INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

LEGAL ADMISSIBILITY OF VOICE RECOGNITION TECHNOLOGY IN THE INFORMATION AND COMMUNICATION TECHNOLOGY – ENABLED CRIMINAL INVESTIGATIONS

AUTHOR – MONALISHA R A, LLM STUDENT AT SCHOOL OF EXCELLENCE IN LAW, TNDALU , CHENNAI

BEST CITATION - MONALISHA R A, LEGAL ADMISSIBILITY OF VOICE RECOGNITION TECHNOLOGY IN THE INFORMATION AND COMMUNICATION TECHNOLOGY – ENABLED CRIMINAL INVESTIGATIONS, *INDIAN JOURNAL OF LEGAL REVIEW (IJLR)*, 5 (13) OF 2025, PG. 156-170, APIS – 3920 – 0001 & ISSN - 2583-2344.

1. INTRODUCTION

The growth of Information and Communication Technology (ICT) has transformed criminal investigations, with voice recognition technology becoming an important forensic tool[216]. It assists in authenticating recorded conversations, identifying suspects, and securing digital evidence[217]. However, its admissibility raises questions of reliability, privacy, and constitutional protections.

In India, the Bharatiya Sakshya Adhiniyam Act, 2023 and the Information Technology Act, 2000 recognise electronic records[218], but challenges persist regarding tampering, accuracy, and safeguards against self-incrimination under Article 20(3) of the Constitution. The Supreme Court in *Ritesh Sinha v. State of Uttar Pradesh* held that compelling a voice sample is not testimonial compulsion[219], yet the absence of procedural standards leaves uncertainties. Comparative jurisprudence in the United States and United Kingdom underscores the need for scientific validation and transparency[220].

This article undertakes a doctrinal study of the admissibility of voice recognition evidence in ICT-enabled investigations, highlighting statutory provisions, judicial approaches, and the need for reform.

[216] R. Togneri & J. Pullella, "An Overview of Speaker Identification: Accuracy and Robustness Issues" (2011) 1 IEEE Circuits and Systems Magazine 23.
[217] Lawrence M. Solan & Peter M. Tiersma, Speaking of Crime (University of Chicago Press 2005).
[218] The BSA Act, sec 62 & 63; IT Act, 2000 sec 2&4
[219] Ritesh Sinha v. State of Uttar Pradesh, (2019) 8 SCC 1.
[220] Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993); R v. Robb [1991] 93 Cr App R 161 (CA).

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

2. TECHNOLOGICAL FOUNDATIONS OF VOICE RECOGNITION

2.1 WHAT IS VOICE RECOGNITION SYSTEM?

A voice recognition system is a technological process that analyses and identifies an individual based on the unique characteristics of their voice. It works by capturing a voice sample, extracting acoustic features such as pitch, tone, frequency, and speech patterns, and then comparing these with stored voiceprints or samples[221].

There are two main categories:

1. Speaker Identification – determining who is speaking by comparing an unknown voice against a database of known samples[222].
2. Speaker Verification (Authentication) – confirming whether a given voice matches the claimed identity, often used in biometric security systems[223].

In forensic investigations, such systems are employed to authenticate recorded conversations, identify suspects in telephonic interceptions, and support evidence in cybercrime and organized crime cases. However, the accuracy of voice recognition depends on factors like recording quality, background noise, language variations, and the robustness of the algorithm[224].

2.2 HOW DOES VOICE RECOGNITION SYSTEM WORK?

A voice recognition system functions by analysing the unique biometric features of a person's voice. The process generally involves four key stages:

1. Voice Signal Acquisition – The system records speech through a microphone or telecommunication device[225].
2. Feature Extraction – Acoustic properties such as pitch, tone, formants, frequency spectrum, and rhythm are converted into digital parameters, producing a "voiceprint."[226]
3. Pattern Matching – The extracted features are compared against stored templates in a database using algorithms like Gaussian Mixture Models (GMM), Hidden Markov Models (HMM), or Deep Neural Networks (DNN).[227]
4. Decision Making – The system either identifies the speaker (speaker identification) or verifies the claimed identity (speaker verification). The output is based on probability scores indicating the likelihood of a match.[228]

The effectiveness of this process depends on factors such as recording quality, background noise, channel distortion, and the variability of the speaker's voice due to health or emotional state[229]. In forensic practice, strict attention to accuracy, validation, and error-rate analysis is essential to ensure admissibility in courts.

2.3 TYPES OF VOICE RECOGNITION SYSTEM

Voice recognition systems are classified according to their function and the manner in which they process spoken input. The major types include:

- Speaker Identification
  This type is used to find out *who* is speaking. An unknown sample is compared against several stored samples to determine the closest match.

