IJFL (International Journal of Forensic Linguistic)
Vol. 3, No. 1 April 2022, Page 98-104

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

IMPLEMENTATION OF FORENSIC METHOD USING VOICE RECOGNITION TECHNIQUE TO ANALYZE VOICE RESEMBLANCE TOWARDS MOBILE PHONE'S VOICE RECORDER

Nur Afnita Asfar
. Universitas Warmadewa
nurafnitaasfar07@gmail.com

How to cite (in APA style):

Asfar, N. A. (2022). Implementation of Forensic Method Using Voice Recognition Technique to Analyze Voice Resemblance Towards Mobile Phone's Voice Recorder. *IJFL (International Journal of Forensic Linguistic)*, Vol. 3 (1), 98-104. Doi: https://doi.org/10.22225/ijfl.3.1.4606.98-104

Abstract- This article is intended to analyze the implementation of the forensic method using voice recognition technology to analyze voice resemblance towards mobile phone's voice recorder, by using Forensic audio method In Audio forensics: Theory and Analysis, namely Pitch Statistical Analysis, Forman and Bandwidth Statistical Analysis, Graphical Distribution Analysis and Spectrogram Analysis. However, in this study, researchers only focus on identifying pitch and formants in the data to be analyzed. This study provides an example scenario of applying digital forensic techniques for voice pain recognition. Greetings compare the sound in the evidence and the sound suspect on mobile phone media with a voice case study woman. In this study, PRAAT application was used to help with the audio comparison process from Known Samples and Unknown Samples. The results of each sample of digital evidence of native speakers' voices and recorded comparisons on the two assessment factors, namely Pitch and Formant, the researcher cannot distinguish between the original speaker's voice and the comparison voice. So it is hoped that the output of this research can become a reference or enrichment material for law enforcement agencies, law, and academics who wish to continue their research related to digital forensics. The researcher can state that the accuracy of the pitch and formant analysis through PRAAT is very low and cannot be used as an indicator to distinguish that a date is correct as of the original voice, as evidenced by the results of the F0 value of the original voice data and the comparison obtained. Very flexible and inconsistent. 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 each person's word pronunciation is different. There may be pitch values from several subjects that are almost the same.

Keywords: Forensic Linguistic, Praat Aplication, Forensic audio.

I. INTRODUCTION

The field is known as "forensic linguistics" has been growing in prominence in the past couple of decades. Forensic linguistics is all about taking linguistic insight, method, and knowledge in the context of law, judicial procedures, police investigations, trials, and in short, about studying the language of law and solving crimes. Olsson (2004) defines it as an application of linguistics in the context of crime, court proceedings, or arguments in law. Coulthard and Johnson (2010) mention that forensic linguistics ranges from courtroom discourse and legal language to plagiarism. Briefly, plagiarism uses another person's work for personal advantage without mentioning their name. Forensic linguistic experts proficient in plagiarism cases and copyright infringements prove which position is based upon another.

The development of multimedia technology is currently increasingly facilitating human activities in daily life, including how technology can store audio digitally. Digital audio storage is typically used for interviews or education, which is commonly used for storage media is a sound recorder or use similar apps found on a cellphone. Voice recording is often used by someone to immortalize a conversation Directly or by telephone. In practice, voice recordings are used as evidence To strengthen the law enforcement charges during the trial process. The ITE Law No. 19 of 2016 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. 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. Digital forensic science is, by definition, a combination of the disciplines of law and computer science in collecting and analyzing 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. (Binyamin Widi Prasetya, dkk, 2008)

In its application, digital forensic science is often helpful to authorities in uncovering related crime cases the suspect is concerned about through the evidence that has been collected. The science of sound forensics focuses on analyzing the suitability or originality of proper content material with the original content for later testing reliability and validity (Detik.com, 2011). With the increasing development of technology, more and more use is increasing. One of which is the discovery of several cases of legal irregularities accompanied by evidence in the form of sound recording media. Case this indicates a possibility ahead of audio digital will be used more as evidence in legal cases. Legal considerations in using evidence in the form of digital files, Including audio, are the ability of digital evidence to manage the impact associated with risk on the process law. One of the risks in question is using witnesses who have not known with certainty the truth, even though he has sworn to speak the truth. Using digital evidence that has been tested and analyzed will support the discipline of action and accuracy of guesses and help inaccuracy Decision-making. For that, it is essential for the parties relevant law enforcement agencies to understand and master digital forensic techniques considering the possibilities of increasing use of sound recordings/good multimedia In terms of variety, quantity, and quality. One of the digital forensic techniques is Voice Recognition, namely digital forensic techniques for detecting records Voice. People who have conversations can identify 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 model). If the result of voice recognition indicates that the sound of the strange piece is identical to the 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). Aligarh (2016) was researched to create an environment as natural as possible, conditions Retrieval, And results of the Forensic method used. In this study, forensic testing of sound evidence is carried out using pitch, formant, and spectrogram values, then comparing the sound of the evidence (unknown samples) with recorded sound as a comparison (known samples).

