Detection of Ethno-Lingual Identity Using Artificial Intelligence, Machine Learning and Voice Tools: Introducing "Automated Criminal Ethnicity Identification System"

Chandra Yadav N.[1], Saini A.[2], Sharma V.[3*]

DOI: https://doi.org/10.58260/j.nras.2202.0108

[1] Nikhil Chandra Yadav, M.Sc Forensic Science, School of Basic and Applied Sciences, Galgotias University, Greater Noida, Uttar Pradesh, India.

[2] Aditya Saini, M.Sc Forensic Science, School of Basic and Applied Sciences, Galgotias University, Greater Noida, Uttar Pradesh, India.

[3*] Vinny Sharma, Assistant Professor, School of Basic and Applied Sciences, Galgotias University, Greater Noida, Uttar Pradesh, India.

Voice evidence is also known as voiceprint like fingerprint, it has been proven to substantiate the findings. Voiceprint is a dissimilar character for different people. In forensic science sometimes, we come across cases where the suspect's or victim's ethnicity has to be identified using various number of identification factors like voice, physical and anthropological features etc. In such cases the examination of an individual's ethnicity may be identified using the other available identification factors but when it comes to the Ethno-Lingual identification then examining the individual's language for the same and that too without any digital tool, i.e., doing it manually, becomes a sturdy task for the examiner. The database of the voice samples of Hindi, English and Mother language has been successfully created by the authors which is named as the "Automated Criminal Ethnicity Identification System" (ACEIS). In this paper, the author has summarised the various studies conducted on the ethno-lingual identification and their acquisition. Based on the studies, it was concluded in the review that the use of Artificial Intelligence and Machine Learning was used in prior studied but in India it hasn't been done yet. When known samples were analysed for their ethnicity, we noticed that an 80% matching was there among the samples belonging from same ethnicity. This matching-percentage was calculated on the basis of Pitch, Amplitude, Formant Frequencies, Frequencies and the average time taken to speak a word/letter etc.

Keywords: Ethno-lingual identification, Artificial Intelligence, Machine Learning, Forensic Science

Introduction

The voice we hear is the result of sound and speech production known as voice theory. It is more convincingly described by voice anatomy. So how speech and sounds are produced, to answer this question, it is produced by the movement of lips and tongue. Like we con vey us messages by making different gestures, the gestures made by the lips and tongues make the voice audible and recognized to the ears. It involves forcing air out of the lungs and making noise in the throat or lips in order to make these movements audible. There are three main organs of speech that are Respiratory, Phonatory and Articulatory.[36]

As this project work will entirely be based on the AI & ML (Artificial Intelligence and Machine Learning), so let's first understand what AI & ML actually is; Artificial intelligence is the ability of machines, particularly computer systems, to mimic human intellectual capabilities. Examples of particular AI applications include expert systems, machine learning, natural language processing, speech recognition, and machine vision.

Artificial intelligence is the term used to describe the intelligence displayed by machines. In the modern world, artificial intelligence has become quite popular. It is the imitation of natural intelligence by devices that have been designed to pick up on and imitate human behaviour. These machines are able to learn from experience and carry out jobs that humans would carry out. AI and other emerging technologies will significantly affect our quality of life as they develop. Everyone nowadays wants to interact with AI technology in some way, whether it is as an end user or by seeking a profession in the field. A subfield of artificial intelligence (AI) and computer science called machine learning focuses on using data and algorithms to simulate how people learn, progressively increasing the accuracy of the system. Machine learning is a branch of science and a potent technology that enables machines to learn from data and improve themselves. It is used in many services that we use every day. [12]

Machine learning is utilised in many apps on our phones, including speech recognition, email filters that separate spam from legitimate emails, websites that offer personalised suggestions, banking software that looks for odd activities,

And internet search engines. Numerous more possible uses for the technology exist, some of which carry greater risk. Future developments will significantly affect society and possibly assist the UK economy. Machine learning, for instance, could give us readily available "personal assistants" to help us manage our lives, it could greatly enhance the transportation system by using autonomous vehicles, and it could greatly enhance the healthcare system by enhancing disease diagnosis or personalising treatment. In security applications, machine learning could be used to examine online activity or email traffic. The ramifications of these and other technological applications must be thought through right away, and steps must be made to assure that usage will be advantageous to community. [12]

Ethnolinguistic identity leads to individual sense of identification or allegiance with a social group that is identified by shared linguistic and ethnic heritage. According to the ethnolinguistic identity theory, people utilise verbal and nonverbal communication accommodations such as verbal and nonverbal convergence to or divergence from their communication partner to highlight affiliation or disaffiliation, respectively. The macrolevel, large-scale language maintenance and change are impacted by these microlevel intergroup interactions. Related theories place a strong emphasis on the contribution of ethnolinguistic identity to the process of acculturation of ethnolinguistic groups as well as the development of communicative competence in foreign languages. Drawing on developmental, socio-cognitive, cultural, and linguistic psychology as well as the psychology of language learning, recent conceptualizations of bicultural and situational identities indicate new areas for research. [05]

Objectives

1. To collect the voice samples in Hindi, English and Mother Language and to create a database of voice samples of 'n' number of individuals belonging from different regions of India.

2. To introduce a system or software based on AI & ML for the purpose of recognition and identification of ethno-lingual identity of an individual.