[221] Campbell, J.P., "Speaker Recognition: A Tutorial" (1997) Proceedings of the IEEE 85(9), 1437–1462.
[222] Rose, P., Forensic Speaker Identification (Taylor & Francis 2002).
[223] Doddington, G., "Speaker Recognition Based on Idiolectal Differences among Speakers" (2001) EUROSPEECH Conference Proceedings 2521–2524.
[224] Togneri, R. & Pullella, J., "An Overview of Speaker Identification: Accuracy and Robustness Issues" (2011) 1 IEEE Circuits and Systems Magazine 23.
[225] Campbell, J.P., "Speaker Recognition: A Tutorial" (1997) Proceedings of the IEEE 85(9), 1437–1462.
[226] Rose, P., Forensic Speaker Identification (Taylor & Francis 2002) 17–23.
[227] Kinnunen, T. & Li, H., "An Overview of Text-Independent Speaker Recognition: From Features to Supervectors" (2010) 52 Speech Communication 12.
[228] Doddington, G., "Speaker Recognition Based on Idiolectal Differences among Speakers" (2001) EUROSPEECH Conference Proceedings 2521–2524.
[229] Togneri, R. & Pullella, J., "An Overview of Speaker Identification: Accuracy and Robustness Issues" (2011) 1 IEEE Circuits and Systems Magazine 23.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

It is often applied in police investigations where intercepted calls need to be linked to suspects[230].

- Speaker Verification
  Also known as authentication, this system checks whether a person is truly who they claim to be. It compares the speaker's live voice with a previously stored voiceprint of the same individual. Common applications include secure banking transactions and mobile log-ins[231].
- Text-Dependent Systems
  In this model, the speaker must say a particular phrase or password. The system analyses both the content of the words and the unique vocal features. For example, systems that ask a user to repeat "my voice is my password."[232]
- Text-Independent Systems
  Unlike text-dependent models, these systems do not rely on fixed phrases. They can identify or verify a speaker regardless of the words spoken, focusing only on vocal characteristics like pitch and frequency. This approach is useful in forensic work with spontaneous recordings[233].
- Forensic Voice Recognition
  This type is tailored for legal and investigative use. It combines expert human analysis with computer algorithms to decide whether a voice in a recording can be linked to a particular suspect. Courts usually rely on this in cases of criminal conspiracy, terrorism, or fraud[234].

2.4 APPLICATIONS OF VOICE RECOGNITION SYSTEMS

Voice recognition technology has moved beyond commercial convenience and now plays an important role in both everyday life and criminal justice. Its applications can be grouped into the following areas:

- **Criminal Investigations**

Used to identify suspects from intercepted calls, ransom recordings, or surveillance audio. Forensic experts rely on voice comparison to support or rule out involvement of individuals in criminal conspiracies[235].

- **Judicial Proceedings**

Courts admit voice samples as evidence when authenticated and scientifically validated. Voice analysis helps in corroborating witness testimony, verifying telephonic threats, and detecting tampering in recordings[236].

- **Security and Authentication**

Widely used in banking, mobile applications, and access control systems where a person's voice serves as a biometric password[237].

- **Telecommunication and ICT Services**

Customer service systems use voice recognition for call routing, identity verification, and hands-free interaction. Virtual assistants like Siri, Alexa, and Google Assistant rely on this technology for user interaction[238].

- **Counter-Terrorism and Surveillance**

Intelligence agencies employ voice databases to monitor communications of suspects across borders. Helps in tracking organised crime networks and preventing cyber-enabled offences[239].

[230] Rose, P., Forensic Speaker Identification (Taylor & Francis 2002).
[231] Kinnunen, T. & Li, H., "An Overview of Text-Independent Speaker Recognition: From Features to Supervectors" (2010) 52 Speech Communication 12–40.
[232] Furui, S., "Recent Advances in Speaker Recognition" (1997) Pattern Recognition Letters 18(9), 859–872.
[233] Campbell, J.P., "Speaker Recognition: A Tutorial" (1997) Proceedings of the IEEE 85(9), 1437–1462.
[234] Togneri, R. & Pullella, J., "An Overview of Speaker Identification: Accuracy and Robustness Issues" (2011) 1 IEEE Circuits and Systems Magazine 23–33.

[235] Rose, P., Forensic Speaker Identification (Taylor & Francis 2002).
[236] R. M. Malkani v. State of Maharashtra, AIR 1973 SC 157.
[237] Kinnunen, T. & Li, H., "An Overview of Text-Independent Speaker Recognition" (2010) 52 Speech Communication 12–40.
[238] Furui, S., "Recent Advances in Speaker Recognition" (1997) Pattern Recognition Letters 18(9), 859–872.
[239] Togneri, R. & Pullella, J., "An Overview of Speaker Identification: Accuracy and Robustness Issues" (2011) 1 IEEE Circuits and Systems Magazine 23–33.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

➢ **Accessibility**

Supports individuals with disabilities by enabling voice-controlled navigation, dictation, and interaction with ICT platforms[240].

These applications highlight the dual role of voice recognition: while it enhances efficiency in security and governance, it also raises questions of privacy, reliability, and legal admissibility.