This study provides an example scenario of applying digital forensic techniques for voice pain recognition. Greetings compare the sound in the evidence and the sound suspect on mobile phone media with a voice case study woman. So it is hoped that the output of this research can become a reference or enrichment material for law enforcement agencies, law, and academics who wish to continue their research related to digital forensics. In this study, a Praat application was used to help with the audio comparison process from Known Samples and Unknown Samples. Praat is a computer program used to sound analysis, synthesis, and manipulation. This app was developed in 1992

by Paul Boersma and David Weenink at the University's Institute of Phoenix Sciences Amsterdam. Several versions are released with customization for some standard operating systems used: Mac, Windows, and Linux. Since 2001, it has been 5000 registered users in 99 countries have been using Praat. Septiyansyah (2015) stated that the Praat app could record sound from a microphone or other audio devices. Besides that, this application can also read sound from an input file or disk. With Praat, then the user can see into the audio. This research uses the Forensic audio method. Based on the background above, the formulation of the problem in this research is What are the results of each sample of digital evidence of native speakers' voices and recorded comparisons on the two assessment factors, namely Pitch and Formant, the researcher can distinguish between the original speaker's Voice and the comparison voice?

II. METHODS

In the method, this research uses the Forensic audio method. In Audio forensics: Theory and Analysis, namely Pitch Statistical Analysis, Forman and Bandwidth Statistical Analysis, Graphical Distribution Analysis, and Spectrogram Analysis. However, in this study, researchers only focus on identifying pitch and formants in the data to be analyzed.

This research uses software, such as praat application and Microsoft excel. This app praat is used to search information from the Comparison between native speaker's voice records and recorded comparison voice. Microsoft Excel to measure the formant pitch of each word spoken original and Comparison.

III. RESULT

3.1 Process with Praat

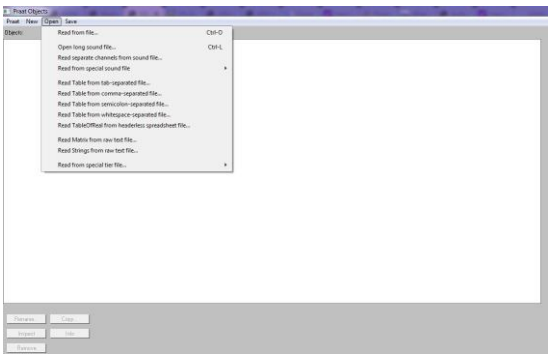
This process is the core process of this research. With the Praat application, you can find out the pitch, formant and spectrogram of each sound recording. The following is the implementation of the pitch, formant and spectrogram:

3.1.1 Pitch

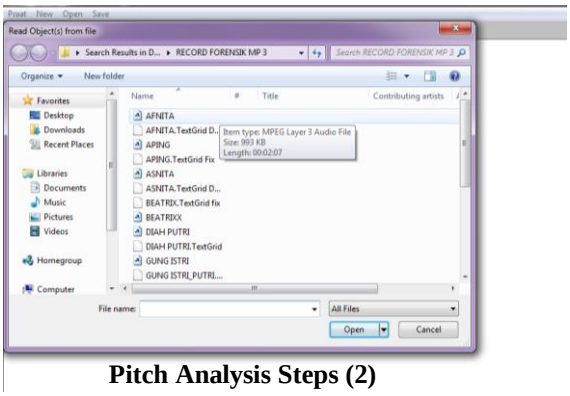
Pitch analysis can be used to perform voice recognition on a person's voice, that is, through statistical analysis of minimum pitch, maximum pitch, and average pitch. Pitch is the frequency vibrating vocal cords (Jose R.L Batubara, 2010). One of the parameters of the sound signal is the frequency fundamentals. Fundamental frequency in terms of musical instrument known as pitch or frequency value of a tone type. The faster the vibration of the vocal cords, the higher the pitch, and vice versa.

