IJFL (International Journal of Forensic Linguistic)
Vol. 3, No. 1 April 2022, Page 41-46

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

VOICE RESEMBLANCE ANALYSIS ON MOBILE PHONE'S VOICE RECORDER (AUDIO FORENSIC METHOD)

Komang Helianti
Universitas Warmadewa
mamaarya86@gmail.com

How to cite (in APA style):

Helianti, K. (2022). Voice Resemblance on Mobile Phone's Voice Recorder (Audio Forensic Method). *IJFL (International Journal of Forensic Linguistic)*, Vol. 3 (1), 41-46. Doi: https://doi.org/10.22225/ijfl.3.1.4995.41-46

Abstract- Audio forensics is one of the sciences that juxtaposes science and scientific methods in sound recording analysis to assist and support the disclosure of a crime required in the trial process. This article is intended to analyze voice resemblance towards mobile phone's voice recorder: an implementation of the forensic method using voice recognition technique, by using Forensic audio method In Audio forensics: Theory and analysis, namely pitch statistical analysis, formant 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 aims to provide an example scenario of the application of digital forensic techniques for voice 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. The results showed that most of the F1-F4 values of the comparison voices showed higher results than the original voices. However, several voices have F1-F4 results close to the original sound, namely PP2 on behalf of Dayu Inten and PP4 on behalf of Chandra.

Keywords: Forensic linguistic, Forensic audio, Formant.

I. INTRODUCTION

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 extremely easy and prone to be manipulated, either intentionally or unintentionally.

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 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.

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 is just that sound recordings cannot be used as evidence without going through a fairly long analysis process, which an expert in audio forensics carries out. 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 efforts to analyze of suitability or originality of sound 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. It is important 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 sample). Suppose the voice recognition result indicates that the sound of the unknown sample is identical to the known sample voice. In that case, 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 aims to provide 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 common 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.

III. RESULT

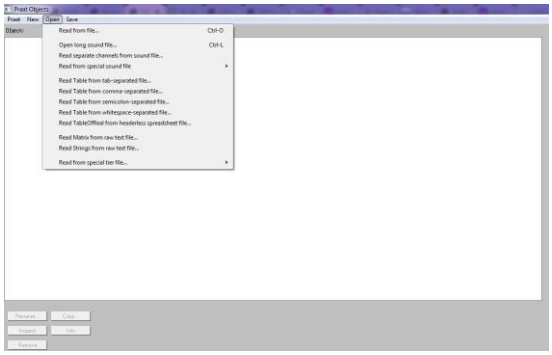
3.1 Process with Praat

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

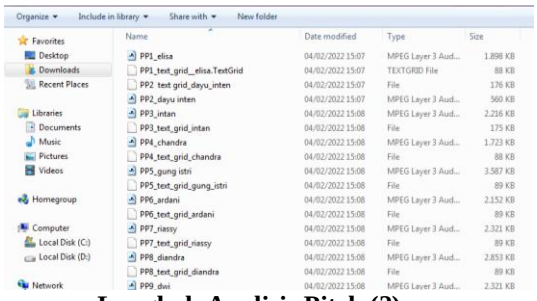
3.1.1 Pitch

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. Pitch is the frequency vibrating vocal cords (Jose R.L Batubara, 2010). 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).

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. 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.

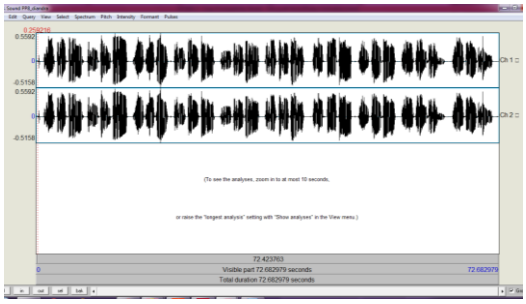


Langkah Analisis Pitch (1)



Langkah Analisis Pitch (2)

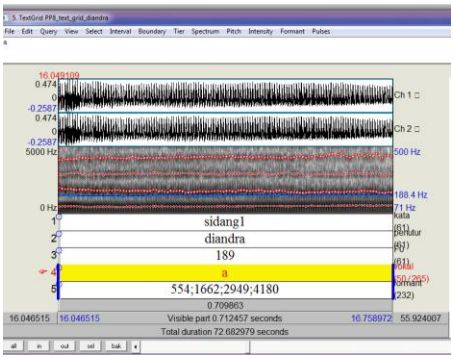
After that select *view and edit*. Then a window like the one below will appear.