3. To compare the collected samples with each other to obtain the average matching percentage.

Literature Review

It is said that Elizabeth Howe Chief, a researcher who worked with Native Americans, was the first to use an acculturative stress scale. Since then, self-report acculturation tools have been utilised often. The majority of prior assessments only included US samples. Publicly available self-report measures of acculturation were searched using several English peer-reviewed electronic journal databases, such as PsycINFO and Psyc Articles, to assess measures not limited to the US population and to extend previous research. The terms "assessment of acculturation," "acculturation," "measurement," and "meta-analysis" were among those used. In addition, a request for new instruments was submitted to the IACCP List for Cross-Cultural Psychologists. Our search found 50 publicly available measurements; in order to methodically review each of them, a categorization system was created. Our three key scale classifications were scale descriptors, psychometric characteristics, and conceptual and theoretical framework.[16]

A. Scale indicators

Target audience: Our review of publicly accessible metrics reveals that60.9% are focused on a particular demographic. Most are directed against various American ethnic communities.

Age: 14 percent of advertisements are produced particularly for the adult immigrant community, while 34 percent identify a particular age range while concentrating on the target group's age bracket.12 percent are for youth and adolescents, and 8 percent are for children.

Subscales: A single scale makes up the bulk of acculturation measures (54%) while two or more subscales make up the remaining 46%. The latter speak of subscales that assess different facets of acculturation. The basis for the subscales is typically a conceptual analysis or factor analytic data.

B. Psychological traits

Reliabilities: 80 percent of the measures have documented psychometric characteristics. 11.1 percent of the scales and 13.3 percent of the subscales had reported reliability values lower than.70. Factorial validity and other additional psychometric features are rarely discussed.

Theory and Conceptual Framework

A. Conditions for acculturation: Acculturation conditions are evaluated by claims such "I have experienced discrimination because I have trouble speaking Spanish." The bulk of the instruments lack any statement that gauges the circumstances of acculturation. [03]

B. Acculturation orientations: Sample items measuring acculturation orientations are "I would prefer to live in an American community" and the majority of the measures (50.5%) do not include items assessing acculturation orientations. [03]

C. Acculturation results: Phrases like uncomfortable since my family members don't know Mexican/Latino ways of doing things" are used to gauge psychological acculturation outcomes. "Having to accept the regional democratic structure" and similar phrases.are used to evaluate behavioural results. That just a small minority (23.4%) of the ratings never include expressions measuring acculturation outcomes, while the bulk of the remaining scales (76.6%) focus on behavioural results (64.9%) rather than psychosocial symptoms. (11.7 percent). [03] Additionally, we looked at how well tools evaluate circumstances, orientations, and outcomes—three aspects of the acculturation process—individually or together. A minor majority of the instrument's 54.7 percent (conditions, orientations, or outcomes) deals with just one aspect, 30.5 percent are measured. [03]

D. Acclimation mindsets: These behaviours, which often relate to acculturation orientations, express preferences of the immigrant group toward the acculturation process. These attitudes might be seen as moderators or mediators between the conditions and results of acculturation. Measures of acculturation attitudes include claims like "I like to be around my conational" and Most of the tests evaluate acculturation attitudes (66.7 percent). [03]

E. Acculturative stress behaviours: Since categories pertaining to acculturative stress activities frequently correlate to clear and explicit experiences of the immigrant and mainstream communities, it is reasonable to assume that acculturative stress activities and quick cultural assimilation outcomes are related. Examples include "Frequently engage in festivals or observe traditional Chinese holidays and festivities" and "In what language do you often watch television?"

The majority of subscales contain items designed to assess acculturation behaviour (86.3 percent). Additionally, we looked at how often attitudes and behaviours are included in assessments, and we discovered that most instruments evaluate both attitudes and behaviours (53.7 percent). Those who assess attitudes and behaviours independently make up the remaining 46.3 percent; subscales measure either attitudes (14%) or behaviours (%). (32.3 percent). [03]

F. Conceptual model: Questions like "In whose culture(s) are you sure you can act?" are found in one-dimensional assessments (41.5 percent). with response options such as Hispanic/Latino language only, Anglo/American language only, or "spouse preference" with all-Mexican statements selected like "I speak English at home" can be used to evaluate bidimensional acculturation tactics. "I eat American stuff at home."

G. Life domains: The majority of scales (91.3 percent) contain statements that evaluate acculturation across many domains. Language is mentioned in a variety of ways in 70% of the measurements, two examples of public comments to gauge acculturation. The following are some examples of questions to gauge acculturation in the private sphere.

According to Sapolsky, the early phases of language learning are more conducive to aptitude as it is now understood. Given that, unlike traditionally defined skills, this view can also enable the learner to advance beyond the stage of basic communication to the phase of more fully perfecting the nuances of a new language, it is conceivable that ethno-linguistic relativity can actually help both the early and later stages of learning. thing can contend so it while second language learners employ universal principles and techniques in the early stages, by the advanced stages they are using second language-specific methods that may be more impacted by ethno-lingual relativity.

Although motivation may not initially appear to have anything to do with ethnolinguistic relativity, it can be closely related and difficult to discern. According to Gardner & Lambert's original definition, integrative motivation is associated with positive feelings for the target language group and the potential to integrate with or socialize with its members. Consequently, how one feels in the group

That speaks the target language will affect how well the word is learned, according to Gardner's more recent socio-educational model, which acknowledges that learning a language requires acquiring parts of behaviour typical of another cultural group. It also acknowledges the influence of cultural attitudes on learning. Ethnic and linguistic relativistic theory seems to relate to all these correlations between cultural relevance to the learner and associations with attitudes toward members of other groups and a desire to learn about the mores of their society to the extent that they will be linked. to openness to different linguistic and cultural styles.

