IJFL (International Journal of Forensic Linguistic)
Vol. 3, No. 1 April 2022, Page 47-51

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

SMARTPHONE'S VOICE RECORDING USING PRAAT: FORENSIC AUDIO CASE STUDY

I Putu Kelvin Meiadiaksa Permana
Universitas Warmadewa
ptkelvin1@gmail.com

How to cite:

Permana, I. P. K. M. (2022). Smartphone's Voice Recording Using PRAAT: Forensic Audio Case Study. *IJFL (International Journal of Forensic Linguistic)*, *Vol*. 3 (1), 47-51. Doi: https://doi.org/10.22225/ijfl.3.1.4994.47-51

Abstract- Nowadays, technological developments have contributed to the sustainability of human life. Especially in the field of law or forensics which continues to evolve with the times. Sound recording evidence can show the identity of the person whose voice is recorded on the evidence by means of perform audio forensic examination for voice recognition. Therefore, this study is aimed to analyze the sounds by using the comparative method. Technically, this method is proposed to compare the sound of goods evidence with recorded sound for comparison. In this study, the analysis was limited to identify the value of voice recognition on the statistics of pitch and formant. If the results of voice recognition indicates that the voice of the conversation that is inside the evidence is identical to the voice of the perpetrator, then it can be concluded that the voice of the conversation inside the recorded evidence is the voice belonging to the perpetrator so that the voice record can be used as strong evidence in court. As conclusion, the difference between the voice of the original speaker and the voice of the comparison speaker are clearly shown in some parts of pitch and formant. On the other hand, there is no similarity of each word in value of pitch (F0) in the low, medium, and high voices of native speakers with non-native speakers. Thus, the indicators of the validity of the original sound proof (native speaker) look vague and can't even be proven.

Keywords: Evidence, Forensic-audio, PRAAT, Sound.

I. INTRODUCTION

The existence of technology in human life has facilitated almost all the human's activities in daily life. Especially, in daily communication which is not relied on face-to-face meeting but it can be facilitated by multimedia technology such as social media, cellphone, computer software, etcetera.

Therefore, today's massive technological developments have had a fairly serious impact on human life. On the other hand, it has a positive impact, but on the other hand it has a negative impact. in this case, digital crimes can be indicated by actions that violate the law both virtually and in real terms. Thus, the study of forensic-linguistic science was raised which was aimed at studying and overcoming legal problems involving linguistic evidence such as text or audio.

According to Coulthard and Johnson (2010) forensic linguistics ranges from courtroom discourse and legal language to plagiarism. Briefly, plagiarism is using another person's work for personal advantage without mentioning his/her name. Forensic linguistic experts proficient in plagiarism cases and copyright infringements provide evidence to find out which work is based upon another. In addition, Olsson (2004) defines it as an application of linguistics in the context of crime, court proceedings, or arguments in law. Thus, this has led to studies that refer to audio forensics or digital forensics. This is intended to find answers and explore legal evidence in the form of language material.

Specifically, audio forensics is one of the sciences that juxtaposes 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. (Subki et al., 2018). 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 straightforward and prone to be manipulated, either intentionally or unintentionally.

Moreover, 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. (Binyamin Widi Prasetya, dkk, 2008). In practice, digital forensic science is often helpful authorities in uncovering related crime cases the suspect concerned through the evidence that has been collected. Thus, the case this indicates that there is a possibility of future audio digital will be used more as evidence in legal cases. Legal considerations in using evidence in the form of digital files, including audio, is the ability of digital evidence in manage the impact associated with risk on the process law.

One of the risks in question is the use of witnesses who are not known with certainty the truth even though he had sworn to speak the truth. By using digital evidence that has been tested and analyzed will be able to support the discipline of action as well as accuracy of guesses and helps in accuracy decision-making. For that, it is important for the parties' relevant law enforcement agencies to understand and master digital forensic techniques considering the possibilities increasing use of sound recordings/good multimedia in terms of variety, quantity and quality.