Each person has a distinctive pitch (habitual pitch) strongly influenced by physiological aspects of the human larynx. In normal conversation conditions, habitual pitch levels range from 50-250 Hz for men and 120-500 Hz for women (M. Nuh AL-Azhar, 2012).

To analyze the pitch on the praat, the first thing to do is import sound files that have been noise filtered into the praat app. Open the Praat application, select the Open menu, select Read from a file, and select the sound recording file you want to download analysis.To analyze the pitch on the praat, the first thing to do is import sound files that have been noise filtered into the praat app. Open the Praat application, select the Open menu, select Read from a file, and select the sound recording file you want to download analysis.

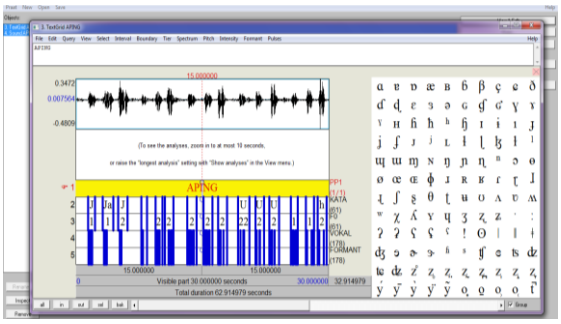


Pitch Analysis Steps (1)



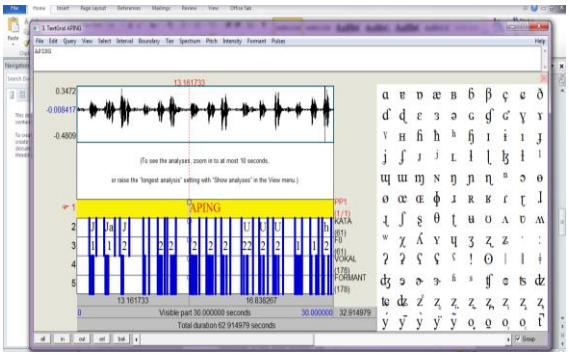
Pitch Analysis Steps (2)

After that select *view and edit*. Then a window like the one below will appear.



View and edit menu window

Because the analysis carried out is word for word which includes the vowels of the recorded sound, then listen carefully to the words you want to analyze, then do the blocks on the graph.

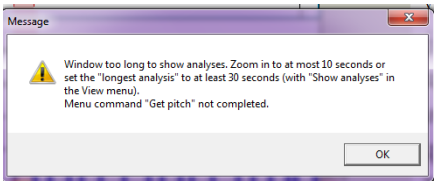


To get the right block, click the enter button in the corner left to zoom in. here's a brief explanation button located in the left corner.

Zoom Button

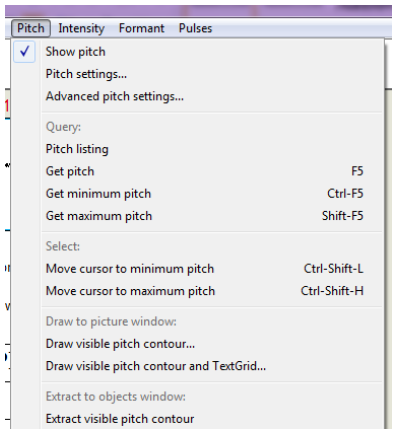
All	All selection (to see the whole chart)
in	Zoom in (to enlarge the graphic view)
out	Zoom Out (to zoom out the graphic)
sel	Selection (to view block graphs only)
back	Back (back to the previous initial view)

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.



Pitch Value Warning Window

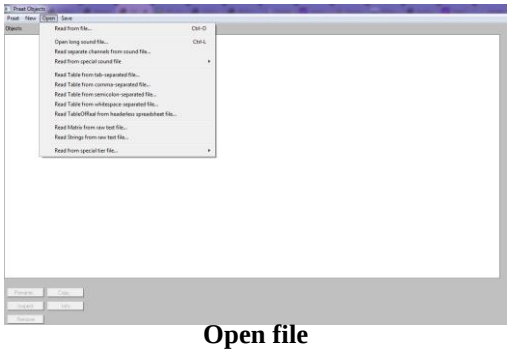
To view the minimum, maximum and mean, select the menu tab Pitch.