The motivational ideas from other fields of education have been attempted to be applied by Crookes and Schmidt to the study of second languages. When they point out that they support the argument for linking motivation with ethnolinguistic relativity. The social attitude discussed here may be related to ethnolinguistic relativity, as students who lack this perspective may be less motivated to learn a new language because it will seem less applicable to them. Crookes and Schmidt challenge and urge more hypothesis testing in their call for further study of motivation.

Numerous social and psychological elements, according to Schumann, might support learning a second or foreign language. The notion of ethno-linguistic relativity places particular emphasis on the acceptance of ambiguity in personality traits and cultural adaptation in affect. Anyone who has tried to learn a new language will be relatable to the fact that one frequently needs to function in perplexing situations when conversation topics and acceptable replies are unclear. According to some theories, students who have a limited tolerance for ambiguity may respond to such circumstances by becoming depressed, disliking, or avoiding them. Naaman, Frolic, and Stern discovered a strong correlation between tolerance for ambiguity and listening comprehension but not with an imitation test. These results, in line with Cohen, suggest that students with a high tolerance for this ambiguity may be able to listen more carefully and understand more of the information provided, in contrast to students with a lower tolerance who are confused by the linguistic input and pay. pay attention less effectively. One might argue that how regulated and constrained or open to new ideas one's outlook is determines a major part of one's tolerance for such uncertainty.

Materials and Methods

A. Materials:

- **Collection of samples:** The 'n' number of samples will be collected from the different regions of India. These collected samples will be used to make the database. The voice samples will be collected using a mobile phone/recorder for better audio quality. The individual will be asked to sign a consent form before taking their voice sample. A transcription will be given to them, having a paragraph in Hindi, English and in their respective mother language also (The transcription will be in such a way that it includes all the vowels and consonants). Then the recording of voice sample will be done.

- **Digital Voice Recorder:** A digital voice recorder was used during this project to collect the voice samples.

- **Samples:** The London letter has been used to call the English sample. And a random line of a newspaper has been called to get the Hindi sample. And a person has been called according to his wishes to call the mother tongue line.

- **Preservation of samples:** The collected voice samples will be preserved in a pen drive/hard disk/memory card along with their respective unique serial number.

- **AI & ML IDE (Jupyter):** JupyterLab is a web-based Interactive Development Environment for Jupyter notebooks, code, and data. JupyterLab is flexible: configure and arrange the user interface to support a wide range of workflows in data science, scientific computing, and machine learning. JupyterLab is extensible and modular: write plugins that add new components and integrate with existing ones. The Jupyter Notebook is an open-source web application that allows us to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more.

- **Audacity:** It is a free, easy-to-use, multi-track audio editor and recorder for Windows, macOS, GNU/Linux and other operating systems. The interface is translated into many languages.

- **Gold Wave:** Gold Wave is a commercial digital audio editing software product developed by Gold Wave Inc, first released to the public in April 1993. Gold Wave has an array of features bundled which define the program.

B. Methodology: A quiet place was first selected for sampling. So that the sound coming from behind in the voice sample can be reduced or removed. The consent form was then filled out by the person giving the sample and the consent form was signed as their consent. Then voice samples were collected by calling the above paragraph with a mic with a digital voice recorder. After the samples were taken, they were arranged according to the serial number of their MoU by providing them sample numbers in the database.

The 50 samples collected for this project were put in folders according to their phone numbers and saved in the database. After this these temples were tested with the help of Audacity, Gold Wave, AIML tools. During the test, the samples were studied at different levels. Features like prosodic features, amplitudes, pitch, frequency, format frequency, etc. were studied thoroughly for personal identification and Ethno-Lingual identification.

To collect, random sampling method was used during this project, keeping an eye on the ethnicity of the sampler. While collecting the samples, we tried to collect as many samples as possible from different parts of India. So that we can accurately trace their Ethno-Lingual identity.

Working of Gold wave

This is how the Gold Wave works and the spectrograph after analysing the particular word looks.

After resampling the audio range from 44,000 Hz to 11025 Hz because the normal audio range for humans is between 20-20,000 Hz but we narrow it down to 11025 Hz because the shortest range available in Gold Wave is 8000 Hz and then another range is 22,050 crossing20,000 Hz.

The bandwidth is set to 200-4000 range. The original samples are in .mp3 format which is then converted to .wav as .mp3 file is not accessible in Gold Wave. This is done to obtain the specific cue word from all the voice samples of a particular person and then for the rest of the samples of different speakers.
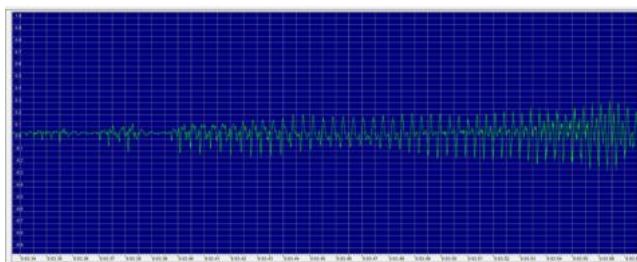
Figure 1: Working and Interface of Gold Wave

Result

Clue Words: Clue words are the similar sounding words that are used for the purpose of identification of a questioned voice sample and their individual features during an examination.