Technically, the digital audio storage is usually used for the need for interviews or education by using which is commonly used for storage media is voice recorder or use a similar application found on cellphone. In fact, the crimes whose evidence is linguistic material such as audio and text are very volatile and really need attention.

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. Moreover, AlAzhar Nuh, (2011) mentioned 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.

Moreover, 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.

In this study, a PRAAT application was used to help with the audio comparison process from native and non-native speakers. Especially, 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.

In addition, Septiyansyah (2015) stated that 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. With PRAAT, then the user is able to see into the audio.

Therefore, based on the background above, the formulation of the problem in this study is what are the results of PRAAT software analysis for each sample of audio evidence from native speaker and non-native speakers' audio recording?

Thus, this study is aimed to explain and describe the scenario of PRAAT application of voice recognize digital forensics techniques in comparing the audio of native and non-native speakers' sound suspect on mobile phone media.

II. METHODS

This study uses descriptive-analysis method which are applied simultaneously because the data and results are displayed and described on number and graphs. However, technique of data analysis uses the Forensic audio technique. 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, 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.

III. RESULT

3.1 The Analysis of PRAAT App Process

In Praat application, the value of pitch, formant and spectrogram of each sound recording can be identified. From the results of the analysis using the PRAAT program, from native and non-native speakers, this study shows a comparison of the differences in pitch and formant 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 using recording application on smartphone.

3.1.1. **Research Data Analysis**

The analysis was carried out on the voice output of respondents in uttering ten words containing vowels that have different frequency and formant characteristics. The results of the data in the form of an analysis of the frequency and formant of sound are displayed in the form of tables and graphs. However, this study has determined the sample of data which consisted of four respondents except the native-speakers and expressed three words and contained each vowel. Thus, the results of the analysis using PRAAT produce the following details:

Table 3.1
The value of pitch (Hz)

No.	Speakers	words	Pitch (Hz)
1.	Native Speaker	Peluang	114
		Memanipulasi	116
		Keuntungan	119
2.	PP1	Peluang	116
		Memanipulasi	110
		Keuntungan	115
3.	PP2	Peluang	125
		Memanipulasi	121
		Keuntungan	133
4.	PP3	Peluang	138
		Memanipulasi	136
		Keuntungan	136
5.	…	…	…

Note:
Audio: lower tone
PP = Penutur Pembanding (Comparison speaker)/Non-Native Speaker.

Based on the table above, it can be seen that, the words used as data sources are *peluang*, *memanipulasi*, and *keuntungan*. The table also shows the value of pitch for each word from different speakers. Obviously, the value of pitch clearly differs from native and non-native speakers. In this case, there are three sample non-native speakers to be compared.

3.1.2. **The Result of Pitch and Formant Value in Comparing Speakers.**

The results of the research show that the pronunciation of the words below gives rise to various spectrograms and graphs, based on the vowels contained in them. However, the data which have been analyzed can be explained below:

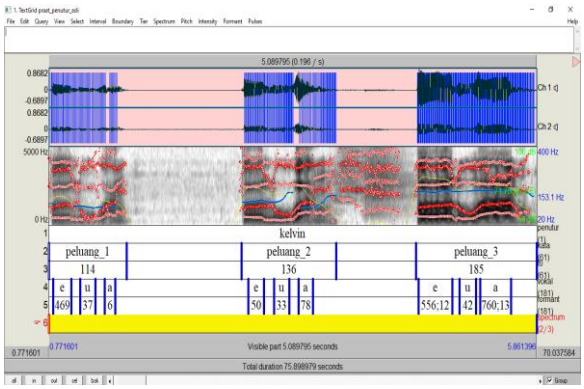
Based on the F0 analysis of the voice data with PRAAT, it can be seen the difference between the voice of the original speaker and the voice of the comparison speaker. There is no similarity F0 in the low, medium, and high pitch voices of native speakers with comparison speakers. The results of the analysis are as follows:

The word *"peluang"* sounds low, medium, and high.

a. F0 native speaker: 114; 136; 185
b. F0 Comparative speaker 1: 116; 118; 141
c. F0 Comparative speaker 2: 125; 122; 185
d. F0 Comparative speaker 3: 138; 154; 178
e. F0 Comparative speaker 4: 119; 142; 187
f. F0 Comparative speakers 5: 143; 170; 202
g. F0 Comparative speaker 6: 145; 155; 233
h. F0 Comparative speakers 7: 149; 161; 192
i. F0 Comparative speaker 8: 133; 173; 208
j. F0 Comparative speaker 9: 109; 122; 117
k. F0 Comparator speaker10: 120; 159; 169

From the results of data analysis above, it can be seen that there are no three F0 in the low, medium, and high tones of the word *"peluang"* from the same comparison speaker as the original speaker.

Figure 3.1
The spectrum of word "Peluang" native speaker



The figure above shows that the spectrogram on each intonation of words is different and the value of pitch and formant are also different.

The word *"manipulasi"* low, medium, high sounds.

F0 native speaker: 116; 124; 169

a. F0 Comparative speaker 1: 110; 118; 141
b. F0 Comparative speaker 2: 121; 131; 267
c. F0 Comparative speaker 3: 136; 154; 160
d. F0 Comparative speaker 4: 116; 145; 145
e. F0 Comparative speaker 5: 118; 141; 186
f. F0 Comparative speaker 6: 158; 171; 224
g. F0 Comparative speaker 7: 158; 176; 149
h. F0 Comparative speaker 8: 166; 196; 217
i. F0 Comparative speakers 9: 121; 121; 124
j. F0 Comparative speaker10: 126; 158; 180

From the results of value Pitch above, it can be seen that there are no three F0's in the low, medium, and high tones of the word "manipulate" from the same comparison speaker as the original speaker.

Based on the F1-F4 analysis of the voice data with PRAAT, it can be seen the difference between the voice of the original speaker and the voice of the comparison speaker. No similarity was found between F1-F4 in the low, medium, and high-pitched voices of native speakers and comparison speakers. The results of the analysis are as follows:

- F1-F4 word "aman" Native speakers
  Aman_1_a: 747; 1334; 2500; 3725
  Aman_1_a2: 666; 1407; 2340; 3533
  Aman_2_a: 889; 1494; 2572; 3788
  Aman_2_a2: 674; 1645; 2495; 3654
  Aman_3_a: 862; 1547; 2499; 372
  Aman_3_a2: 793; 1561; 2708; 3800
- F1-F4 word "aman" Comparator speaker 1
  Aman_1_a: 468; 1332; 2758; 3626
  Aman_1_a2: 625; 1302; 1657; 3513
  Aman_2_a: 679; 1295; 2666; 3646
  Aman_2_a2: 674; 1311; 2671; 3528
  Aman_3_a: 715; 1367; 2704; 3716
  Aman_3_a2: 738; 1315; 2818; 3642
- F1-F4 word "aman" Comparator speaker 2
  Aman_1_a: 627; 1207; 2583; 3727
  Aman_1_a2: 624; 1276; 2948; 3536
  Aman_2_a: 645; 1193; 2637; 3671
  Aman_2_a2: 627; 1210; 2768; 3462
  Aman_3_a: 703; 1250; 2594; 3695
  Aman_3_a2: 655; 1197; 2745; 3572
- F1-F4 word "aman" Comparator speaker 3
  Aman_1_a: 767; 1721; 3100; 4232
  Aman_1_a2: 538; 1315; 2853; 3567
  Aman_2_a: 751; 1228; 2633; 3692
  Aman_2_a2: 614; 1227; 2922; 3541
  Aman_3_a: 751; 1302; 2580; 3685
  Aman_3_a2: 767; 1397; 2913; 3524
- F1-F4 word "aman" Comparator speaker 4\
  Aman_1_a: 923; 1544; 2742; 3871
  Aman_1_a2: 705; 1488; 2495; 3318
  Aman_2_a: 858; 1477; 2887; 3931
  Aman_2_a2: 697; 1574; 2585; 3600
  Aman_3_a: 1019; 1639; 3012; 3914
  Aman_3_a2: 762; 1563; 2714; 3717
- F1-F4 word "aman" Comparator speaker 5
  Aman_1_a: 842; 1312; 2776; 3867
  Aman_1_a2: 772; 1493; 2583; 3595
  Aman_2_a: 818; 1308; 2777; 3749
  Aman_2_a2: 732; 1389; 2481; 3592
  Aman_3_a: 862; 1355; 2729; 3796
  Aman_3_a2: 675; 1455; 2416; 3417

