Kurdish Studies
May 2024
Volume: 12, No: 4, pp. 720-732
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i4.3050

A Phonetic Forensic Analysis of Imran Khan's Speeches

Muhammad Atif[1]*, Dr. Aniqa Rashid[2], Muhammad Farukh Arslan[3], Faiz Ullah[4], Mahnoor Amjad[5], Hira Haroon[6]

[1] *Email: atif.ali35@gmail.com M.Phil scholar, NUML (Faisalabad Campus)
[2] Email: anrashid@numl.edu.pk Assistant Professor, NUML (Faisalabad Campus)
[3] Lecturer, NUML (Faisalabad Campus) Email: Farukhgill99@gmail.com
[4] Email: faizp057@gmail.com M.Phil scholar, NUML (Faisalabad Campus)
[5] Email: mahnoor.amjad2508@gmail.com Visitng Lecturer, University of Okara
[6] Email: hiraharoon44@gmail.com MPhil, Department of Applied Linguistics, GC, Faisalabad

***Corresponding Author:** Muhammad Atif
*Email: atif.ali35@gmail.com M.Phil scholar, NUML (Faisalabad Campus)

Abstract
The objective of this research was to analyze the speeches made by Al Tools and Imran Khan. Praat played a crucial role in conducting this analysis. Nowadays, there are numerous fake videos and audios associated with specific individuals. For instance, speeches made by Al Tools, such as Imran Khan's speech after being imprisoned, were released. The researcher obtained these videos from online sources and compared them. The first video focused on Israel's attack on Gaza, which was a speech generated by artificial intelligence (AIGS), while the second speech addressed the Kashmir issues at the United Nations (UNS). Certain words, namely "people," "world," "devised," and "when," were selected for study. The first two words were present in both videos, the third word was only in the AIGS video, and the fourth word was from the UNS video. These words were analyzed based on intensity, pitch, pulse, spectrogram, and formants. The analysis involved comparing the graphs and values of these parameters. The study yielded significant results in analyzing both speeches. The sound features of AIGS consistently exhibited low values, whereas the UNS features displayed high values. In terms of intensity, the pattern was reversed, with AIGS exhibiting higher values compared to UNS for all the aforementioned words. Therefore, it can be concluded that all praat features consistently have higher values, except for intensity, when comparing AIGS and UNS. Consequently, praat can effectively distinguish between human-based and Al-based speeches. This study holds important implications for future researchers. The findings can be applied to various fields related to forensic linguistics, such as analyzing threatening calls, voice detection, judicial procedures, and law enforcement agencies like police departments in apprehending criminals.

Keywords: Phonetic, Forensic, Praat, Artificial Intelligence Generated Speech (AIGS), United Nations Speech (UNS)

1. INTRODUCTION

This research focuses on the field of phonetics, which involves the study of human speech sounds. According to Haralambous (2024), phonetics is concerned with the sounds produced by humans for communication purposes. Acoustic phonetics plays a crucial role in analyzing these speech sounds, as it deals with their physical properties. Acoustic Phonetics (n.d.) defines it as the study of the physical characteristics of speech, aiming to analyze sound wave signals that occur during speech, such as varying frequencies, amplitudes, and durations. In Ladefoged's (1996) description, he emphasizes the importance of understanding concepts related to acoustic analysis using modern laboratory techniques, including the resonances of the vocal tract and the relationship between formants and different cavities.

1.1. BACKGROUND OF THE STUDY

The parameters which are very important to investigate the speech can be many in numbers. Intensity is one of the parameters which shows the energy level of a particular sound. Intensity level is measured in decibels and its unit of measurement is dB. Intensity varies as time passes. Different speeches have different intensities which is a very important clue for the analysts of researchers. The next parameter for the evaluation of a speech is the Spectrogram. The spectrogram of a particular sound or speech is different from the other sound or speech. In this research, this parameter also worked effectively - to differentiate AIGS from UNS by the comparison of the two. The

spectrogram shows different graphical views for different speakers even if they pronounced the same sound or utter the same word.

The other parameter which has been explored in this study is pitch. Pitch shows the frequency of lower level and upper level for any sound analyzed by praat. The number of waves passing through a particular area in a second talk about the pitch of the sound one sound produced by AIGS may have different values of pitch as compared to the values extracted by praat in the case of UNS. The pulse of the sound is a very important factor or parameter in conducting the do the analysis. In the AIGs and UNS, the pulses can differentiate both of the speech. The distance between two dotted lines may be helpful to check the pattern of both speeches.

Lastly, the parameters wed to differentiate human speech from Al speech as any two different speeches were formants. Formants showed the resonance at different levels of the sound pattern Four formants were used to find the differences between the two speeches of AIGS and UNS.

Forensic linguistics gave the researchers a lot of techniques and tools for the analysis of speech or writing. It equipped the researchers to evaluate the action of forgery in speech or written forms of language. One tool used to analyse the sounds is known as praat. The researcher adopted this to analyse the two speeches related to Imran Khan. One of the speeches was prepared by Al tool, available on the internet and shortly named AIGS. The other speech used for the comparison was named UNS in a short form and was delivered by him at the United Nations. Pareniuk (2021) talked about the software known as Praat used for the analysis of human speech comprehensively. It can be used for multiple tasks and to study the sounds produced by human beings or any other source.

Seong (2022) described that more indepth analysis could be done in object window of the praat for example, the analysis of spectogram, pitch, formuotes pulses and intensity of the speech Sultana (2023) has investigated with the help of praat the pronunciation of the same words phrases and sentences pronounced by two non-native speakers. The researcher got the idea to research because Al tools were launched to produce human-like speech. The research conducted the research with the help of AIGS and UNS speech comparison on multiple parameters given in the soft wake known as praat.

To conduct the research, the two words for example, 'people', and 'world' were used from both the speeches to study and all the parameters were used to analyse these words. But two words, which were not common in both the speeches, for example, 'devised' from AIGS and 'When' from UNS were also taken to study different words through praat. An in-depth analysis of intensity, spectrogram, pitch, pulses and formants was done with the help of taking graphs and values. All the parameters other than intensity gave higher values for UNS and lesser to the AIGS. The differences in the values were notable and could easily be understood. From all the results, it was concluded that a speech produced by an AI tool always produced high-intensity and low-pitch, pulse and formants of the sounds.

1.2. RESEARCH QUESTIONS

1. How does Imran Khan's AIGS differ from the UNS?
2. What factors influence the variation between Imran Khan's AIGS and the UNS?

1.3. PROBLEM STATEMENT

It was found that some people made AI speeches to deceive others or to achieve some other political targets. So, the research was conducted because the researcher wanted to apply the techniques and ways used in forensic linguistics for analysing Imran Khan's AIGS with the help of Praat.

1.4. RESEARCH OBJECTIVES

The objectives of the study are:

- To find the differences between Imran Khan's AIGS and UNS
- To explore the implications of the differences observed between Imran Khan's AIGS and UNS

1.5. SIGNIFICANCE OF THE STUDY

The study of forensic investigation of the speeches with the help of Praat was very significant because, with the help of Praat, each kind of sound produced by Al or any other source can be differentiated based on similarities and differences. The results of this study help check voice changers on telephonic calls, and any types of threats given by a person to someone. The results can be applied in the judiciary and the law-implementing agencies.