There are two types of variations that may be seen in a voice sample during the examination of a speaker or a questioned voice and standard sample:

1. Intra-speaker Variations: The variations that occurs within the voice sample of a speaker are referred as the Intra Speaker Variations.

2. Inter-speaker Variations: The variations that occurs within the voice samples of two or more different speakers are referred as the Inter-speaker variations.

Significance: The significance of Clue words during an examination is to reduce the "Intra-speaker Variations".

Temporal Parameters:
- Temporal parameters are the 'Time Dependent' parameters.
- Temporal parameters are 'Not Speaker Dependent' parameters.

Spectrographic Analysis: Now, let us see the examination done for the ethno-lingual identity of the speakers using the samples collected for this study via spectrographic analysis of the samples using the selected clue words. The spectrographic analysis done for the voice samples are shown below using the photographs and screenshots taken during the examination. The screenshots of spectrogram have the details of the voice samples, where the sample marked as 'A' contains the sample spoken in English whereas, the samples marked as 'B' contains the sample spoken in Hindi and the samples marked as 'C' (present in database) contains the samples spoken in the individual's mother tongue.

Information in the below observation tables: The examination of the clue words is shown using a table which contains the parameters observed and examined for the various voice samples present in the database.
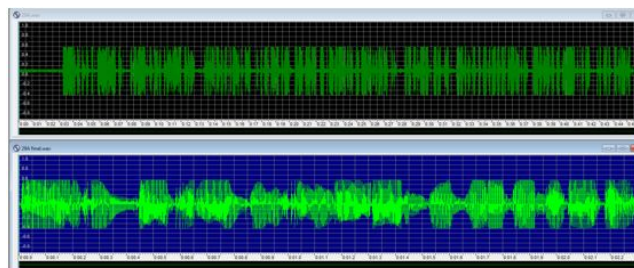
Spectrographic Analysis:

PERSON-01
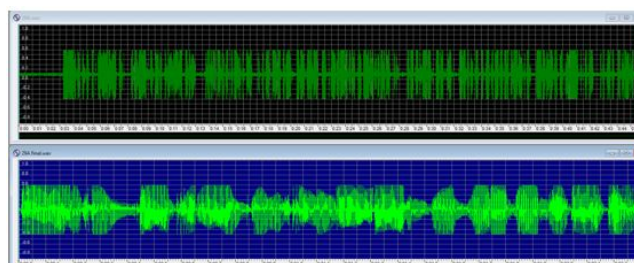

Figure 2: (1A) Spectrographic Analysis

Figure 3: (1B) Spectrographic Analysis

Details of Samples Voice and Clue Words 7

PERSON – 01A

Sr. No.	Clue Words (English)	Wave File 01A			Wave file 01A_CW		
		From (Seconds)	To (Seconds)	Duration (Seconds)	From (Seconds)	To (Seconds)	Duration (Seconds)
01	Our	3.270	3.500	0.230	00.00	0.230	0.230
02	Is	4.670	4.872	0.202	0.230	0.432	0.202
03	But	5.375	5.520	0.145	0.432	0.577	0.145
04	Are	7.040	7.160	0.120	0.577	0.697	0.120
05	Has	9.680	9.870	0.190	0.697	0.887	0.190
06	To	10.375	10.540	0.165	0.887	1.052	0.165
07	And	11.400	11.560	0.160	1.052	1.212	0.160
08	I	11.560	11.700	0.140	1.212	1.352	0.140
09	For	12.270	12.481	0.211	1.352	1.563	0.211
10	He	13.730	13.850	0.120	1.563	1.683	0.120
11	Will	13.870	14050	0.180	1.683	1.863	0.180
12	Be	14.085	14.225	0.140	1.863	2.003	0.140
13	We	--	--	--	--	--	--
14	The	42.785	42.865	0.080	2.003	2.083	0.080
15	Fox	43.650	43.790	0.140	2.083	2.223	0.140
16	Dog	45.120	45.245	0.125	2.223	2.348	0.125

Table 1 - Analysis of clue words, person 01(A)

Details of Samples Voice and Clue Words

PERSON – 01B

Sr. No.	Clue Words (Hindi)	Wave File 01B			Wave file 01B_CW		
		From (Seconds)	To (Seconds)	Duration (Seconds)	From (Seconds)	To (Seconds)	Duration (Seconds)
01	रोड	0.650	0.910	0.260	00.00	0.260	0.260
02	कार	3.035	3.200	0.165	0.260	0.425	0.165
03	गयी	6.120	6.270	0.150	0.425	0.575	0.150
04	रूप	8.370	8.495	0.125	0.575	0.700	0.125
05	लिए	13.635	13.840	0.205	0.700	0.905	0.205
06	जहाँ	16.600	16.790	0.190	0.905	1.095	0.190
07	पर	16.890	16.990	0.100	1.095	1.195	0.100
08	कर	21.090	21.245	0155	1.195	1.350	0.155
09	दिया	21.290	21.507	0.217	1.350	1.567	0.217
10	चार	24.985	25.185	0.200	1.567	1.767	0.200
11	लोग	25.210	25.465	0.255	1.767	2.022	0.255
12	रहे	28.545	28.700	0.155	2.022	2.177	0.155
13	बाद	38.785	38.975	0.190	2.177	2.367	0.190
14	उड़	40.200	40.400	0.200	2.367	2.567	0.200
15	गए	40.400	40.600	0.200	2.567	2.767	0.200
16	तेज	41.745	41.960	0.215	2.767	2.982	0.215
17	गयी	44.250	44.480	0.230	2.982	3.212	0.230
18	गया	57.920	58.150	0.230	3.212	3.442	0.230
19	है	1:17.450	1:17.680	0.230	3.442	3.672	0.230