View and edit menu window

Because the analysis carried out is a 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..

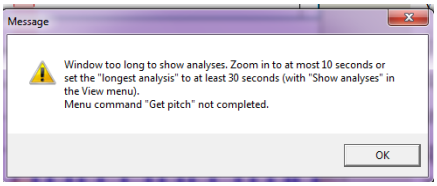


To get the right block, click the enter button in the corner left to zoom in. here's a brief explanation button located in the left corner.

Zoom Button

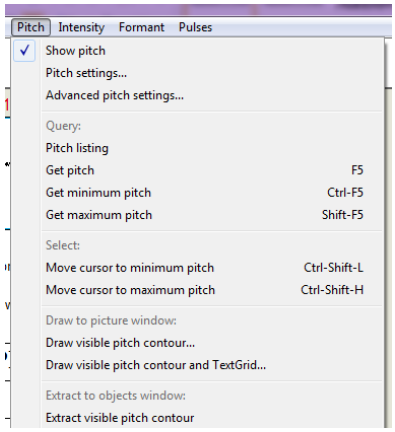
All	All selection (to see the whole chart)
in	Zoom in (to enlarge the graphic view)
out	Zoom Out (to zoom out the graphic)
sel	Selection (to view block graphs only)
back	Back (back to the previous initial view)

After the one-word graphic block, 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, there will be 30 save files 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 checking the Show pitch option. The method to find out if the option is checked is to click the pitch menu tab. If you have not already checked, click Show pitch. Because if not enabled or check show pitch, The pitch value will not come out on the menu tab Pitch.



Pitch Value Warning Window

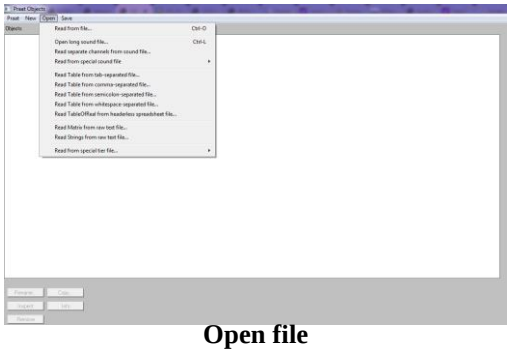
To view the minimum, maximum, and mean, select the menu tab Pitch.



Menu minimum, maximum and mean pitch

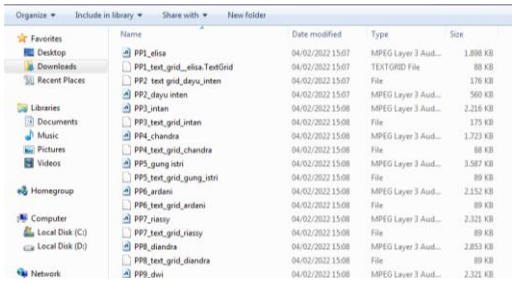
3.1.2 Formant

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



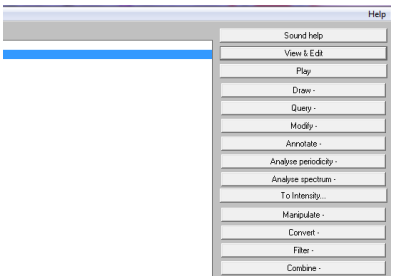
Open file

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



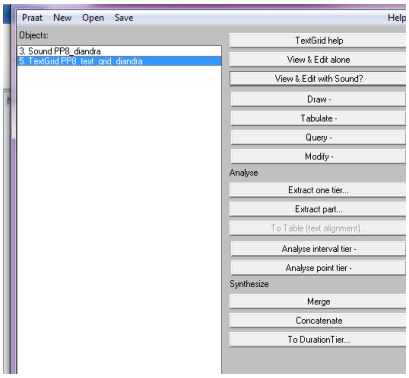
Import file dialog box

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



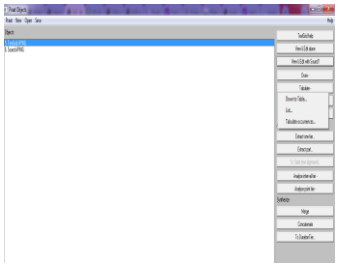
Pre-Praat Window

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.