2. LITERATURE REVIEW

In this section, the past works in this field were reviewed. Pareniuk (2021) describes that praat is software, that can be used for comprehensive speech analysis. The Software can perform plenty of jobs and we can compare the sounds based on criteria built into this software. Its functions contain spectrograms, formant analysis, pitch analysis,

jitter, shimmer, intensity analysis and voice breaks. We can get high-quality images of sounds about different features performed by Praat.

Human speech is very important for analysts to check the psychological condition or state of mind. Emotional, physical, unconscious or intentional Voice shows much about the internal behaviour of the internal state of speakers (Sondhi et al., 2016).

The real-life natural state of restlessness and the voice of the criminal, who is suspected is recorded when the interrogation starts at the police station. This will be analysed through a software known as praat. The important tools in this regard are fundamental frequency (Fo) and formant frequencies F1 and F_2 and they help recognize suspects. Similarly, (Sondhi et al., 2016) points out pitch contours and they show a considerable rise in the pitch when crime answering specific questions. So these frequencies of, and formant are very much linked with stress.

The Praat is a very versatile software and has also become standard in the scientific community relative to language. With the help of Easy Align used in praat, an automatic speech segmentation is developed and it requires text which means separated sentences (Pleva et al. 2015) In many fields of speech-language, doctor's, engineering, pathology and linguistics praat is & playing a very important role to help the people working in these settings of day to day life. Seong (2022) discusses more in-depth analysis will be done in the advanced menu of the object window. In this window, spectrum analysis, formant analysis, pitch anlalysis, pulses and intensity analyses can be checked. They are very useful in the clinical field because, with the help of quality measurement, the diagnosis of the patient can also be fulfilled.

Ko (2020) describes that time. Frequency analysis software (TF32), MDVP and praat are being used for voice analysis but they show a lot of difference in results because of the differences in the use of algorithms in their programming. So before using a particulars analyzer, an analyst must have to take the knowledge of the specific software working.

Sultana (2023) researched to investigate the pronunciations of the same words, phrases and sentences uttered by two non-natives with the help of praat. The study shows that there is a difference in the speaking of both the teachers speakers. The same word has been produced by different. pronunciation or sound by the creation of several phonemes in the speech which they produced. With the help of Praat, the study observed many differences while they were using the same words in their speech.

IAFPA stands for International Association for Forensic Phonetics and Acoustics. This is an important task to analyse speaker identification in the field of forensic phonetics. and This an Acoustics (Coulthard Association & Johnson, 2010.). annually conducts meetings to represent the journal International Journal of Speech and Language and speech law and Speech. Coulthard and Johnson (2010) describe the steps and points for forensic-speaker identification. Coulbard and Johnson (2010) give an example of the Yorkshire Ripper. case. The two stages were discussed in that case. In the first case, the speaker profile was given where the caller claimed that he was a Yorkshire Ripper and he murdered 13 women in the cities of Leed, Huddersfield and Bradford. After that, it was known to the investigator that had been made by the heaxer, in the biological investigation, and found through DNA. A speaker and a call made in 1970 were compared and then the decision was taken that Both the speaker and the caller were the same individual.

Benus (2021) says that the crucial part of the analysis in speech analysis is the notion of prominence. Different words in the phrases stay with the comparison to the other words in that phrase and what are the communicative meanings of this pattern highlight? The fundamental frequency is marked as for and it corresponds with the rate of the vocal cords vibration. By varying the tension of vocal cords and muscles Vibrations are controlled. (Benus, 2021).

Boersma (2002) suggests that the technology in the field of phonetics and phonology has changed the way of deserving and analysis of the data many linguists use the recording in interviews when they conduct research and after completing interviews, they use. IPA symbols to annotate the speech. But in phonetics and psychology different parameters like tempo. or intensity is often changed. The introduction of computers in the fields of linguistic research related to phonetics drastically changed the ways people use these for speech recording and analyzing through praat software.

Albin (2014) describes that many researchers use R programming language to visualize the statistical modelling of the data about acoustic phonetics R's digital processing has a lot of limitations. Nowadays praate has been designed to overcome the inefficiency of the R functions. The analysis of spectrographic consists of getting a narrow-hand spectrogram from the digitalized. voice sample by the 2 independent observers. The final result of the acoustic parameters like shimmer, jitter and harmonic to the noise ratios with the fundamental frequency were obtained.

Barreda (2021) came up with the idea of fast-tracking formant implemented in praat and gave the explanations in this way:

Fast Track is a formant tracker implemented in praat that attempts to automatically select the best analysis from a set of candidates. The best track is selected by modelling smooth formant contours across the entirety of the sound, providing the researcher with rich information about static and dynamic formant

properties. Fast Track returns text files containing acoustic information (formant frequencies, formant bandwidths, fundamental frequency, etc.) sampled every 2 ms, generates images showing the winning analysis and comparing alternate analyses and creates log files detailing analysis information for each file.

3. METHODOLOGY

The research methodology of this paper is composed of the following sections which discuss the ways and methods of collection of the data, tools and analysis in a detailed manner to show the proper road map for conducting the research that will provide ease for the readers.

This chapter talks about the methodology used in this research article. the author explains here briefly the purpose of the research, the nature of the research, the framework, sampling and its size, data collection and its analysis. Let us consider about different steps used by the writer in this particular chapter of methodology.

3.1. PURPOSE OF THE STUDY

The purpose of the study is to analyse the AI-generated speech by Imran Khan after getting jailed. With the help of praat software, the speech has been analysed by studying the different perspectives like spectrogram, intensity, frequency and formants. The speech would be analysed by comparing it with another human-based speech by him at the UN.

3.2. RESEARCH FRAMEWORK

The research was conducted with the help of an analysis of speech delivered by Imran Khan by the praat software. The praat is a very strong tool to analyse speech whether it is delivered by human or AI-generated tools. The AI-generated speech was compared with another speech of human-based.

The software is equipped with different tools for the analysis of speech at different levels. Spectrogram, frequency, formants and intensity were the tools to be involved in the analysis of the speech. The praat software gave pictorial evidence for each level of analysis. This made the readers' view very broad to understand how the speeches are different from each other and which part was most deviated from each other if we compare all of the factors while analyzing.

The praat can perform a lot of duties for us. For the analysis of sound produced by a human and by Al tool of any other way can easily be judged with this strong software. The multiple features of the Praat were used to analyse the sounds produced by both of the speech. All of the features of Praat are being discussed here as part of the research methodology.

Intensity is the important parameters that differentiate the two speeches. The researcher evaluated the words from both of the speeches for example "people", world, devised and when. The first two words were comen for both of the speeches while the third and fourth words are unique in both of the speechs whether it is AIGS or UNS. The upper, lowes and intensity of the two speeches were taken to be analysed.

The next parameter which was adopted to evaluate the above-mentioned words was the spectrogram. This tool gave a clear image of the word in graphical. The researchers interpret the images for the differences of similarities. The other feature of praat was pitch which helped a lot in analysing the speeches of AIGS and UNS respectively. UNS speech has greater values as compared to the values of AIGS.