2.5 FEATURES OF VOICE RECOGNITION SYSTEM

Voice recognition systems, also known as automatic speech recognition (ASR) systems, are designed to convert spoken language into text or to identify speakers based on their voice characteristics. These systems have become critical tools in information and communication technology (ICT)-enabled criminal investigations, offering efficient and non-invasive methods of identifying suspects or authenticating individuals. The major features of such systems are outlined below:

➢ **Acoustic Signal Processing**

Voice recognition begins with capturing the speaker's voice through a microphone. The raw audio input is processed to remove background noise, normalize volume, and extract relevant acoustic features such as pitch, tone, and frequency. These acoustic features are crucial for distinguishing between different speakers or recognizing spoken words accurately[241].

➢ **Speech-to-Text Conversion**

Most ASR systems use language models and acoustic models to interpret spoken language and convert it into text. Language models help predict the sequence of words based on grammar and context, while acoustic models map sound patterns to phonetic units. This dual modelling enhances accuracy, especially in noisy or complex linguistic environments[242].

➢ **Speaker Identification and Verification**

Speaker recognition functions in two primary ways: speaker identification, where the system determines who is speaking from a group of known voices, and speaker verification, which confirms if a voice matches a claimed identity. These features are especially valuable in criminal investigations involving intercepted calls, ransom demands, or anonymous threats.[243]

➢ **Use of Machine Learning and AI**

Modern voice recognition systems are powered by machine learning, particularly deep neural networks (DNNs) and convolutional neural networks (CNNs). These models are trained on large datasets to improve their ability to recognize speech patterns, accents, and speaker-specific traits. Over time, the system adapts and becomes more accurate as it encounters more diverse inputs.[244]

➢ **Biometric Voiceprint Analysis**

Each individual has a unique vocal tract, which results in distinct voice characteristics. Voice recognition systems can create a voiceprint, a digital model of these unique features for biometric identification. This makes voice recognition a form of biometric technology, similar to fingerprints or retinal scans.[245]

➢ **Language and Accent Adaptability**

Advanced systems are capable of recognizing multiple languages and adapting to various accents and dialects. They use contextual modelling and large vocabulary databases to maintain accuracy across diverse speech inputs[246]. This feature is essential in multilingual countries like India, where suspects may speak regional dialects.

➢ **Integration with Forensic Tools**

Voice recognition technologies can be integrated with forensic audio tools that

[240] Campbell, J.P., "Speaker Recognition: A Tutorial" (1997) Proceedings of the IEEE 85(9), 1437–1462.
[241] Jurafsky, D., & Martin, J. H. (2023). Speech and Language Processing (3rd ed.). Stanford University Press.
[242] Young, S. et al. (2015). "The HTK Book: Hidden Markov Models for Speech Recognition." Cambridge University Engineering Department.
[243] Campbell, J. P. (1997). "Speaker Recognition: A Tutorial." Proceedings of the IEEE, 85(9), pp. 1437-1462.
[244] Hinton, G. et al. (2012). "Deep Neural Networks for Acoustic Modelling in Speech Recognition." IEEE Signal Processing Magazine, 29(6), pp. 82-97.
[245] Kinnunen, T., & Li, H. (2010). "An Overview of Text-Independent Speaker Recognition." Speech Communication, 52(1), pp. 12–40.
[246] Yu, D., & Deng, L. (2016). Automatic Speech Recognition: A Deep Learning Approach. Springer.

perform spectrogram analysis, formant frequency mapping, and phonetic segmentation. These tools assist forensic experts in authenticating recordings, enhancing speech quality, and matching voice samples to known speakers.[247]

➢ **Real-Time and Batch Processing Capabilities**

Some systems are capable of real-time recognition, which is beneficial for live surveillance, emergency response, or monitoring of telephonic communication. Others work in batch mode, analysing pre-recorded audio files. This flexibility makes them adaptable to various stages of an investigation.[248]

➢ **Security and Anti-Spoofing Mechanisms**

To prevent deception through replay attacks or synthetic (AI-generated) voices, modern voice recognition systems incorporate anti-spoofing techniques. These include liveness detection, speech pattern consistency checks, and synthetic speech detection algorithms.[249]

2.6 ADVANTAGES AND DISADVANTAGES

Voice recognition technology, built upon advanced algorithms, signal processing, and machine learning techniques, has revolutionized the way humans interact with digital systems. In the context of criminal investigations, it has become a vital ICT enabled tool for both surveillance and evidence collection. However, like all technological solutions, it comes with both strengths and limitations based on its foundational technologies.

2.6.1 ADVANTAGES

✓ **Non-Invasive Biometric Identification**

Voice recognition enables biometric identification without physical contact. Unlike fingerprint or retina scans, it only requires a sample of the person's speech, which can even be captured passively in some cases[250]. This feature makes it ideal for remote surveillance and covert investigations.

✓ **Cost-Efficiency and Ease of Integration**

Compared to other biometric systems, voice recognition technology is relatively inexpensive to deploy, especially when integrated into mobile devices, phone lines, or virtual assistants. It also does not require specialized hardware apart from a basic microphone[251].

✓ **Remote and Real-Time Functionality**

One of the core technological benefits is the ability to function remotely and in real time. Investigators can monitor suspects' calls, verify identities during digital transactions, or assess threats from voice messages without physical proximity[252].