Menu minimum, maximum and mean pitch

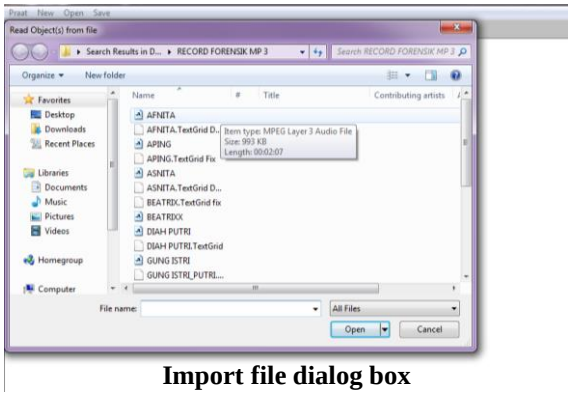
3.1.2 Formant

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



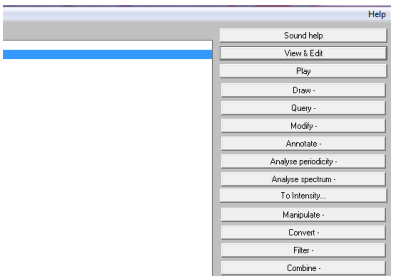
Open file

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



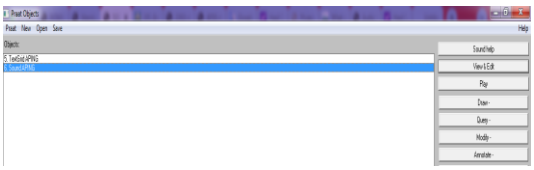
Import file dialog box

After selecting a file, the file will be listed in the left-field but still in audio/sound format. Meanwhile, to process formants, the required data must be in informant format. To change it, click on the right menu of Analyze Spectrum and select formant.



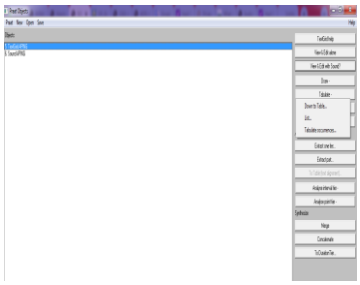
Pre-PRAAT Window

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.



Output Forman

To find the value, click on the file that has been formant, then on the right menu, select tabulate – list.