The pulses were also very important phenomenon of the analysis done by the praat. This was the representation of the graph as well as the representation of the values of both of the speeches. The shows more width for Al for human speech. graph with speech and less width

Lastly, formants were adopted to analyse the speeches. The formants were of four kinds for example F1, F2, F3 and F4. In addition to the formants, bandwidth was also taken to analyse the speeches.

3.3. SAMPLE TYPE

The author took Imran Khan's AI-generated speech after he got into jail. Only one speech was selected to be analysed and after that another speech for reference was also selected which was delivered by Imran Khan himself and the UN assembly.

3.4. SAMPLING TECHNIQUE

In this study, the AI speech related to Imran Khan was taken from YouTube and extracted the sound from into the form of MP3 which was supported by the Praat software. Another speech by Imran Khan was also selected which was delivered at the UN assembly it was totally human-based and the same process was also done to make the sound compatible with Praat. This speech was used as a reference for comparing AI speech with this. The reference speech was also taken from YouTube videos and converted into MP3 sounds to make it compatible with the Praat software.

Two similar words 'people', 'world' and two different words 'devised', 'when' were taken from both of the speeches to be analysed with the help of parameters given in the software.

3.5. DATA COLLECTION TOOL

The data collection tool in this study is YouTube videos and after getting the speech the sound from the MP4 type file was extracted with the help of online software for converting the file into MP3. Both the files of sound went through the same process to get compatibility.

3.6. LIMITATIONS

The study only discusses the comparison of the spectrogram, intensity, frequency and formants of the two speeches. The praat software could do more than only the comparison at four levels.

4. ANALYSIS

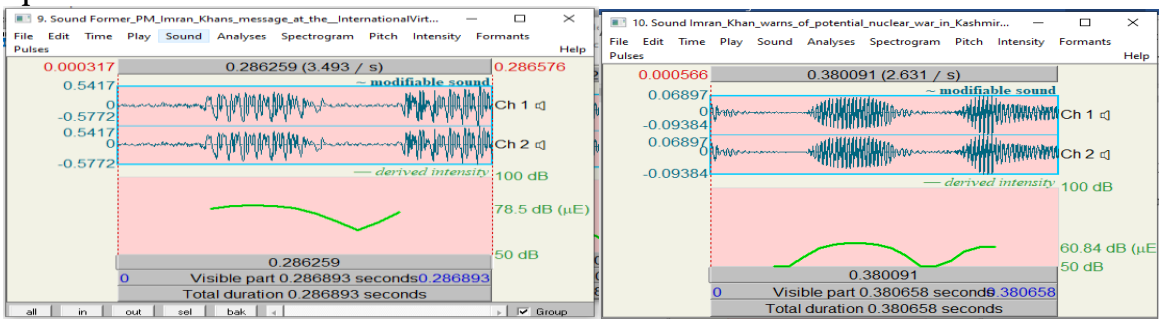
This section talks about the speech analysis delivered by Imran Khan after going to jail. The speech was generated by artificial intelligence about Israel brutality in Gaza. The speech was prepared for an international conference about the conflict. In this specific chapter, our task is to evaluate the different aspects of speech based on forensic linguistics by using a speech analysis to tool known as praat. The research tried to find whether the speech was given by a human being or was generated through some AI software. For this, the researcher also took another speech by Imran Khan delivered at the UN about Kashmir issues a few years ago. With the comparison between the two speeches' parameters like waveform, amplitude, intensity, frequency, beep etc we can evaluate whether the speaker is a human being or not. Now all the aspects of speech are being compared one by one to differentiate that the speech about Israel was made by an AI tool and not by a human.

4.1. Intensity

Intensity is a significant factor in differentiating human speech from machine-based speech. These are two speeches here which would be compared through praat.

In the beginning the words which are in both of the files occurred are discussed to check the differences in the intensities of both AIGS and UNS.

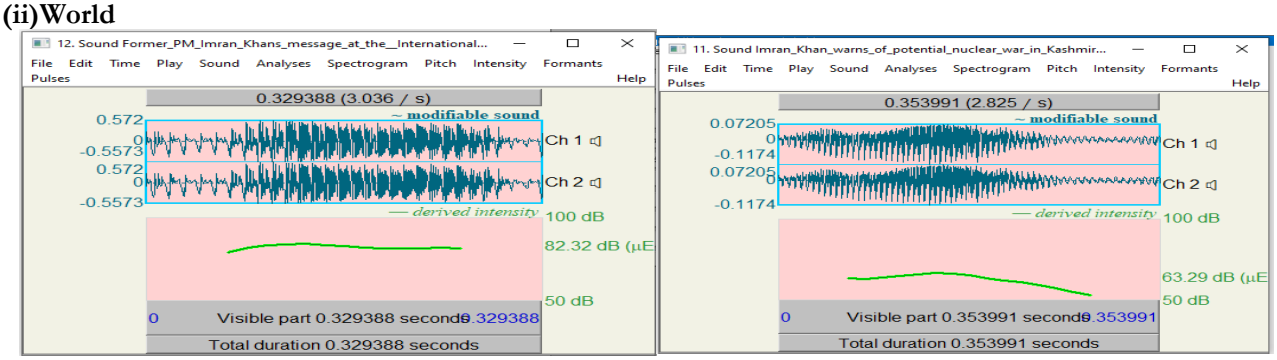
(i) People



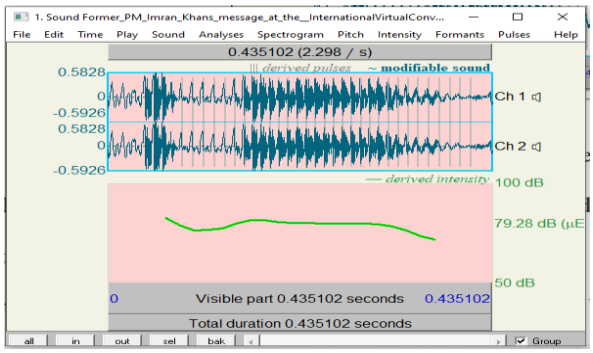
The intensity-time graph can be seen from the picture. The green line shows the variation of the left side picture which is about AIGS. The values of the intensity are 65. 5 dB for a minimum level and 81.5dB for a maximum value. The average value of the intensity is 78.5 dB.

Now the UNS is being analysed on the same grounds. The min intensity is 39.7dB While the max. value of the intensity in this situation is 64.8dB and the average value of intensity is 60.84 dB. There is a huge difference between the intensities of the same word uttered by humans and AI tools.

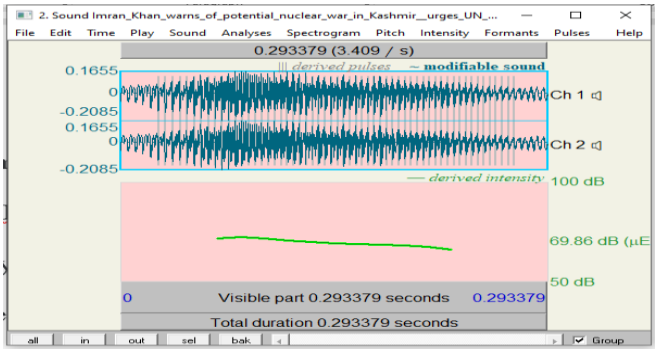
(ii)World



This word has been uttered in AIGS and shows min. intensity as 79.12dB. The higher intensity is 84.04 dB in this speech if the word 'world' is taken to be studied. The average of both the intensity is 82.32dB.
Now the UNS has been evaluated with the praat. The min. intensity level of this word is 52.98dB and the max. is 66.52dB. The software shows the average of min. and max intensities as of 63.29 dB.
A huge difference can also be this analysis of the same word by praat. The human-based and AI-based speeches show very significant differentiations in this regard to know about the originality of the speeches.
Now the different words for example, the word 'devised' from the AIGS and 'when' from the UNS have been taken to be analysed.