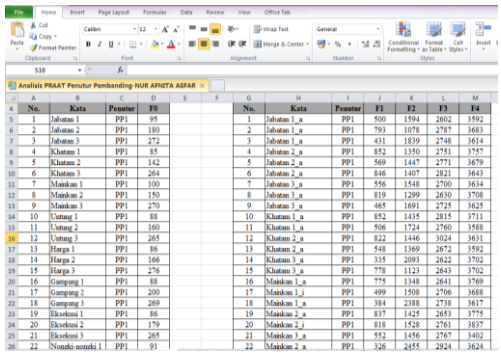
Selecting Tabulate-List Menu

File Edit Search Convert Font

tmin	tier	text	tmax	
1.217008		formant	540;1642;2787;4009	1.469431
1.217008		vokal	u1	1.469431
1.217008		penutur	diandra	2.902756
1.217008		F0	187	2.902756
1.217008		kata	usulkan1	2.902756
1.469431		vokal	s	1.620991
1.620991		vokal	u2	1.805027
1.620991		formant	347;1324;2980;3986	1.805027
1.805027		vokal	l	1.894339
1.894339		vokal	k	1.994477
1.994477		vokal	a	2.708973
1.994477		formant	522;1512;2910;4079	2.708973
2.708973		vokal	n	2.902756
3.336234		formant	467;1433;2997;4425	3.566514
3.336234		vokal	u1	3.566514
3.336234		F0	266	4.781160
3.336234		penutur	diandra	4.781160
3.336234		kata	usulkan2	4.781160
3.566514		vokal	s	3.756737
3.756737		vokal	u2	3.923762
3.756737		formant	494;1391;3208;4767	3.923762
3.923762		vokal	l	3.991036
3.991036		vokal	k	4.095427
4.095427		formant	661;1582;3014;4599	4.603461
4.095427		vokal	a	4.603461
4.603461		vokal	n	4.781160
5.208186		formant	646;1393;2467;3777	5.501412
5.208186		vokal	u1	5.501412
5.208186		kata	usulkan3	7.375576

Forman Values in Numeric

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



No.	Kata	Penutur	F0
1	Jabatan 1	PP1	95
2	Jabatan 2	PP1	180
3	Jabatan 3	PP1	272
4	Khatam 1	PP1	85
5	Khatam 2	PP1	142
6	Khatam 3	PP1	264
7	Maiskan 1	PP1	100
8	Maiskan 2	PP1	150
9	Maiskan 3	PP1	270
10	Untung 1	PP1	88
11	Untung 2	PP1	160
12	Untung 3	PP1	265
13	Harga 1	PP1	86
14	Harga 2	PP1	166
15	Harga 3	PP1	276
16	Gampang 1	PP1	88
17	Gampang 2	PP1	200
18	Gampang 3	PP1	269
19	Eksekusi 1	PP1	86
20	Eksekusi 2	PP1	179
21	Eksekusi 3	PP1	265
22	Nonaki-nonaki 1	PP1	91