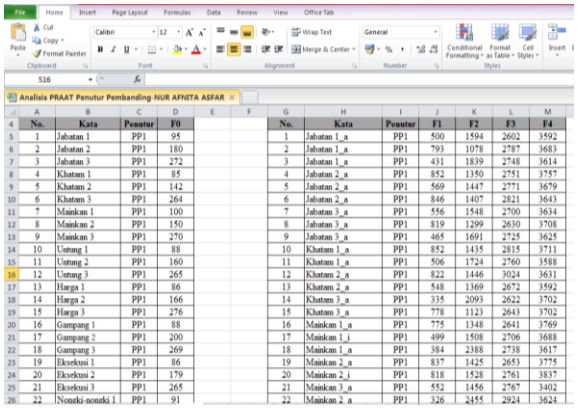
Selecting Tabulate-List Menu

File Edit Search Convert Font

```
tmin    tier    text    tmax
0       PP1     APING   62.914979
1.465467        KATA    Jabatan 2.281521
1.465467        F0      188     2.281521
1.491058        FORMANT 296;1048;2654;4370      1.543526
1.491058        VOKAL   A       1.543526
1.743102        FORMANT 678;1630;3379;3947      1.775749
1.743102        VOKAL   A       1.775749
1.890011        FORMANT 803;1968;3298;4354      1.940147
1.890011        VOKAL   A       1.940147
1.940147        FORMANT                         3.346860
3.210407        F0      192     4.444011
3.210407        KATA    Jabatan 2       4.444011
3.346860        FORMANT 250;1197;2952;4216      3.435884
3.346860        VOKAL   A       3.435884
3.498845        FORMANT 490;2198;3310;4414      3.554810
3.498845        VOKAL   A       3.554810
3.601394        FORMANT 876;1950;3410;4475      3.669019
3.601394        VOKAL   A       3.669019
5.014199        F0      208     6.146417
5.014199        KATA    Jabatan 3       6.146417
5.210070        VOKAL   A       5.287985
5.210070        FORMANT 703;2171;3349;4460      5.287985
5.299347        VOKAL   A       5.351290
5.299347        FORMANT 925;2039;3466;4558      5.351290
5.380508        VOKAL   A       5.429205
5.380508        FORMANT 950;1880;3506;4562      5.429205
7.780206        F0      209     8.423771
7.780206        KATA    Khatam 1        8.423771
7.881515        FORMANT 1231;1864;3324;4535     7.935081
7.881515        VOKAL   A1      7.935081
7.944967        FORMANT 1155;1866;3057;3382     7.949482
7.944967        VOKAL   A       7.949482
9.433679        F0      246     10.281341
9.433679        KATA    Khatam 2        10.281341
9.545365        VOKAL   A       9.678029
9.545365        FORMANT 1029;1854;3397;4524     9.678029
9.856084        VOKAL   A       9.913706
9.856084        FORMANT 1001;1871;3530;4606     9.913706
10.562256       KATA    Khatam 3        11.558496
10.562256       F0      259     11.558496
```

Forman Values in Numeric

Copy and paste the results into the Microsoft Excel, so that the tabular formant results can be processed in the Microsoft Excel as shown below:



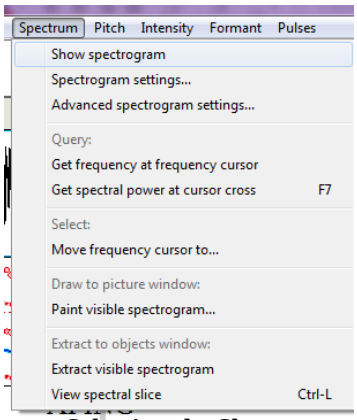
No.	Kata	Penutur	F0	No.	Kata	Penutur	F1	F2	F3	F4
1	Jabatan 1	PP1	95	1	Jabatan 1_a	PP1	500	1594	2602	3592
2	Jabatan 2	PP1	180	2	Jabatan 1_a	PP1	793	1078	2787	3683
3	Jabatan 3	PP1	272	3	Jabatan 1_a	PP1	431	1839	2748	3614
4	Khatam 1	PP1	85	4	Jabatan 2_a	PP1	852	1350	2751	3757
5	Khatam 2	PP1	142	5	Jabatan 2_a	PP1	569	1447	2771	3679
6	Khatam 3	PP1	264	6	Jabatan 2_a	PP1	846	1407	2821	3643
7	Mainkan 1	PP1	100	7	Jabatan 3_a	PP1	556	1548	2700	3634
8	Mainkan 2	PP1	150	8	Jabatan 3_a	PP1	819	1299	2630	3708
9	Mainkan 3	PP1	270	9	Jabatan 3_a	PP1	465	1691	2725	3625
10	Untung 1	PP1	88	10	Khatam 1_a	PP1	852	1435	2815	3711
11	Untung 2	PP1	160	11	Khatam 1_a	PP1	506	1724	2760	3588
12	Untung 3	PP1	265	12	Khatam 2_a	PP1	822	1446	3024	3631
13	Harga 1	PP1	86	13	Khatam 2_a	PP1	548	1369	2672	3592
14	Harga 2	PP1	166	14	Khatam 3_a	PP1	335	2093	2622	3702
15	Harga 3	PP1	276	15	Khatam 3_a	PP1	778	1123	2643	3702
16	Gampang 1	PP1	88	16	Mainkan 1_a	PP1	775	1348	2641	3769
17	Gampang 2	PP1	200	17	Mainkan 1_i	PP1	499	1508	2706	3688
18	Gampang 3	PP1	269	18	Mainkan 1_a	PP1	384	2388	2738	3617
19	Eksekusi 1	PP1	86	19	Mainkan 2_a	PP1	837	1425	2653	3775
20	Eksekusi 2	PP1	179	20	Mainkan 2_i	PP1	818	1528	2761	3837
21	Eksekusi 3	PP1	265	21	Mainkan 3_a	PP1	552	1456	2767	3402
22	Nonaki-nonaki 1	PP1	91	22	Mainkan 3_a	PP1	326	2455	2924	3624

Ms. 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.