✓ **Scalability and Automation**

The underlying algorithms of voice recognition systems are scalable. With cloud computing and AI frameworks, voiceprints of thousands of individuals can be processed, stored, and matched at scale, making it suitable for large-scale criminal databases[253].

✓ **Continuous Learning Through AI**

Modern systems based on deep learning can improve over time as more voice data is fed into them. They learn to accommodate accent variations, dialects, and speaking styles through continuous training[254].

[247] Rose, P. (2002). Forensic Speaker Identification. Taylor & Francis.
[248] Rabiner, L., & Juang, B. H. (1993). Fundamentals of Speech Recognition. Prentice-Hall.
[249] Todisco, M., Delgado, H., & Evans, N. (2019). "ASVspoof 2019: Future Directions in Spoofed Speech Detection." Proc. INTERSPEECH, pp. 1008–1012.
[250] Faundez-Zanuy, M. (2007). "Biometric Security Technology." IEEE Aerospace and Electronic Systems Magazine, 22(6), pp. 15–26.
[251] Jain, A., Ross, A., & Prabhakar, S. (2004). "An Introduction to Biometric Recognition." IEEE Transactions on Circuits and Systems for Video Technology, 14(1), pp. 4–20.
[252] Kinnunen, T. & Li, H. (2010). "An Overview of Text-Independent Speaker Recognition: From Features to Supervectors." Speech Communication, 52(1), pp. 12–40.
[253] Zhang, C., & Koishida, K. (2017). "End-to-End Text-Independent Speaker Verification with Triplet Loss on Short Utterances." Proc. INTERSPEECH 2017, pp. 1487–1491.
[254] Yu, D., & Deng, L. (2016). Automatic Speech Recognition: A Deep Learning Approach. Springer.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

2.6.2 DISADVANTAGES

✓ **Vulnerability to Background Noise and Quality Degradation**

Voice recognition systems often rely on clean, high-quality input. In real-world scenarios especially criminal investigations, audio samples are frequently distorted due to ambient noise, compression artifacts, or recording limitations. This can lead to misidentification or failed recognition[255].

✓ **Susceptibility to Spoofing and Deepfake Attacks**

One of the critical technological vulnerabilities lies in the system's potential to be deceived by synthetic or cloned voices generated through AI (deepfakes). Such spoofing attacks can bypass security or falsely incriminate individuals if not properly detected[256].

✓ **High False Acceptance and Rejection Rates**

Despite advancements, no voice recognition system is 100% accurate. Depending on the dataset and model, systems may exhibit high False Acceptance Rates (FAR) or False Rejection Rates (FRR) leading to both wrongful access and denial of legitimate users[257].

✓ **Performance Affected by Physiological or Health Changes**

A person's voice can be affected by temporary factors such as illness (cold, sore throat), emotional stress, or aging. These changes may degrade the accuracy of the system in matching the voice to previously stored samples[258].

✓ **Linguistic and Accent Bias**

Some systems perform better for certain languages or accents, especially if they were trained on biased datasets. This creates a significant disadvantage in multilingual or demographically diverse regions such as India[259].

3. LEGAL FRAMEWORK

3.1 OVERVIEW OF INIDIAN LAWS ON VOICE RECOGNITION

Voice recognition technology has increasingly found relevance in Indian criminal investigations, particularly in the era of digital communication and cyber-enabled crimes. However, its use and admissibility are governed by a combination of statutory provisions, constitutional mandates, and judicial interpretations. While India does not have a single, dedicated legislation on voice recognition, several laws and case precedents provide a legal foundation for its application.

1. THE BHARATIYA SAKSHYA ADHINIYAM ACT, 2023 (BSA)

The BSA Act, 2023 originally did not account for digital or biometric evidence like voice samples. However, through judicial interpretation and amendments, it has evolved to accommodate such evidence under electronic records.

- Sections 62 & 63 of the act governs the admissibility of electronic records, including audio recordings. A voice recording, whether stored on a digital device or transmitted electronically, is admissible if accompanied by a proper certificate under this section[260].
- Section 39 of the act deals with expert opinion and is applicable in cases where forensic experts analyse voice samples using spectrograms or other tools. Their findings are treated as expert testimony.
- Section 72 of the act, though originally meant for handwriting or signature comparisons, has been extended by courts to allow the comparison of voice

[255] Rabiner, L. R., & Schafer, R. W. (2011). Theory and Applications of Digital Speech Processing. Prentice Hall.
[256] Wu, Z. et al. (2015). "Spoofing and Countermeasures for Speaker Verification: A Survey." Speech Communication, 66, pp. 130–153.
[257] Hansen, J. H. L., & Hasan, T. (2015). "Speaker Recognition by Machines and Humans: A Tutorial Review." IEEE Signal Processing Magazine, 32(6), pp. 74–99.
[258] Reynolds, D. A. (2002). "An Overview of Automatic Speaker Recognition Technology." Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).
[259] Sreeram, V., & Shetty, S. (2020). "Accent Bias in Speech Recognition Systems: A Case Study of Indian English." International Journal of Speech Technology, 23(4), pp. 903–911.
[260] The BSA Act, 2023 sections 62,63,39 & 72

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]
VOLUME 5 AND ISSUE 13 OF 2025
APIS – 3920 - 0001 *(and)* ISSN - 2583-2344
Published by
Institute of Legal Education
https://iledu.in

samples by experts or judges themselves in appropriate cases[261].