The varying intensity-time graph can be seen in the picture of this speech. The Chunk of the speech has been taken and the graph shows 71.79 dB intensity for the minimum level of this particular word and if the maximum intensity is taken into consideration, the value would be 82.47 dB. Intensity is related to the loudness of the sound. AI-generated speech, all the graphs represent, is more loaded as compared to human-based speech.
Let us see Imran Khan's UN speech on the issues of the lockdown in Kashmir.

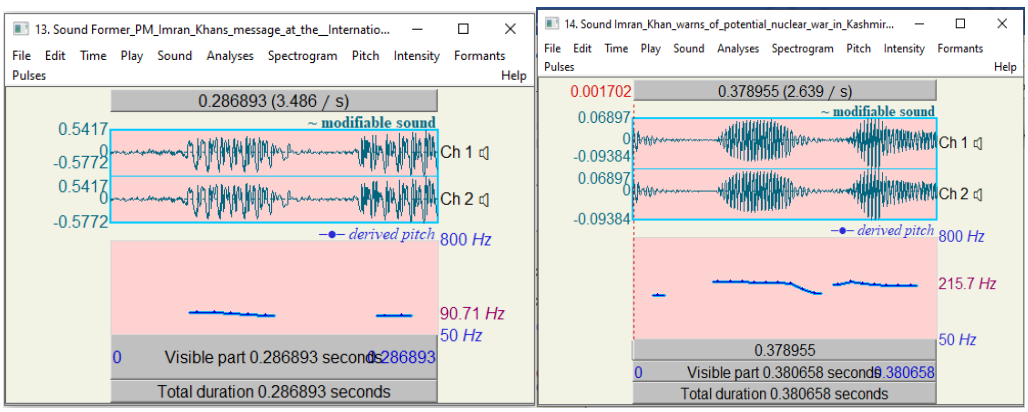


In this case, the minimum intensity is 66.21 dB. The maximum - intensity in this case is 72.23 dB. Both the diagrams have shown significant results for the researcher. Specially. If the upper level of the intensity is taken to be analysed, the AI-generated speech shows a huge difference as compared to human-based Speech. The lower level intensity also has some notable differences The AI speech is more. louder than H.B.'s speech.

4.2. Pitch

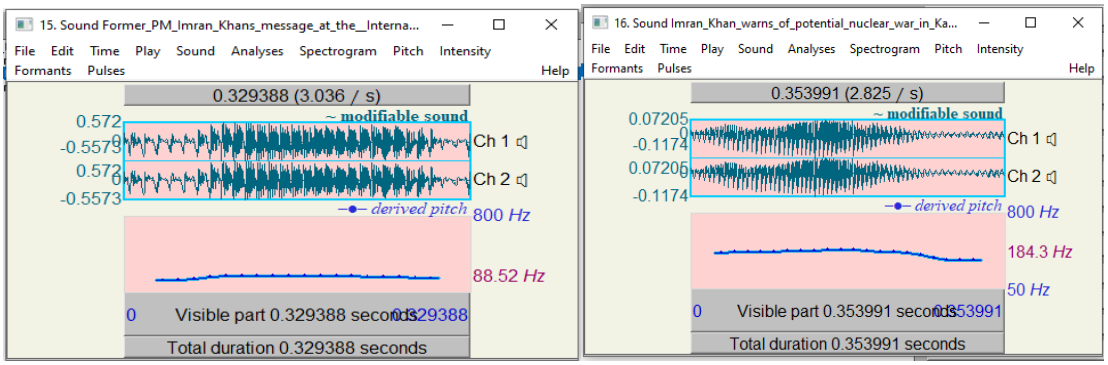
The researcher's task is to check the pitch of both speeches and compare them based on similarities and differences.

(i) People



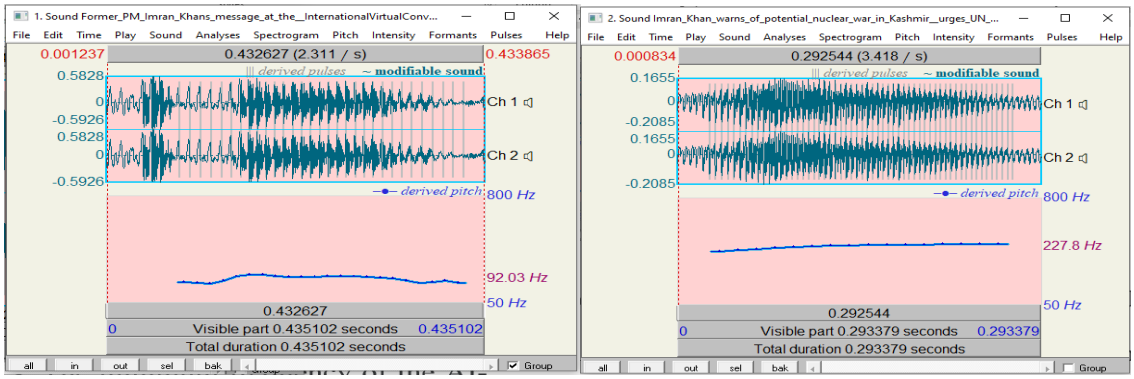
The min. pitch of this word can be gotten as 86.85Hz for AIGS by using this software and the value of max. pitch in this regard is 85.99 Hz and that of the average value of both of them is 90.7/Hz. These values are being compared with the values taken from the analysis of UNS speech. Here, are the values for min. and max. pitches are 163.91 and 236 Hz respectively. The average mean value can be seen from the praat window which is 215.7 Hz.

(ii)World



The word 'world' shows a 78.31 Hz pitch value when it is spoken through the Al tool and similarly, the max. value can be seen in the intensity tab which is 93.13 Hz. If the mean value has been taken under consideration, it will show 88.52 Hz for that. In the second picture of praat the min. value for the above-mentioned word is 141.09 Hz and for the max. value it is 207. 21Hz. The mean valve has been. shown on the praat window for analysis and that is 184.3 Hz

From both of the above examples, it has been clear to the reader that huge differences are there in the pitch values of both kinds of speeches. UNS-based words show more frequencies. as compared to AIGS-based words. These pitch values are very important clues to check whether the sound is produced by a human or not.



The minimum frequency of the AI-generated speech for the word is 81.13 Hz while the upper-level frequency of the speech delivered at the United Nations assembly is 198 Hz if the lower frequency of this chunk would be taken. There is a large difference between the lower frequencies of both of the speeches. It is a hint that both of the speeches are not delivered by the same person. In the case of upper frequencies, the AI generated speech shows 103.00 and UN speech h shows 240.85 Hz frequency. Both of these have a lot differences.