- F1-F4 word “aman” Comparator speaker 6
  Aman_1_a: 732; 1251; 2308; 4254
  Aman_1_a2: 713; 1460; 2233; 3769
  Aman_2_a: 820; 1258; 2424; 4046
  Aman_2_a2: 820; 1486; 2562; 4125
  Aman_3_a: 855; 1603; 2941; 4413
  Aman_3_a2: 747; 1585; 2287; 3980
- F1-F4 word “aman” Comparator speaker 7
  Aman_1_a: 846; 1464; 2681; 4383
  Aman_1_a2: 739; 1376; 2483; 4315
  Aman_2_a: 906; 1497; 2868; 4208
  Aman_2_a2: 822; 1312; 2567; 4034
  Aman_3_a: 836; 1328; 2538; 3853
  Aman_3_a2: 805; 1346; 2624; 3816
- F1-F4 word “aman” Comparator speaker 8
  Aman_1_a: 823; 1452; 2680; 3790
  Aman_1_a2: 490; 1274; 2332; 3492
  Aman_2_a: 803; 1519; 2717; 3776
  Aman_2_a2: 548; 1300; 2319; 3447
  Aman_3_a: 783; 1455; 2762; 3822
  Aman_3_a2: 718; 1381; 2714; 3308
- F1-F4 word “aman” Comparator speaker 9
  Aman_1_a: 481; 1260; 2359; 3655
  Aman_1_a2: 619; 1297; 2215; 3692
  Aman_2_a: 835; 1366; 2470; 3607
  Aman_2_a2: 778; 1340; 2331; 3573
  Aman_3_a: 767; 1367; 2570; 3625
  Aman_3_a2: 705; 1286; 2300; 3552
- F1-F4 word “aman” Comparator speaker 10
  Aman_1_a: 626; 1186; 2472; 3639
  Aman_1_a2: 635; 1162; 2494; 3724
  Aman_2_a: 622; 1256; 2414; 3681
  Aman_2_a2: 674; 1276; 2334; 4201
  Aman_3_a: 689; 1202; 2525; 3828
  Aman_3_a2: 717; 1262; 2375; 4130

From the results above, it can be seen that there are no F1-F4 vocal formants in the word "Aman" from the same comparison speakers as native speakers so that the difference becomes clear.

IV. CONCLUSION

Based on the discussion above, the results showed that the words pronounced with high intonation indicate the high value of pitch and formant, while in words with low and medium intonation, the results of pitch analysis do not show significant differences. However, the difference between the voice of the original speaker and the voice of the comparison speaker are clearly shown in some parts of pitch and formant. On the other hand, there is no similarity of each word in value of pitch (F0) in the low, medium, and high voices of native speakers with non-native speakers. Thus, it can be concluded that the indicators of the validity of the original sound proof (native speaker) look vague and can't even be proven. Moreover, the value of formant (F1-F4) analysis of the voice data with PRAAT show that there is no similarity was found between F1-F4 in the low, medium, and high pitch voices of native speakers and native speaker. Therefore, in sound analysis using the PRAAT application still requires more in-depth research with the support of other relevant media to prove the validity of an audio is correct and recognized by law and forensic linguistic studies

REFERENCES

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

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

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.

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