2. THE BHARATIYA NAGARIK SURAKSHA SANHITA ACT, 2023 & THE CRIMINAL PROCEDURE (IDENTIFICATION) ACT, 2022

The BNSS Act governs procedural aspects of investigation and evidence collection. Until recently, there was no express provision authorizing police to compel voice samples. This changed with the Criminal Procedure (Identification) Act, 2022, which replaced the Identification of Prisoners Act, 1920.

- Section 3 of the 2022 Act allows police officers and prison authorities to collect "measurements," which are defined to include biological and physical samples, and expressly includes voice samples[262].
- The accompanying rules (Criminal Procedure (Identification) Rules, 2022) outline how voice samples may be recorded and stored, though some practical clarity and standardization are still lacking.

3. THE CONSTITUTION OF INDIA, 1950

Two key constitutional protections must be considered when collecting voice samples:

- Article 20(3) states that no person accused of an offence shall be compelled to be a witness against themselves. The Supreme Court has clarified in Ritesh Sinha v. State of Uttar Pradesh (2019) that providing a voice sample does not amount to self-incrimination, as it is not "testimonial" in nature, but merely a physical attribute like fingerprints.[263]
- Article 21, which guarantees the right to life and personal liberty, has been interpreted to include the right to privacy. The Puttaswamy judgment (2017) emphasized that any biometric data collection must meet standards of legality, necessity, and proportionality[264]. Thus, voice sample collection must be lawful, for a legitimate aim, and with adequate safeguards.

4. THE INFORMATION TECHNOLOGY ACT, 2000

The IT Act, 2000, particularly after its 2008 amendment, regulates electronic communications and digital evidence. While it does not explicitly mention voice recognition, the Act supports the legal recognition of electronic records and digital signatures.

- Section 65B of the Evidence Act (as inserted via the IT Act amendment) remains central for authenticating audio recordings captured electronically.
- Further, under Sections 43 and 66, unauthorized access to voice data or its misuse (e.g., synthetic voice frauds) may attract civil and criminal liability[265].

3.2 ADMISSIBILITY CRITERIA OF VOICE RECOGNITION EVIDENCE IN INDIA

➢ **RELEVANCE**

The voice evidence must be relevant to the facts in issue, as per Sections 4–7 of the Bharatiya Sakshya Adhiniyam, 2023[266].

➢ **IDENTIFICATION OF VOICE**

The voice must be clearly attributed to the speaker, either by:

- A person who is familiar with the speaker's voice,
- An expert (e.g., forensic phonetician), or
- A validated voice recognition system.

If the speaker denies the voice, stricter proof (expert or scientific) is required[267].

[261] *Rakesh Bisht v. CBI*, (2007) 13 SCC 804.
[262] Criminal Procedure (Identification) Act, 2022, Section 3
[263] *Ritesh Sinha v. State of Uttar Pradesh*, (2019) 8 SCC 1.
[264] *Justice K.S. Puttaswamy v. Union of India*, (2017) 10 SCC 1.
[265] The Information Technology Act, 2000, Sections 43 and 66.
[266] Bharatiya Sakshya Adhiniyam, 2023, Sections 4–7. Replaced Indian Evidence Act, 1872 from July 1, 2024.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]
VOLUME 5 AND ISSUE 13 OF 2025
APIS – 3920 - 0001 *(and)* ISSN - 2583-2344
Published by
Institute of Legal Education
https://iledu.in

➢ **AUTHENTICITY AND INTEGRITY OF RECORDING**

There must be no signs of tampering, editing, or distortion. The device used to record must have been working properly and the data securely stored[268]. Chain of custody documentation strengthens admissibility.

➢ **CERTIFICATE UNDER SECTION 63(4) OF THE BSA ACT, 2023**

Voice recordings stored digitally must be accompanied by a certificate under Section 63(4) of the Bharatiya Sakshya Adhiniyam (similar to the old Section 65B of the Indian Evidence Act)[269].
The certificate must:

- Identify the device used,
- Describe how the data was produced,
- Confirm the accuracy and integrity of the recording.

➢ **AUDIBILITY AND CLARITY**

The recording must be clear and free from distortion, background noise, or overlap. The court must be able to clearly hear and understand the voice in question[270].

➢ **LEGAL ACQUISITION**

While evidence obtained through illegal or unethical means is not automatically inadmissible, Indian courts prefer that recordings are obtained lawfully[271]. However, in R.M. Malkani v. State of Maharashtra (1973), the Supreme Court held that even improperly obtained voice recordings can be admitted if:

- They are relevant,
- They are authentic,
- They do not violate the right to a fair trial (Article 21 of the Constitution).