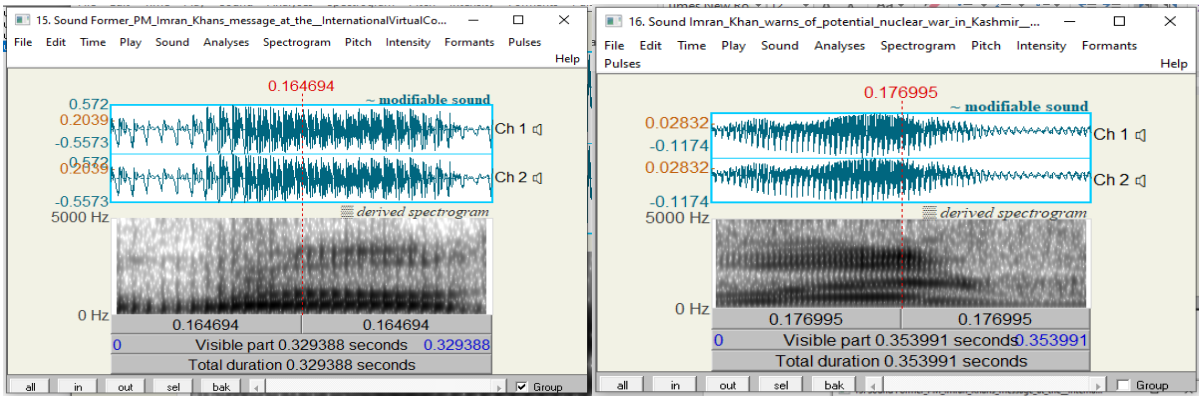
The Results of many experiments show that most of the time the human speech frequency is above 165 if we talk about the upper-frequency range. It can be said that first speech is not human-made and is software. is the product some.

4.3. Spectrogram

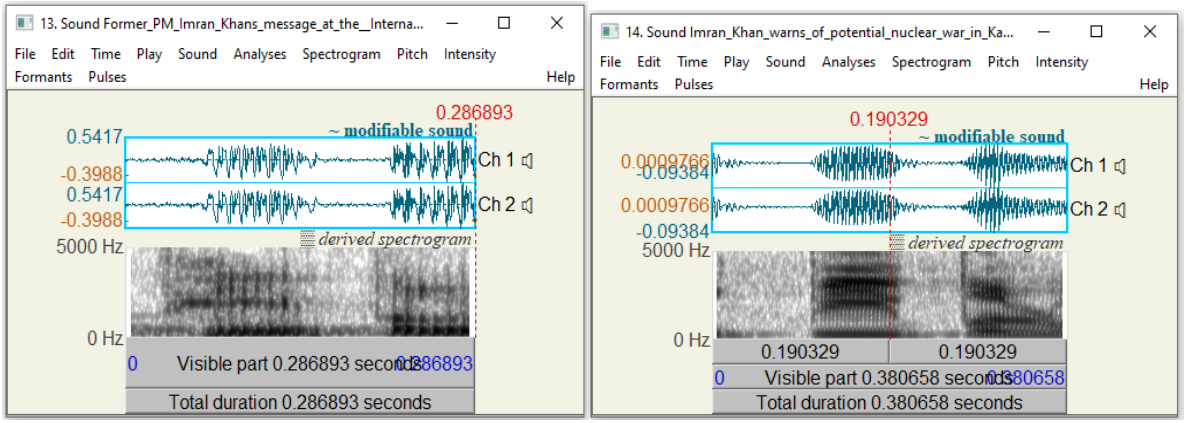
(i) People

The left side picture shows the spectrum of the people produced by AIGS. From the picture. it can be seen that the high energy area and low energy area cannot be separated most of the fast of the window. This is a mixture of black and white areas. We can see three layers easily. The fourth layer is not very clear in this picture. The second picture is actually about UNS-based word presentation of the same word people. In the first part of the word people in the dark areas are very much clear which shows high energy. All the four formants can easily be seen in this part of the picture.

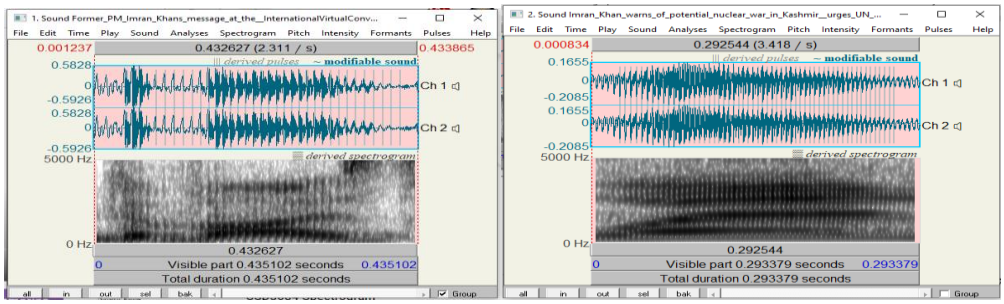
(ii)World



The AIGS spectrogram of the word the world' has been mentioned left side of the graph. Different formats can be seen in the plot of the window, but the third and fourth formants are dim which shows less energy in that area. On the other hand, the UNS Spectrogram on the right side of the picture shows more dark areas of the formants. The third and fourth formants are mixed because of the darkness in that area. The darkness also indicated the high energy of the formants of the given word the 'world'.



Spectograms of both the speeches can be seen in the figures below.

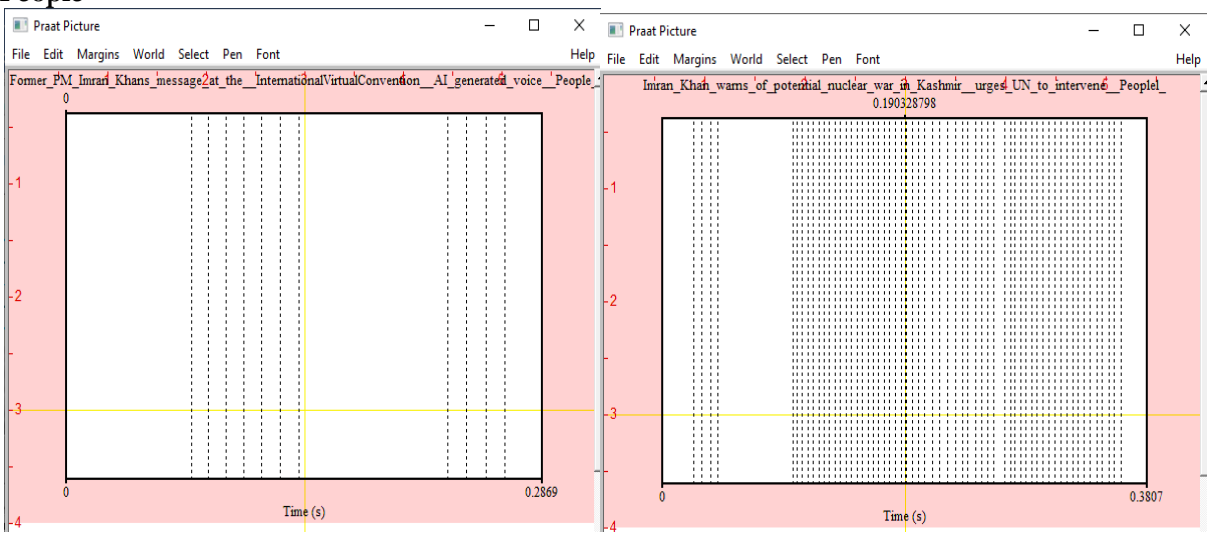


Let us first discuss the spectrogram of AI-generated speech. The dark and white areas on the picture 1 can be seen easily. There are broken parts of dark and white areas show low amplitude of the speech. The picture shows the sound becomes abruptly high and not a particular can be seen. It does not show all the formants because the third formant is missing. If we see closely a very thin line can be seen.

