_1_a1 | 2 | 702 | 1258 | 2891 | 3611 |
| 18 | Kewalahan_1_a2 | 2 | 754 | 1258 | 2827 | 3656 |
| 19 | Kewalahan_1_a3 | 2 | 662 | 1441 | 2841 | 3808 |
| 20 | Kewalahan_2_e | 2 | 393 | 1257 | 2802 | 3669 |
| 21 | Kewalahan_2_a1 | 2 | 676 | 1213 | 2902 | 3602 |
| 22 | Kewalahan_2_a2 | 2 | 770 | 1329 | 2951 | 3804 |

As result the values of pitch and formant are ready to be compared and analyzed to identify the native speaker and non-native speaker depended on the selected words which has been arranged.

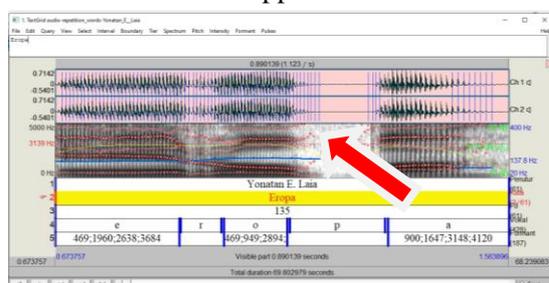
3.1.3 Spectrogram

Spectrogram is also one of the important components in analyzing the voices on praat. Here are several steps to get the spectrogram, 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 spectrogram option is checked.

Picture 3.12
Spectrogram



Then a window will appear as below:



3.2. The Data Result of the Research

From all the data which have been analyzed. The researcher has taken the sample to be displayed in this part. Thus, the sample data which has been chosen can be seen on the explanation below:

Sample data:

- Word : Eropa
- Frequency of voice recording: 3 times on each speaker.
- PP (Penutur Pembanding)/ Comparative Speaker

| Speaker | Pitch (F0) | | |
|----------------|------------|-----|-----|
| | W-1 | W-2 | W-3 |
| Native speaker | 134 | 138 | 137 |
| PP1 | 114 | 127 | 206 |
| PP2 | 100 | 152 | 251 |
| PP3 | 140 | 183 | 281 |
| PP4 | 104 | 171 | 297 |
| PP5 | 123 | 189 | 257 |
| PP6 | 168 | 215 | 292 |
| PP7 | 154 | 211 | 285 |
| PP8 | 110 | 149 | 300 |
| PP9 | 112 | 179 | 350 |
| PP10 | 107 | 191 | 309 |

The data above was obtained by analyzing the Pitch (F0) value of the respondents (male comparison speakers) three times with variations in tone starting from low, medium, and high notes. Except for the native speaker only at one tone level. Therefore, from the data above, it can be stated that:

- The original speaker's voice has a ratio of F0 values in a not too far range in the same pitch,
- The comparison voice data (PP1-PP10) mostly does not have a pitch value that is identical to the original speaker's F0 voice, this is because the comparison speaker's pitch value does not approach the original speaker's F0 value with a wide range. However, some data also have a value ratio that is almost the same as the voice data of native speakers.
- As conclusion, the voice data of the identical comparison speaker has a pitch value that is close to the pitch value of the original speaker's voice data.
- On the other hand, F0 analysis through praat cannot show the accuracy that a data is true voice data of the original speaker. This is because the data obtained from the comparison speaker's pitch value which has the same value and the range of values is not too far away.

Therefore, based on the results of the data analysis above, it can be stated that the accuracy

of the pitch analysis (F0) through Praat is very low and cannot be used as an indicator to distinguish that a data is correct as the original voice, as evidenced by the results of the F0 value of the original voice data and a very flexible and inconsistent comparison, some are the same, close to and even far from the F0 value of the subject data. This shows that everyone has a different pitch value because the intonation of each person's word pronunciation is different. It is possible that there are pitch values from several subjects that are almost the same.

Briefly, the value of formant can be seen on the sample data below:

Word : Eropa

PP (Penutur Pembanding)/ Comparative Speaker

| Speaker | Formant | | | |
|----------------|---------|------|------|------|
| | F1 | F2 | F3 | F4 |
| Native Speaker | 469 | 1960 | 2638 | 3684 |
| PP1 | 509 | 1903 | 2553 | 3477 |
| PP2 | 429 | 1814 | 3084 | 3775 |
| PP3 | 571 | 2029 | 2826 | 3637 |
| PP4 | 522 | 1866 | 2707 | 3522 |
| PP5 | 494 | 2035 | 2682 | 3911 |
| PP6 | 473 | 2096 | 2642 | 3790 |
| PP7 | 406 | 1824 | 2513 | 3368 |
| PP8 | 446 | 2128 | 2798 | 3696 |
| PP9 | 400 | 1991 | 2601 | 3837 |
| PP10 | 476 | 2044 | 2450 | 3449 |

The sample data above shows the results of the formant value on the word "EROPA_1" from the voice data of native speakers and 10 voice data of comparison speakers. Therefore, from the results of the data analysis above, researcher found that:

The comparison voice data (PP1-PP10) has a different formant value from the original speaker's voice data. However, if the argument is on a range of values, then some of the comparison speaker's formant values have a value range that is quite close to the original speaker's formant value, although the accuracy is still not met.

The formant value of the native speaker's voice data can be clearly distinguished because it has a different value from the formant value of the comparison speaker's data. On the other hand, indicators for the difference in formant values have not been found because the formant values of each data are inconsistent. Starting

from a lower formant value, approaching, and even higher than a native speaker's formant value.

IV. CONCLUSION

Based on the discussion above, it can be concluded that the comparison of native speakers and comparison speakers does not show the accuracy of the value because the range of values is sometimes very close and even the same and the range is quite far. However, the implementation of digital devices in forensic linguistics can provide a comparative perspective when analyzed more deeply with other collaboration device. Thus, it can be stated that the character of the voice of each participant both as the original voice and the voice of comparison. this proves that humans have variations in intensity, intonation, and frequency. so, it is possible to give the same value of pitch and formant as well as the difference.

REFERENCES

- 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.
- M. Nuh Al-Azhar. (2011, Juni) scribd.com. [Online].<https://www.scribd.com/doc/58248062/AudioForensic-TheoryandAnalysis-By-Muhammad-Nuh-ALAzhar>
- 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.