➢ **EXPERT EVIDENCE (IF DISPUTED)**

If the voice is disputed, expert analysis may be required. The court may call upon:

- Forensic phonetics labs (e.g., CFSLs),
- Voice spectrogram analysis,
- Biometric comparison via AI tools (must be validated)[272].

➢ **SYSTEM RELIABILITY (FOR AUTOMATED VOICE RECOGNITION SYSTEMS)**

If a voice recognition algorithm or biometric system is used (e.g., AI matching):

- The system must be proven reliable and widely accepted,
- Accuracy and error rate must be demonstrated,
- The process used to match the voice must be transparent and scientifically valid[273]

3.3 JUDICIAL INTERPRETATIONS

Indian courts have addressed the admissibility of voice recordings and voice identification through several landmark judgments. The judiciary has emphasized the importance of authenticity, reliability, and relevance of such electronic evidence. Key rulings are summarized below:

❖ **R.M. Malkani v. State of Maharashtra (1973)**

The Supreme Court held that a tape-recorded conversation is admissible as evidence if it is relevant and not tampered with. It further held that the consent of the speaker is not necessary for admissibility, as long as the conversation was voluntarily made and the recording is authentic[274].

[267] Ram Singh v. Col. Ram Singh, AIR 1986 SC 3. Voice identification must be proved beyond doubt if disputed.
[268] Yusufalli Esmail Nagree v. State of Maharashtra, AIR 1968 SC 147: Tampering renders recordings inadmissible.
[269] Section 63(4), Bharatiya Sakshya Adhiniyam (similar to old Sec. 65B, Indian Evidence Act).
[270] R.M. Malkani v. State of Maharashtra, AIR 1973 SC 157: Voice must be clear and recognizable.
[271] Pooran Mal v. Director of Inspection (1974) 1 SCC 345: Illegally obtained evidence can be admissible.
[272] Ibid. Voice recording made without consent was admitted, provided it was authentic and relevant.
[273] No direct Indian case yet on AI voice recognition, but courts demand reliability and explainability of scientific tools (Ref: Selvi v. State of Karnataka, 2010).
[274] *R.M. Malkani v. State of Maharashtra*, AIR 1973 SC 157; (1973) 1 SCC 471.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]
VOLUME 5 AND ISSUE 13 OF 2025
APIS – 3920 - 0001 *(and)* ISSN - 2583-2344
Published by
Institute of Legal Education
https://iledu.in

"There is no violation of Article 21 if a person's voice is recorded without his knowledge, provided it is not under duress or coercion." – SC in Malkani

❖ **Ziyauddin Burhanuddin Bukhari v. Brijmohan Ramdass Mehra (1976)**

In this case, the Supreme Court accepted tape-recorded speeches as evidence during election trials and emphasized that the voice must be properly identified, either by the person who recorded it or by other reliable witnesses[275].

❖ **Ram Singh v. Col. Ram Singh (1986)**

This case laid down conditions for admissibility of tape-recorded evidence. The Court stressed:

- The voice must be clearly identifiable,
- The tape must be free from tampering,
- The recording device must have been in proper working condition,
- The recording must be relevant to the case[276].

These have become standard criteria in courts when dealing with voice evidence.

❖ **Yusufalli Esmail Nagree v. State of Maharashtra (1968)**

The Supreme Court held that electronic recordings can be treated as documents, and when produced from proper custody and shown to be authentic, they are admissible. However, the case also warned against accepting such evidence without sufficient safeguards[277].

❖ **Tukaram S. Dighole v. Manikrao Shivaji Kokate (2010)**

The Court reiterated that voice evidence must be corroborated and that a voice sample comparison by an expert (under Section 39 of the BSA Act) can be permitted. This shows growing reliance on forensic voice identification[278].

❖ **Selvi v. State of Karnataka (2010)**

Though focused on narco-analysis and brain mapping, the Court laid down strong principles around the right against self-incrimination and involuntary extraction of evidence. It is relevant when voice samples are compelled without consent[279].

4. ICT- ENABLED INVESTIGATIVE PRACTICES

4.1 FORENSIC APPLICATIONS IN VOICE ANALYSIS

Voice analysis is a vital tool in modern ICT-enabled criminal investigations, allowing law enforcement to authenticate voices, detect tampering, and identify suspects through unique speech characteristics. Forensic phonetics and biometric systems are widely used in both investigation and prosecution stages. Some key forensic applications are:

❖ **Speaker Identification**

- Purpose: To match a questioned voice with a known sample.
- Method: Analysts examine acoustic features such as pitch, formants, intensity, and duration, comparing them visually using spectrograms (voiceprints)[280].
- Tools:
  - Speech signal analysis software (e.g., Praat, Adobe Audition)
  - Spectrographic matching systems

❖ **Voice Authentication**

- Used in: Access control, telecom surveillance, and Aadhaar-based authentication.
- Types:

[275] *Ziyauddin Bukhari v. Brijmohan Mehra*, (1976) 2 SCC 17.
[276] *Ram Singh v. Col. Ram Singh*, AIR 1986 SC 3; (1985) Supp SCC 611.
[277] Yusufalli Esmail Nagree v. State of Maharashtra, AIR 1968 SC 147; (1968) 3 SCR 477.
[278] Tukaram S. Dighole v. Manikrao Shivaji Kokate, (2010) 4 SCC 329.
[279] Selvi v. State of Karnataka, (2010) 7 SCC 263
[280] Ram Singh v. Col. Ram Singh, AIR 1986 SC 3 – Emphasized the use of spectrographic analysis in identifying speakers.