Table 2 - Analysis of clue words, person 01(B)
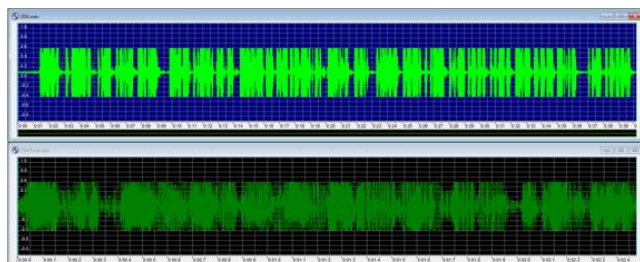
Spectrographic Analysis

PERSON – 02

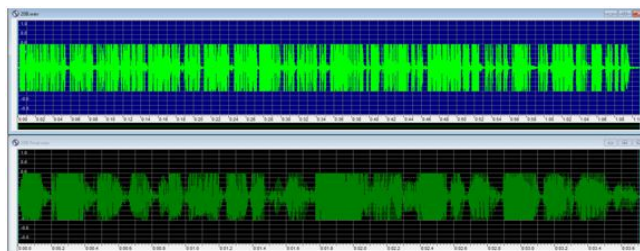**Figure 4: (2A) Spectrographic Analysis**

Figure 5: (2A) Spectrographic Analysis

Details of Samples Voice and Clue Words

PERSON – 02A

Sr. No.	Clue Words (English)	Wave File 02A			Wave file 02A_CW		
		From (Seconds)	To (Seconds)	Duration (Seconds)	From (Seconds)	To (Seconds)	Duration (Seconds)
01	Our	1.430	1.645	0.215	00.00	0.215	0.215
02	Is	2.285	2.480	0.195	0.215	0.410	0.195
03	But	3.115	3.265	0.150	0.410	0.560	0.150
04	Are	4.315	4.440	0.125	0.560	0.685	0.125
05	Has	6.390	6.520	0.130	0.685	0.815	0.130
06	To	6.875	6.945	0.070	0.815	0.885	0.070
07	And	7.795	7.945	0.150	0.885	1.035	0.150
08	I	7.945	8.150	0.205	1.035	1.240	0.205
09	For	8.445	8.560	0.115	1.240	1.355	0.115
10	He	9.855	10.020	0.165	1.355	1.510	0.165
11	Will	10.020	10.220	0.200	1.510	1.710	0.200
12	Be	10.220	10.340	0.120	1.710	1.830	0.120
13	We	28.040	28.170	0.130	1.830	1.960	0.130
14	The	37.030	37.160	0.130	1.960	2.090	0.130
15	Fox	38.145	38.335	0.190	2.090	2.280	0.190
16	Dog	39.585	39.775	0.190	2.280	2.470	0.190

Table 3 - Analysis of clue words, person 02(A)

Details of Samples Voice and Clue Words

PERSON – 02B

Sr. No.	Clue Words (Hindi)	Wave File 02B			Wave file 02B_CW		
		From (Seconds)	To (Seconds)	Duration (Seconds)	From (Seconds)	To (Seconds)	Duration (Seconds)
01	रोड	0.685	0.885	0.200	00.00	0.200	0.200
02	कार	2.480	2.730	0.250	0.200	0.450	0.250
03	गयी	5.260	5.470	0.210	0.450	0.660	0.210
04	रूप	7.500	7.660	0.180	0.660	0.840	0.180
05	लिए	12.085	12.255	0.170	0.840	1.010	0.170
06	जहाँ	14.065	14.290	0.225	1.010	1.235	0.225
07	पर	14.370	14.515	0.145	1.235	1.380	0.145
08	कर	18.070	18.215	0.145	1.380	1.525	0.145
09	दिया	18.275	18.490	0.215	1.525	1.740	0.215
10	चार	21.700	21.915	0.215	1.740	1.955	0.215
11	लोग	22.010	22.170	0.160	1.955	2.115	0.160
12	रहे	23.700	23.890	0.190	2.115	2.305	0.190
13	बाद	32.435	32.710	0.275	2.305	2.580	0.275
14	उड़	33.825	33.925	0.100	2.580	2.680	0.100
15	गए	33.925	34.125	0.200	2.680	2.880	0.200
16	तेज	34.900	35.155	0.255	2.880	3.035	0.255
17	गयी	37.500	37690	0.190	3.035	3.325	0.190
18	गया	50.300	50.525	0.225	3.325	3.550	0.225
19	है	1:9.480	1:9.640	0.160	3.550	3.710	0.160

Table 4 - Analysis of clue words, person 02(B)