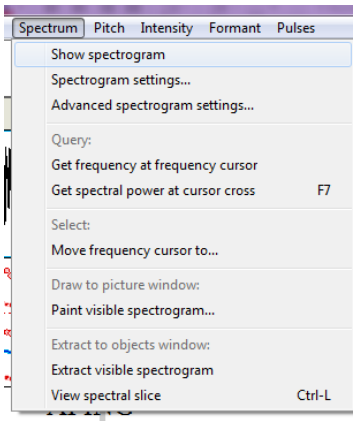
No.	Kata	Penutur	F1	F2	F3	F4
1	Jabatan 1_a	PP1	500	1594	2602	3592
2	Jabatan 1_a	PP1	793	1078	2787	3683
3	Jabatan 1_a	PP1	431	1839	2748	3614
4	Jabatan 2_a	PP1	852	1350	2751	3757
5	Jabatan 2_a	PP1	569	1447	2771	3679
6	Jabatan 2_a	PP1	846	1407	2821	3643
7	Jabatan 3_a	PP1	556	1548	2700	3634
8	Jabatan 3_a	PP1	819	1299	2630	3708
9	Jabatan 3_a	PP1	465	1691	2725	3625
10	Khatam 1_a	PP1	852	1435	2815	3711
11	Khatam 1_a	PP1	506	1724	2760	3588
12	Khatam 2_a	PP1	822	1446	3024	3631
13	Khatam 2_a	PP1	548	1369	2672	3592
14	Khatam 3_a	PP1	335	2093	2622	3702
15	Khatam 3_a	PP1	778	1123	2643	3702
16	Maiskan 1_a	PP1	775	1348	2641	3769
17	Maiskan 1_i	PP1	499	1508	2706	3688
18	Maiskan 1_a	PP1	384	2388	2738	3617
19	Maiskan 2_a	PP1	837	1425	2653	3775
20	Maiskan 2_i	PP1	818	1528	2761	3817
21	Maiskan 3_a	PP1	552	1456	2767	3402
22	Maiskan 2_a	PP1	326	2455	2924	3624

Ms. Excel

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

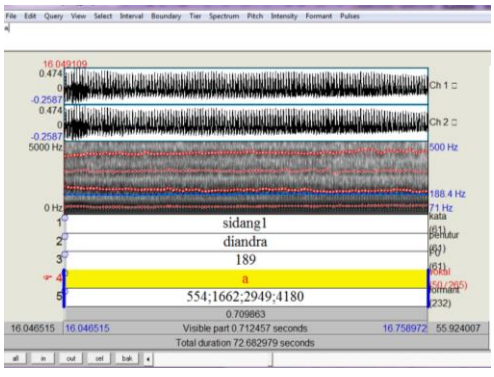
3.1.3 Spectogram

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



Selecting the Show

Then a window will appear as below:



Then screenshot the spectrogram section to compare with the Spectrogram of the evidence manuscript. With the way, after opening the sound file, select Analyze Spectrum, then select To Spectrogram. 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:

Based on the results of the F0 analysis, the researcher can distinguish between the voice of the original speaker and the voice of the comparison. This is because all F0 results in the comparison sound differ from the original sound. For example, the word 'usulkan' the result of F0 (for a medium tone) in a native speaker, which is 198, while PP1 = 259; PP2=249 ; PP3=371 ; PP4=267 ; PP5=279 ; PP6=251 ; PP7=251; PP8=264 ; PP9=264 ; PP10=261. From this value, it is clear that the F0 value of native speakers is lower than the other comparison speakers. Based on the results of the F1-F4 analysis on the PRAAT voice data, the researcher can distinguish the original speaker's voice from the comparison voice 1-10. One example can be seen from the results of F1-F4 when the original speaker and comparison pronounce the vowel 'e' in the word 'revisi' in a medium tone. The results showed that most of the F1-F4 values of the comparison voices showed higher results than the original voices. However, several voices have F1-F4 results close to the original sound, namely PP2 on behalf of Dayu Inten and PP4 on behalf of Chandra. Where the results of F1-F4 native speakers are: 506, 2019, 2773, 3922; PP2 on behalf of Dayu Inten the results are 586, 2068, 2511, 4109; and PP4 on behalf of Chandra, namely 503, 2100, 2691, 4243.

IV. CONCLUSION

Based on the discussion above, it can be concluded that Digital voice forensic techniques can not validate evidence because no standard validation is determined. Digital forensic techniques can only provide results from the sound similarity analysis of good evidence by the suspect's voice. Therefore, based on the results of the data analysis above, it can be stated that it is clear that the F0 value of native speakers is lower than the other comparison speakers. The results showed that most of the F1-F4 values of the comparison voices showed higher results than the original voices. However, several voices have F1-F4 results close to the original sound, namely PP2 on behalf of Dayu Inten and PP4 on behalf of Chandra.

REFERENCES

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