- Text-dependent systems use fixed passphrases,
- Text-independent systems analyze natural speech[281].

❖ **Tampering Detection**

- Voice experts use waveform and frequency analysis to detect editing, splicing, or artificial modification (e.g., deepfake audio)[282].
- Authenticity is verified by checking for discontinuities in pitch or waveform.

❖ **Audio Enhancement**

- Techniques include noise filtering, gain control, echo cancellation, and de-reverberation to improve intelligibility of poor-quality recordings[283].
- Enhanced recordings are used for courtroom presentation or further forensic analysis.

❖ **Linguistic and Acoustic Profiling**

- Experts analyse speech characteristics like accent, dialect, grammar, vocabulary, and pacing to infer geographic, social, or ethnic background[284].
- Useful when the speaker is unknown and only a voice recording is available.

❖ **Voice Biometrics in Surveillance**

- Law enforcement and intelligence agencies use voiceprint databases to identify or track persons of interest in intercepted communications (e.g., terrorism or organized crime)[285].
- AI-powered speaker recognition systems (like i-vector and x-vector models) are deployed in high-security zones and cybercrime units.

INSTITUTIONS INVOLVED IN INDIA

- Central Forensic Science Laboratories (CFSLs) under DFSS, Ministry of Home Affairs provide voice comparison reports admissible in court[286].
- State Forensic Science Labs (FSLs) – handle regional cases.
- Voice ID units in law enforcement agencies including NIA, CBI, and State ATS.
- Private forensic labs must follow guidelines and certification standards for admissibility.

4.2 ICT INTEGRATION WITH OTHER TOOLS

The integration of Information and Communication Technology (ICT) with various investigative tools has transformed criminal investigations in India. Voice recognition systems are increasingly used in conjunction with other technologies, enabling law enforcement agencies to gather, cross-verify, and present digital evidence more effectively**.**

❖ **Integration with CCTV & Audio Surveillance Systems**

- CCTV cameras with built-in microphones can record audio, which can be extracted for voice analysis and synchronized with visual evidence[287].
- AI-based audio analytics can detect shouting, gunshots, or keywords in real-time, triggering alerts in sensitive zones (e.g., airports, public rallies).

❖ **Linkage with Call Detail Records (CDRs) and IPDRs**

- Voice recordings from intercepted calls are correlated with Call Detail Records

[281] S. K. Verma, Forensic Science and Criminology, 2015.
[282] Forensic Science Laboratory Manual, Directorate of Forensic Science Services (DFSS), Ministry of Home Affairs, Govt. of India.
[283] ibid. – Chapter on Audio-Video Forensics.
[284] Dr. H. M. Trivedi, "Forensic Linguistics and Profiling," Journal of Indian Law & Technology, 2019.
[285] National Crime Records Bureau (NCRB), Annual Report on Cybercrime Trends, 2021.

[286] Ministry of Home Affairs, Govt. of India, Circular No. MHA/DFSS/2020/VC – Authorizing CFSLs to provide expert reports under Section 39 of the BSA Act, 2023.
[287] **Delhi Police Smart City Surveillance Project Report (2022), Ministry of Home Affairs.**

(CDRs) or Internet Protocol Detail Records (IPDRs) to trace communication chains and confirm speaker identity[288].

- This integration supports timeline construction, geolocation mapping, and pattern analysis.

❖ **Use with GPS and Geofencing Tools**

- Voice data collected from mobile phones can be linked with GPS metadata to determine the speaker's location at the time of call[289].
- Geo-fencing tools are used to monitor individuals entering sensitive zones; voice input is used to verify their identity or intention.

❖ **Combined with Biometric Databases (e.g., Aadhaar, NATGRID)**

- Voiceprints can be linked to Aadhaar-linked identity records, particularly in welfare fraud and impersonation cases[290].
- NATGRID and Crime and Criminal Tracking Network System (CCTNS) allow seamless data sharing between police, intelligence, and other departments, often integrating voice data as one component among others (e.g., fingerprints, facial recognition).

❖ **Coupled with Digital Forensics Tools**

- Voice files recovered from devices (phones, laptops) are analysed along with metadata, deleted files, and social media data using forensic software like FTK, EnCase, or Oxygen Forensics[291].
- This integrated digital trail helps in establishing intent, motive, and links between suspects.

4.3 RELIABILITY AND SECURITY

As voice recognition technologies become central to law enforcement and forensic procedures, ensuring reliability and security of these systems is essential. Courts require not only accurate results but also secure handling of voice data to maintain evidentiary value and prevent misuse.

❖ **Reliability of Voice Recognition Systems**

Voice recognition systems must demonstrate consistent performance and scientific validity to be admissible in court.