Discussion

A visual depiction of a certain set of sounds in terms of time, frequency, and amplitude is produced by the instrument. The four fundamental components of an analogue spectrograph are a magnetic tape recorder/player, a tape scanning mechanism with a drum that carries the paper to be marked, an electronic variable filter, and an electronic stylus that transfers the processed data to the paper. The magnetic tape recording's energy levels in a narrow frequency range are sampled by the analogue sound spectrograph and recorded on electrically sensitive paper. The following brief frequency range is then analysed, and samples and markings of the energy levels are made. Until the whole target frequency range is examined for that section of the recording, this process is repeated. A visual representation of the patterns, in the form of bars or formants, of the acoustical occurrences throughout the time period investigated is the final result, known as a spectrogram. A spectrogram will be generated by the machine in around 80 seconds. The spectrogram is a graph with an X axis representing time (often 2.4 seconds) and a Y axis representing frequency (typically 0 to 4000 or 8000 Hz). The marks' level of darkness provides an estimate of the relative amplitude of the energy present at a particular frequency and time.

The spectrographic examination is accomplished for every one of the speakers and their individual voice and just the outcome and information of an individual speaker has been referenced in the experiment. The remainder of the work is done similarly. The idea of spectrographs relies upon the Fourier hypothesis. Incorporating of the method relies upon the refined utilization of the electronic separating or on the methodology of complex computational calculations. The Fourier hypothesis certifies that any occasional waveform can be examined into a progression of sine waves with various frequencies, amplitudes and stage connections.

The most conjoint strategy for spectrographic sifting of the discourse signal is bandpass channels that conducts frequencies inside lower and higher scope of frequencies passing. The lower and higher restrictions of the bandpass are separated in those frequencies where decrease is contrasted with the focal point of the band.

These are called known as the cut off frequencies of the channel. Channels with thin bandpass are dormant or inert in their reaction though Wide band channels react in a quick way. Their time goal is very acceptable aside from their recurrence goal which is very poor.

The range of an acoustic wave is essentially the consequence of a Fourier investigation of the waves under assessment. That is, it is a decree of what frequencies are available and what their amplitudes are. Every recurrence segment (consonant)of the wave is addressed by a line sited roughly situated on the recurrence hub. The tallness of every line shows its sufficiency in dB.

The diagram isn't persistent and there are no focuses between the music. The square waves are made out of discrete recurrence parts. The highest point of the consonant lines can't be combined to frame a persistent and a smooth bend. The clear spaces or the clear lines suggests the shortfall of frequencies and not the shortfall of any information. This sort of range is known as line range.

Cue words are comparable sounding words which have been chosen from all the voice for one speaker and this cycle is revised for the accounts. The cue words have been chosen based on CV (consonant-vowel), CVC (consonant-vowel-consonant), CVVC (consonant-vowel-vowel-consonant) format. The arch depends on getting similar words accessible in the voice test of a speaker.

This is accomplished for each conceivable word found from the examples. This validates that the specific voice test has a place with a specific speaker. The pitch of the female speakers is normally high. The normal pitch and forces have been noted down for each voice test of every speaker exclusively to show that they differ each time even in the voice tests of a similar speaker because of regular variety among them.

Conclusion

In forensic science sometimes, we come across cases where the suspect's or victim's ethnicity has to be identified using various number of identification factors like voice, physical and anthropological features etc. In such cases the examination of an individual's ethnicity may be identified using the other available identification factors but when it comes to the

Ethno-Lingual identification then examining the individual's language for the same and that too without any digital tool, i.e., doing it manually, becomes a sturdy task for the examiner.

In this paper, the author has summarised the various studies conducted on the ethno-lingual identification and their acquisition. Based on the studies, it was concluded in the review that the use of Artificial Intelligence and Machine Learning was used in prior studied but in India it hasn't been done yet.

After all the literature review and reviewing the prior studies, the author has already concluded in the review that this study hasn't been done yet in India, and the author thought that the use of Artificial Intelligence and Machine Learning can be used in the field of Forensic Science for Ethno-Lingual Identification and we have also started the research over this topic for the sake of the future of forensic science.

Now, when the study has been successfully done, the author conclude that the Voice recognition is a very significant and constructive tool in the field of forensic science. It is as valuable as any other evidence found to prove or solve a crime. Voice evidence is also known as voiceprint like fingerprint, it has been proven to substantiate the findings. Voiceprint is a dissimilar character for different people. Like DNA, it is unique to everybody. The research is done to understand that there is always a difference between the voice of the individuals belonging from different ethnicity. Though some people may try to hide their true identity and may try to change their voice (disguising) but they can't change their individual features of their voice like formant frequencies, time taken to speak the word/letter etc (the disguise will be noticed instantly as the change in word length, vowels and consonants will be noticed).

By taking a look at various features of speech sound and by investigating it on various programming software, it is reasoned that they are extraordinary and is a useful exploration instrument in cases, for example, recover calls, tapped telephone discussions and so on. This examination depends on digging of cue words of the speaker of all the 16/50 voice tests recorded directly for every speaker. The research depends on choice of the comparable sounding words accessible in the voice tests of the speaker.

To validate the outcomes found from gold wave and spectrographic examination carried out on Pratt, a spectrographic examination of the voice tests.

Some the objectives proposed in the research has been successfully achieved which are as follows whereas some are partially achieved. We can't say anything on those objectives until and unless we conduct a more detailed and through study on those objectives. We will be conducting the research on those particulars in details soon.