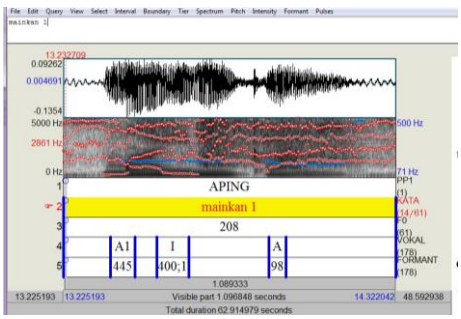
3.1.3 Spectogram

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.



Selecting the Show Spectogram Menu

Then a window will appear as below:



Then screenshot the spectrogram section to compare with the spectrogram of the evidence manuscript. Can also with way after opening the sound file, select Analyze Spectrum then select To Spectogram. Then click views. Perform this process on all evidence files with suspect voice. Then compare which suspect voices are has a Formant value close to that of a native speaker.

Sample data:

Word : Jabatan

PP (Penutur Pembanding)/ Comparative Speaker

Speaker	Pitch
Native Speaker	195 210 267
PP1	153 224 265
PP2	188 192 208
PP3	224 264 265
PP4	96 150 205
PP5	156 141 289
PP6	99 128 212
PP7	148 252 285
PP8	213 106 155
PP9	132 194 264
PP10	136 140 184

Based on the Data above, The results of each sample of digital evidence of native speakers' voices and recorded comparisons on the two assessment factors, namely Pitch and Formant, the researcher cannot distinguish between the original speaker's voice and the comparison voice. The researcher can state that the accuracy of the pitch and formant analysis through Praat is very low and cannot be used as an indicator to distinguish that a date is correct as of the original voice, as evidenced by the results of the F0 value of the original voice data and the comparison obtained. Very flexible and inconsistent. 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 each person's word pronunciation is different. There may be pitch values from several subjects that are almost the same.

Word : Jabatan

PP (Penutur Pembanding)/ Comparative Speaker

Penutur	F1	F2	F3	F4
Penutur Asli	822	1388	2718	3956
PP1	523	2053	3343	4350
PP2	296	1048	2654	4370
PP3	446	1533	2824	3566
PP4	452	1637	2925	3918
PP5	527	2631	3705	4563
PP6	444	1814	2749	3215
PP7	1079	2281	3484	4632
PP8	382	1723	2643	3183
PP9	475	1506	2444	3430
PP10	446	1533	2824	3566

Based on the F1-F4 analysis of the voice data above with PRAAT, the researcher also cannot easily distinguish the original speaker's voice from the comparison voice of 1-10 because the native speaker's voice has a formant value (F1-F4) which is not the same as the formant value of the data comparison speaker.

IV. CONCLUSION

Based on the discussion above, it can be concluded that Digital voice forensic techniques cannot validate evidence because no standard validation is determined. Digital forensic techniques can only provide sound similarity analysis of good evidence by the suspect's voice. This shows that everyone has a different pitch value because each person's word pronunciation is different. There may be pitch values from several subjects that are almost the same.

REFERENCES

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

aligarh, A., Hidayanto, C., Si, S., & Kom, M. (2016). *Implementasi Metode Forensikdengan Menggunakan Pitch , Formant, danSpectrogram untuk Analisis Kemiripan Suara Melalui Perekam Suara Telepon*

Genggam Pada Lingkungan yang Bervariasi. *5*(2).

Binyamin Widi Prasetya, Budi Santoso, and J. P. (2008). *Identifikasi Suara Pria Dan Wanita Berdasarkan Frekuensi Suara*. Universitas Kristen Duta Wacana, Yogyakarta.

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

Detik.com. (2011). *Kenapa Suara Laki-Laki Berubah Saat Puber*.

Jose R.L Batubara. (2010). *Adolescent Development (Perkembangan Remaja)*.

M. Nuh AL-Azhar. (2012). *Digital Forensic : Panduan Praktis Investigasi Komputer*. Salemba Infotek.

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

Septiyansyah, H. (2015). *Implementasi Metode Forensik Suara Pria Menggunakan Teknik Voice Recognize Untuk Analisis Kemiripan Suara Pada Media Alat Rekam Telepon Selular*. Universitas Teknologi sepuluh November, Surabaya.

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