Now, let us talk about the second picture of the spectrogram. In the spectrum, it can be seen the dark parts of the picture have specific patterns. All the layers of the frequencies are there to show the human voice has a regular pattern. The picture is also a white part but not in disordered form.

4.4. Pulses

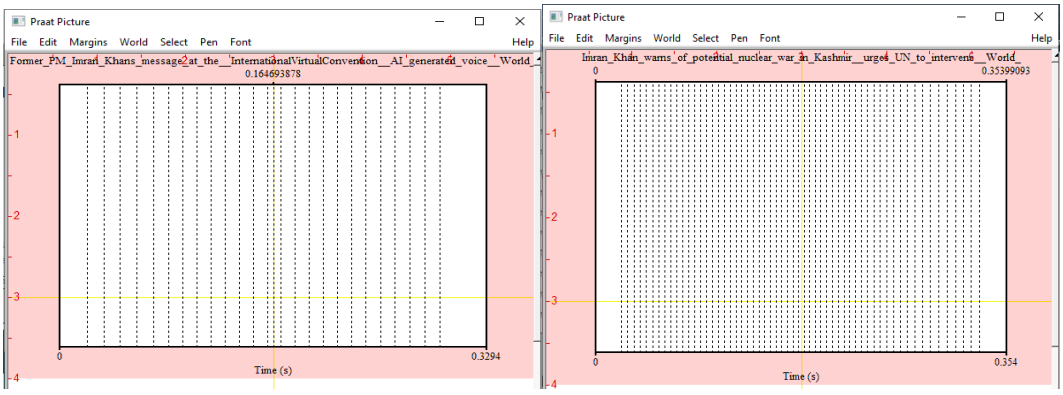
(i) People



The above picture shows the pulled time graph. The left side picture shows the pulses for the AIGS. In this pictorial representation, the dotted lines are far from each other and the gap between the pulses can be seen clearly. If the comparison is made between the two speeches, the right side of the picture has a narrow width shown among the lines. Pulse are very close to each other in the case of UNS. This analysis is also more relevant to the graphical representations.

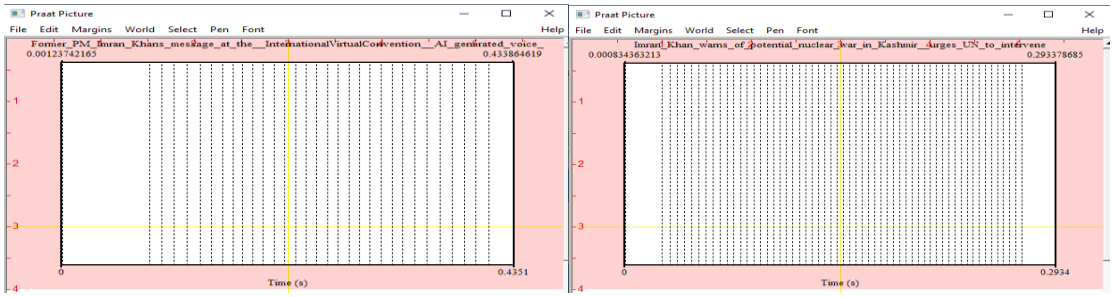
If the values are discussed, both speeches also represent different values in this case. In AlGS, the number of pulses is 11 and the No. of periods are 9. But we can see that the UNS has No. of the pulse is equal to 60 and No. of periods is equal to 55. There is a huge difference between the pulses and period values in both cases.

(ii)World



A similar pattern can be seen in the pictures for the world' of both speeches. The dotted lines are apart from each other in AIGS while the lines are very close to each other if the picture on the right side is taken into consideration. The right virtual speech delivered by Imran Khan has 25 No. of pulses and 24 periods and that of the speech delivered shows 58 pulses and 57 periods of the same word which is the world' in this situation.

Both speeches have been shown graphically parallel to each other. The left side picture shows the pulses of AI-generated speech while the right side picture shows the pulses of Imran Khan's original voice in the speech delivered at the plate form of the United Nations Assembly.

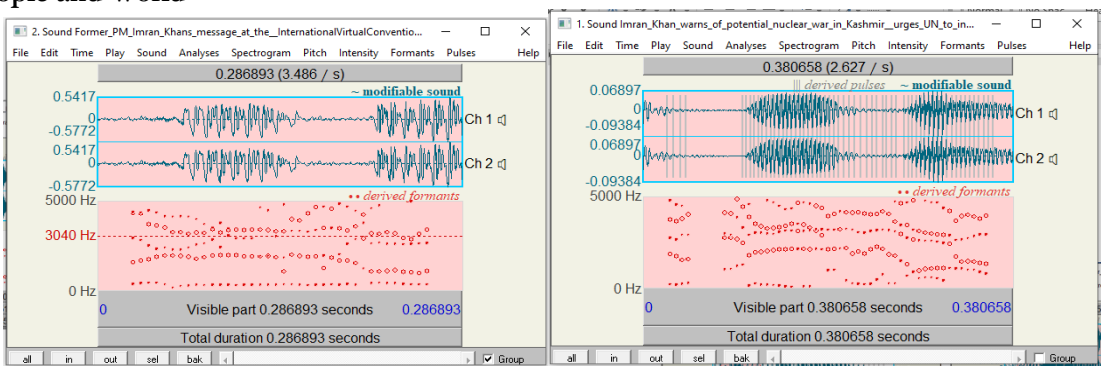


The AI-generated speech shows 30 pulses and the number of periods in this case is 31. The gap of the pulses can be seen in this picture. The dotted lines are very apart from each other. But if the second picture is discussed, the number of pulses is 56. The number of periods. are 57 in this human-generated speech. There is a small gap among the lines shown in dotted form.

4.5. Formants

The formants are the important characteristics of human speech. They are the acoustic resonance of speech. Formants help differentiate the vowel sounds. The researcher has found all the four formats of the words 'people', 'world', 'devised' and 'when'.

(iii) People and World



The words "people" and "world" are being analysed here based on formants in two speeches of AIGS and UNS. The pictorial evidence of the words for both speeches has shown that formants have significantly different values represented by dated waves.

Table: 1

Speech/ Formants/ Bandwidth	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	B.W1 (Hz)	B.W2 (Hz)	BW3 (Hz)	B.W4 (Hz)
AIGS	345.83	1663.91	2671.72	3516.75	51.82	2037.72	2531.64	207.47
UNS	640.12	1970.08	3005.65	3840.63	1843.80	275.39	134.70	1078.34

First of all, the word 'people' is being discussed in both of the cases of AIGS and UNS speeches. The formants formed by the AIGS are significantly higher than the formants of UNS. The same pattern has been represented by BWI and BW2. Both the bandwidths are higher in the speech of UNS and the difference is very high. But for the bandwidths of three and four AIGS show higher values than the values of UNS for the same bandwidths.