1. The database of the voice samples of Hindi, English and Mother language has been successfully created by the author which is named as the **"Automated Criminal Ethnicity Identification System" (ACEIS).** The database is currently having 50 voice samples of the individuals belonging from different ethnicity in India. Out of these 50 samples, we have examined all the 50 samples and the result of 02 samples has been shown in this research paper.

2. As far as the second objective of this research is concerned, that was to introduce a system or software-based o AI & ML for the purpose of recognition and identification of ethno-lingual identity of an individual, we are still working on it. We hope that this particular objective will also be achieved in the further studies. We did make a programme using Python and Jupyter based on Artificial Intelligence and Machine Learning but after a count of 17 total tries, it still isn't working that efficiently. We are trying our level best to make it work efficiently and once it is functional properly, we will introduce this system for the sake of the future of forensic science.

3. The collected samples have been compared and analysed successfully.
- The comparative analysis of two voice samples of an individual shows 100% matching with each other, confirming the accuracy of the system.

- The comparative analysis of two different individual's voice samples shows negligible (~1 - 2%) matching with each other, which again confirms the sensitivity and accuracy of the system.

- When known samples were analysed for their ethnicity, we noticed that an 80% matching was there among the samples belonging

- from same ethnicity. This matching-percentage was calculated on the basis of Pitch, Amplitude, Formant Frequencies, Frequencies and the average time taken to speak a word/letter etc.

Though there were only 3 main objectives of this study but there were some other features and other parameters of the voice samples and the tool that were notices during the study. We would like to mention them also as the conclusion of this study. These were as follows:

- The samples have been freed from distortion to maximum level with the use of Gold wave and by using its different features like resampling, adjusting the bandwidth, noise cancellation, adjusting the volume to 75%, double, half etc.

- There are different physical, emotional and mental parameters for distorted voice such as having a cold and cough, shivering, fever, feeling sad, sorrowful, happy, excited, some people stammer, some have a mental disorder etc.

Future Scope:

- This technology can be used like AFIS in the coming times. Just as fingerprints data is present in AFIS, in the same way data of voice samples will also be present in it, which we can use to find out the ethnicity of a person.

- This 'Automated Criminal Ethnicity Identification System' (ACEIS) will help in ending any investigation as quickly as possible in the time to come.

- ACEIS (Automated Criminal Ethnicity Identification System) will be able to narrow down the list of suspects during a trial and time will be decreased for the Law Enforcement Agencies to caught the guilty or criminal.

- The system will be able to identify the merged voice samples by selecting the keywords own its own using Artificial Intelligence and will apply the same in future using the machine learning.

- The way this project is going on, we are sure that it will be modified in way that it would be able to do the Speaker Identification and Speech Recognition with a great accuracy and precision for the sake of the future of Forensic Science.

Limitations and Problems faced:

- If the voice is distorted or filled with background noise interference, it causes a lot of problems for the expert in his examination.

- It takes time for the expert to analyse the voice with different accents.

- It takes time for the expert to differentiate between original and disguised voice.

- Physical parameters affect the voice very much so, the words or sentences are not very much clear and hard to understand and analyse.

Reference

1. Bokayev, B., Kazhenova, A., Zharkynbekova, S., Beisembayeva, G., & Nurgalieva, S. (2014). Adjustment and ethno-lingual identification of Kazakh repatriates: Results of sociolinguistic research. Journal of Sociology, 50(4), 545-559.

2. Gqibitole, K. M. (2014). Ethno-lingual Issues. Counter Discourse in African Literature, 127.

3. Citron, J. L. (1995). Can cross-cultural understanding aid second language acquisition? Toward a theory of ethno-lingual relativity. Hispania, 105-113.

4. Citron, J. L. (1993). The Role of Ethno-Lingual Relativity in Second Language Acquisition. Working Papers in Educational Linguistics, 9(2), 29-41.

5. oels, K. A. (2017). Identity, ethnolinguistic. The International Encyclopaedia of Intercultural Communication, .

6. Brown, H. D. (1980). The Optimal distance model of second language acquisition's Quarterly, 14(2), 157-164.

7. Clavijo, F. J. (1984). Effects of teaching culture on attitude change. Hispania, 67, 88-91.

8. Crookes, G. & Schmidt, R. (1991). Motivation: Reopening the research agenda. Language Learning,41(4), 469-512.

9. Fantine, A. E. (1993). Becoming better global citizens: The promise of intercultural competence. Odyssey, Fall, 1993, 17-19.

10. Fishman, J. (1981). Language policy: Past, present and future. In C.A. Ferguson & Sebific Heath (Eds.), Languages in the USA (pp. 516-526). New York: Cambridge University Press.

11. Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second language acquisition. Canadian Journal of Psychology, 13, 266-272.

12. "Artificial Intelligence and Machine Learning Made Simple"

13. Kauffman, N. L., Martin, J. N., Weaver, H. D., & Weaver, J. (1992). Students abroad: Strangers at home. Yarmouth, ME: Intercultural Press.

14. Owen, Jennifer. Voice Identification the Aural/Spectrographic Method | Owen Forensic Services, LLC. https://www.owenforensicservices.com/voice-identification-the-aural-spectrographic-method/. Accessed 05 June 2022.

15. Kohls, L.R. (1984). The values Americans live by. Washington, D.C.: Meridian House International.

16. Larsen-Freeman, D. & Long, M. (1991). An introduction to second language acquisition research. New York: Longman.

17. Pease, D. M., Berko Gleason, J., & Pan, B.A. (1993). Learning the meaning of words: Semantic development and beyond. In J. Berko Gleason (Ed.), The development of language (pp.115-149). New York: Macmillan.

18. Schumann, J. H. (1978). Social and psychological factors in second language acquisition. In J. C. Richards, (Ed.), Understanding second and foreign language learning issues and approaches (pp.163-178). Rowley, MA: Newbury House.

19. Skehan's, P. (1991). Individual differences in second language learning. SSLA, 12 (2),275-98.

20. Sapolsky, B. (1989). Conditions for second language learning. Oxford: Oxford University Press.

21. Whorf, B.L. [1967 (1956)]. Language, thought, and reality. Cambridge, MA: MIT Press.

22. Wolfson, N. (1989). Perspectives: Sociolinguistics and TESOL. New York: Harper &Row.

23. "The Branches of Forensic Science - An Overview of Its Various Disciplines." IFF Lab, 23 Feb. 2018,

24. Williamson, G. (2022). Effective Voice Production.

25. Analyzing of the vocal fold dynamics using laryngeal videos - Scientific Figure on ResearchGate.

26. LING520: Lecture Notes 2. (2022). Retrieved 28 June 2022,

27. Physics- Sound Waves Simulation Quiz - Proports' Quiz. (2022). Retrieved 28 June 2022,

28. French, J. P. (2017). A developmental history of forensic speaker comparison in the UK. English Phonetics, 271-286.

29. Ladefoged, P., & Johnson, K. (2014). A Course in Phonetics (7th ed.). Cengage Learning.

30. Bhuta, 1T. Patrick, L. &Garnett, J. D. (2004). Perceptual evaluation of voice quality and its correlation with acoustic measurements.

31. Hirano. M. Hibi, S. Yoshida. T. Harada, Y. Kazuya, H., & Kikuchi, Y. (1988). Acoustic Analysis of Pathological Voice:

32. (PDF) Automatic Speech Recognition System for Isolated. (n.d.). Automatic Speech Recognition System for isolated and Connected Words of Hindi Language by Using Hidden Markov Mo del Toolkit HTK

33. M. M., J. G., Subhashini, D. P., & Krishnan, D. M. (n.d.). Automated Speech Recognition System A Literature Review.

34. Sivakumar. N.S. (n.d.). Acoustic Analysis for Human Voice Disorder Classification Using Optimization and Machine Learning Technique

35. Magdin, M., Sulka, T., Toma nova, J., & Voser, M. (2019). Voice Analysis Using PRAAT Software and Classification of User Emotional State. International Journal of Interactive Multimedia and Artificial Intelligence, 5(6),

36. Ladefoged, P., & Johnson, K. (2014). A Course in Phonetics (7th ed.). Cengage Learning.

37. Dunn. H. K. (1961). Methods of Measuring Vowel Formant Bandwidths. The Journal of the Acoustical Society America,

38. Kersta. L. G. (1948). Amplitude Cross Section Representation with the Sound Spectrograph. The Journal of the Acoustical Society of America,

39. Davis, L. I. (1964). Biological Acoustics and the Use of the Sound Spectrograph. The Southwestern Naturalist, 9(3), 118.

40. Parse. V... & Jamieson. D. G. (2000). Identification of Pathological Voices Using Glottal Noise Measures. Journal of Speech. Language, and Hearing Research,

41. Parsa, V., & Jamieson, D. G. (2001). Acoustic Discrimination of Pathological Voice. Journal of Speech, Language, and Hearing Research,

42. Zhang. Y., & Jiang. J. J. (2008). Acoustic Analyses of Sustained and Running Voices from Patients with Laryngeal Pathologies. Journal of Voice,

43. Zhang. Y. Wang, S. Sun, P. & Phillips, P. (2015). Pathological brain detection based on wavelet entropy and Hu moment invariants. Bio-Medical Materials and Engineering.

44. Zhang. Z., Xia. Y., Xing. F, Me Gough. M., & Yang, L. (2017). Medent: A Semantically and Visually Interpretable Medical Image Diagnosis Network. 2017 IEEE Conference on Computer Vision and Pattern Recognition

45. Gu, D., Li, Y., Jiang, F, Wen, Z., Liu, S., Shi, W., Lu, G., & Zhou, C. (2020, July). VINet: A Visually Interpretable Image Diagnosis Network.

46. C. M. V. & Radha, V. (2012). Speaker Independent Isolated Speech Recognition System for Tamil Language using HMM, Procedia Engineering,

47. Kurian, C., & Balakrishnan, K. (2012). Development & evaluation of different acoustic models for Malayalam continuous speech recognition. Procedia Engineering,

48. "Criminology Vs. Criminalistics: What's the Difference?".

49. "Job Description for Forensic Laboratory Scientists". Crime Scene Investigator EDU. 12 November 2013. Archived from the original on 6 September 2015. Retrieved 28 August 2015.

50. "Prosecutors just got millions of pages of Trump documents. His taxes are only the beginning". NBC News. Retrieved 27 February 2021.

51. "Sections". American Academy of Forensic Sciences. 27 August 2015. Archived from the original on 30 August 2015. Retrieved 28 August 2015.

52. Shorter Oxford English Dictionary (6th ed.), Oxford University Press, 2007, ISBN 978-0-19-920687-2