(i) World

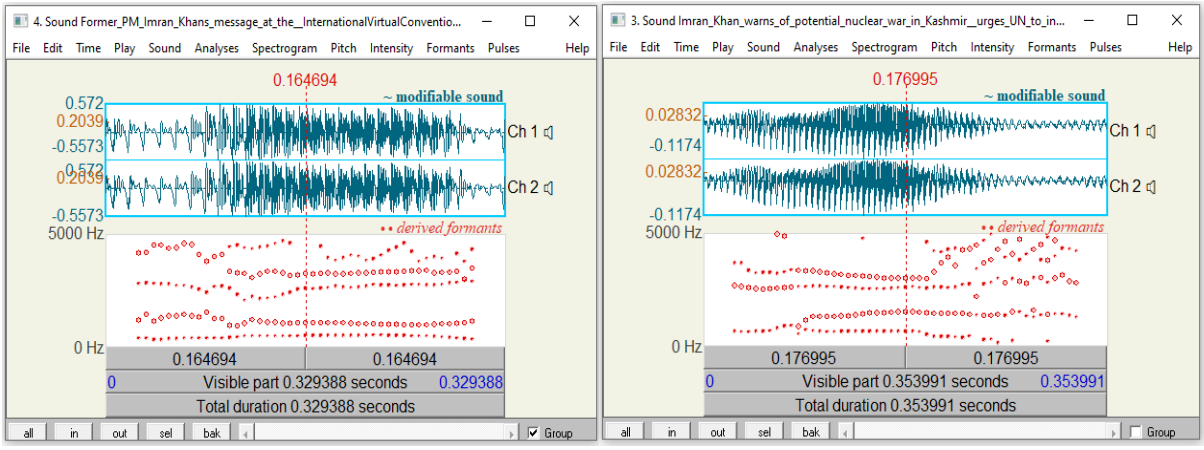


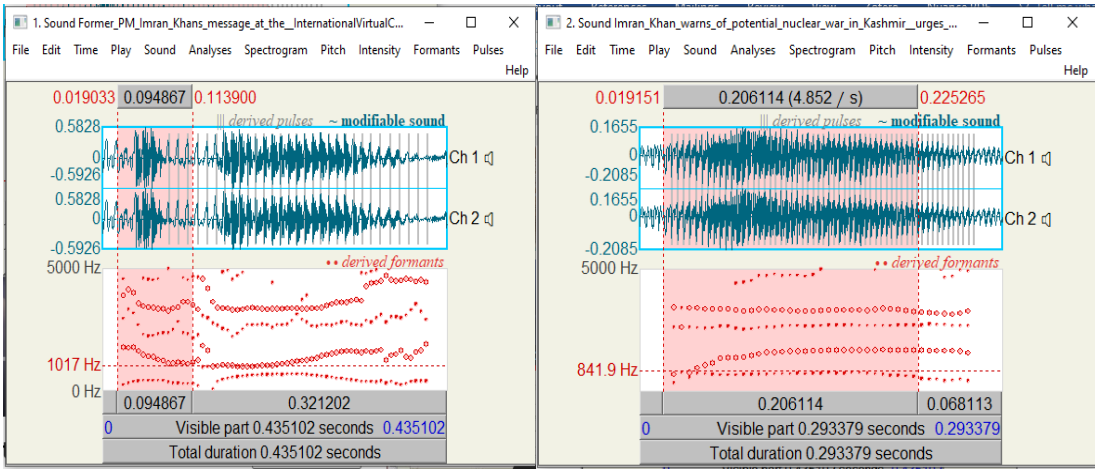
Table 2

Speech/ Formants/ Bandwidth	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)		B.W1 (Hz)		B.W2 (Hz)	BW3 (Hz)	B.W4 (Hz)
AIGS	432.41	1101.98	2622.30	3541.61		26.89		44.12	321.69	107.60
UNS	709.88	1861.22	3026.21	3788.99		152.66		159.94	158.44	66.65

The words "people" and "world" are being analysed here based on formants in two speeches of AIGS and UNS. The pictorial evidence of the words for both speeches has shown that formants have significantly different values represented by dated waves.

Here the researcher's focus is on the numerical representation and comparison of formants more than just showing pictures of both of the words. In the case of the word the 'world', the table values have made the concept clear that all the focus formants have different values if they are compared one to one joining The UN'S values are greater in each situation or formant than the formants produced by AIGS. If the bandwidth is compared to this word, it can be seen that BW1 and BW2 and BW4 are greater in the case of UNs and less significantly in the case of AIGS case. But for bandwidth of four, the patterns have been changed. BW3 for UNS is almost one fourth if the comparison has been done with AIGS.

Different words 'devised' and 'when' for the analysis were taken here to be analysed.



F1- Comparison

The first formant in the case of Al generated Speech is 338.53 Hz. It shows F1 of the vowel 'i' 17 words devised. Between 'd' and 'V' consonants we have the 'i' sound. The bandwidth in this case is 36.3716 Hz.

Contrary to that the UN speech the word 'when also has a short vowel sound between 'w' and 'n' which is 'e'. The first formant in this case is 687.1056 Hz which is far from the first formant of AIGS. The bandwidth also differs from F1 of AIGS. UNS show bandwidth 192.6548Hz.

F2-Comparison

The second formant of AIGS is 1356.43 Hz and the bandwidth of F2 for this speech is 266.33 Hz. In the case of UNS, the F_2 shows a value of 1525. 92 Hz. This is again a huge difference between human speech and Al speech of the same person. The bandwidth in this case also has a huge difference from AIGS. This speech shows band width of the vowel mentioned earlier is 113.34 Hz.

F3-Comparison

Here the comparison of F3 has been done to differentiate human speech from Al speech. For AIGS, the F3 has 2547.9335 Hz and the bandwidth for the formant is 507. 1753 Hz. If the comparison has been done between the two speeches, the Fs for UNS is 2671-4222Hz. Again there is a reasonable difference between the two speeches. The bandwidth for this speech 168.7816Hz which is quite different from the bandwidth of AIGS. In this case, the bandwidth is less while in previous cases bandwidths, were more than AIGS.

F4-Comparison

For the evaluation of F4, the value shown by praat for AIGS is 3513. 1263 Hz. This speech has a bandwidth value for Fs is 237.0988 Hz. In case of UNS F3 is 3326.4173 H2 and the bandwidth for UNS is 152.4523 Hz. Both of the values for UNS are less as compared to the values shown by AIGS. But the differences are quite noticeable to differentiate one from the other.

5. RESULTS AND DISCUSSION

The analysis was done at multiple levels to find the originality of a speech based on Al technology with a comparison of another speech based on the human voice. The following discussion on the analysis has been done concerning the research question mentioned in the introductory part of this paper.

The research question in this study is: How does Imran Khan's AIGS differ from the UNS? "This question was answered with the help of analysis of the words on the basis of parameters give in Praat. In this study, two words and two different words were analysed based on all the parameters discussed above. The researcher used the same words 'people' and 'world' to evaluate the two speeches. Similar parameters were applied to the words 'devised' and 'when' for the analysis of different. Let us discuss the results taken from the software one by one.

The intensity time graphs of all the words shown by screenshots and the numerical values obtained have shown that the intensity. by prant of AlGS is higher than that of UNS in the case is Power. This can be seen throughout the graph. The min. values and the max values in both speeches. Because the results were taken in three different ways each time, the same pattern was observed. It can be said that the speech produced by the Al tool shows higher intensity as compared to the intensity level of human sound.

The next parameter of our research is pitch, which shows the results totally different from the intensity shown earlier. All of the same words and different words gave huge differences. For example, the word people' has an average value of 90.71Hz and 215.7 He for AIGS and UNS respectively. In other words that is the world' has also shown the same pattern and has values of 88.52 Hz and 184.3 Hz respectively for AIGS and UNS. The word 'devised was used in Akas and when' for UNS and the important. the thing that the proat mentioned here was the Same pattern as the earlier examples had described for the pitch. The word 'devised 92.03 Hz and that of when has 227.8 Hz.

From all the results, it can be concluded that AIGS shows low pitch values as that of the pitch values shown by UNS. The speech generated by the Al tool shows less value while a human being shows a higher value in this specific parameter.

This parameter can also be compared to the intensity because the intensity and pitch in both of the speeches behaved oppositely for all the samples. In the case of intensity AIGS has greater values and in the case of UNS pitch has greater values.

The next parameter of our research is pitch, which shows the results totally different from the intensity shown earlier. All of the same words and different words gave huge differences For example, the word people' has an average value of 90.71 Hz and 215.7 He for AIGS and UNS respectively. The other word that is the world' has also shown the same pattern and has values of 88.52 Hz and 184.3 Hz respectively for AIGS and UNS. The word 'devised' was used in Akas and 'when' for UNS and the important thing that the praat mentioned here was the same pattern as the earlier examples had described for the pitch. The word 'devised 92.03 Hz and that of when has 227.8 Hz.

From all the results, it can be concluded that AIGS shows less pitch as compared to the pitch values shown by UNS. The speech generated by the Al tool& shows less value while human beings show more value in this specific pattern.

This parameter can also be compared to the intensity because the intensity and pitch in both of the speeches behaved oppositely for all the samples. In the case of intensity AIGS has greater values and in the case of UNS pitch has greater values.

The analysis of the pulses of all the word tokens from the speeches of AIGS and UNS gives important clues to decide about the non-human nature of a particular sound. The words which have 11, 25 and 30 pulses for the words 'people', 'world' and 'devised' respectively. The low values of pulses can also be easily seen through pictorial representation in the figures for AIGS. Now if comparison has been done, the values of UNS are far greater than the values of AIGS. The words "people", world and when have pulse values of 60, 58 and 56 respectively.

These results are also very significantly important to conclude that Al speech cannot produce proper pulses while a human being produces proper pulses of every word. The same results have been found in the values of periods. AIGS has less no. of periods and UNS has a greater no. of periods.

The last parameter based on which both other speeches, for example, AIGS and UNS can be differentiated very easily is formant analysis of the speeches. These were four formants which were taken for the analysis of the words people, "world", "devised' and 'When'. Table 1 and Table 2 clearly - show the comparison of the results. F1, F2, F3 and F4 for UNS have the larger values and AIGS have low values. So we can say that human-based speech has

values greater than those of non-human based speech. BW1 and BW4 also show very values as compared to the AIGS case for the word "people". BW2 and BW3 deviate from the pattern followed by BWI and BW2. In this situation, AIGS shows greater values than UNS. If the discussion was done for the word 'world', the results of formants are the same for UNS and AIGS. BW1 and BW2 give higher values than UNS and lower values for AIGS. BW3 and BW2 reverse the results.
But we can conclude that the formants are also the best way to judge whether the speech is original or not with the help of comparison with the reference speech.

6. CONCLUSION

From all the above discussion and analysis, it can be concluded that the Al speech always shows low values as compared to the speech delivered by a human being for the comparison of all the parameters except intensity. Intensity is the only parameter of sound which has larger values than a human speech sound.

Future Suggestions

The researcher conducted the study with the help of only two speeches by Imran Khan one was targeted and the other was a referenced speech. The researchers, who are very much interested in this field of forensic linguistics, can take a larger sample. They can take -more than on speech of the two categories. The one most important thing, the other persons' speeches should also be taken for Sampling for the analysis through praat in increase the generalizability.

References

1. *Acoustic phonetics*. (n.d.). The University of Sheffield. https://www.sheffield.ac.uk/linguistics/home/all-about-linguistics/about-website/branches-linguistics/phonetics/what-do-phoneticians-study/acoustic
2. Albin, A. (2014). PraatR: An architecture for controlling the phonetics software "Praat" with the R programming language. *Journal of the Acoustical Society of America, 135*, 2198-2199.
3. Barreda, S. (2021). Fast Track: fast (nearly) automatic formant-tracking using Praat. *Linguistics Vanguard*, *7*(1), 20200051.
4. Batalla, F.J., Márquez, R.G., Gonzalez, M.P., Laborda, I.G., Fernández, M.F., & Galán, M.M. (2014). Acoustic Voice Analysis Using the Praat programme: Comparative Study With the Dr. Speech Programme ☆. *Acta otorrinolaringológica española, 65*, 170-176.
5. Beňuš, Š. (2021). *Investigating spoken English: A practical guide to phonetics and phonology using Praat*. Springer Nature.
6. Boersma, P. (2002). Speak and unSpeak with P RAATRAAT.
7. Coulthard, M., & Johnson, A. (2010). *The Routledge handbook of forensic linguistics*. Routledge.
8. Haralambous, Y. (2024). Phonetics/Phonology. In *A Course in Natural Language Processing* (pp. 15-37). Cham: Springer International Publishing.
9. Ko, H., Woo, M.R., & Choi, Y. (2020). Comparisons of voice quality parameter values measured with MDVP, Praat, and TF32. *Phonetics and Speech Sciences*.
10. Ladefoged, P. (1996). *Elements of acoustic phonetics*. University of Chicago Press.
11. Pareniuk, A.V. (2021). Analysis of Speech Sounds In the Praat Program. *Electronic and Acoustic Engineering*.
12. Pleva, M., Juhár, J., & Thiessen, A.S. (2015). Automatic Acoustic Speech segmentation in Praat using cloud based ASR. *2015 25th International Conference Radioelektronika (RADIOELEKTRONIKA)*, 172-175.
13. Seong, C. (2022). Guidance to the Praat, a Software for Speech and Acoustic Analysis. *Journal of The Korean Society of Laryngology, Phoniatrics and Logopedics*.
14. Sondhi, S., Vijay, R., Khan, M., & Salhan, A.K. (2016). Voice analysis for detection of deception. *2016 11th International Conference on Knowledge, Information and Creativity Support Systems (KICSS)*, 1-6.
15. Sultana, R. (2023). Are non-native speakers' English pronunciations similar?: Evidence through PRAAT recording software. *Qeios*.