✓ **Accuracy and Error Rates**

- Accuracy depends on:
  - Quality of input (microphone, noise),
  - Environmental factors (background noise, speech variation),
  - Speaker variability (accents, illness, stress)[292].
- Systems must report False Acceptance Rate (FAR) and False Rejection Rate (FRR), especially in biometric authentication[293].

✓ **Validation and Testing**

- Forensic voice comparison tools must be empirically validated as per international standards (e.g., ISO/IEC 19795 for biometric performance testing)[294].
- Indian forensic labs (e.g., CFSLs) often use spectrographic methods alongside software tools like *PRATT*, *SASpeechPro*, etc., tested in controlled environments[295].

[288] Indian Telegraph Act, 1885, Sec. 5(2); Lawful interception rules under IT Act, 2000.
[289] CFSL Training Module on GPS and Telecom Forensics, 2020.
[290] UIDAI Technical Framework for Voice Biometrics Integration, 2021.
[291] Directorate of Forensic Science Services (DFSS), Handbook on Digital and Cyber Forensics, 2022 Edition.
[292] Department of Electronics and IT (DeitY), Govt. of India, Standards for Voice Biometric Systems, 2020.
[293] ISO/IEC 19795: Biometric Performance Testing and Reporting.
[294] Bureau of Indian Standards (BIS), Draft Guidelines for Biometric Systems Testing, 2021.
[295] Directorate of Forensic Science Services (DFSS), Forensic Audio & Video Analysis Manual, 2020.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

✓ **Expert Testimony**

- Courts require expert witnesses to explain:
  - Method of analysis,
  - Degree of certainty,
  - Margin of error,
  - System limitations[296].

✓ **Security of Voice Data and Systems**

Given the sensitive nature of voice recordings (often tied to identity, location, or intent), robust data protection protocols are required.

✓ **Data Encryption**

- Voice recordings and biometric voiceprints must be encrypted during:
  - Capture,
  - Transmission (e.g., cloud storage),
  - Storage (databases or forensic lab servers)[297].

✓ **Access Control and Audit Trails**

- Only authorized personnel (e.g., forensic experts, IOs) should access voice data.
- Systems should maintain audit logs to track who accessed or modified any audio file[298].

✓ **Chain of Custody**

- To ensure admissibility, a clear and tamper-proof chain of custody must be maintained:
  - From collection (e.g., phone tap, voice sample),
  - Through analysis and reporting,
  - Up to submission in court[299].

✓ Protection Against Tampering and Deepfakes

- Voice cloning and deepfake audio pose new risks.
- Advanced forensic tools now include tamper detection algorithms and spectral fingerprinting to differentiate real vs AI-generated voices[300].

✓ Cybersecurity of ICT Systems

- Voice data repositories (like in CCTNS, NATGRID, or police forensic labs) must be protected against hacking, ransomware, and unauthorized duplication[301].

5. CHALLENGES AND RECOMMENDATIONS

While the integration of voice recognition systems into investigative practices marks a major advancement in digital forensics, it also presents technical, legal, procedural, and ethical challenges. Addressing these challenges is essential to maintain the credibility, admissibility, and fairness of the criminal justice process.

CHALLENGES

✓ **Lack of Standardization**

- **Problem**: No uniform forensic protocols exist across state and central labs for voice analysis.
- **Impact**: Results may vary depending on equipment, software, or interpretation by experts[302].

✓ **Absence of Clear Legal Framework for AI-based Voice Systems**

- Current Indian laws (e.g., Bharatiya Sakshya Adhiniyam, IT Act) do not specifically address AI-based voice recognition or biometric systems.

[296] *Ram Singh v. Col. Ram Singh*, AIR 1986 SC 3; courts must be satisfied about method and expertise.
[297] UIDAI Guidelines on Encryption & Security of Biometric Data, 2021.
[298] NCRB Guidelines on Data Protection in CCTNS Systems, 2022.
[299] *State of Punjab v. Baldev Singh*, (1999) 6 SCC 172 – Emphasized importance of proper chain of custody.
[300] National Cyber Forensics Lab (CERT-In), Advisory on Deepfake Audio Threats, 2023.
[301] Ministry of Home Affairs Cyber Security Cell, Audit Report on Law Enforcement IT Systems, 2021.
[302] Directorate of Forensic Science Services (DFSS), Internal Audit Report on Lab Practices, 2022.

INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 13 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

- Legal uncertainty exists regarding admissibility, especially for automated decisions made by AI[303].

✓ **Privacy and Consent Issues**

- Recording a person's voice or collecting voice samples without consent may raise privacy violations under Article 21 of the Constitution[304].
- Mass surveillance using voice biometrics (e.g., in public spaces) risks misuse and lack of transparency.

✓ **Voice Cloning and Deepfake Threats**

- AI-generated voices can closely mimic individuals, creating scope for:
  - False implication,
  - Misinformation,
  - Fabrication of evidence[305].

✓ **Low Awareness and Capacity Among Law Enforcement**

- Many investigating officers lack training in digital evidence handling and forensic voice evidence procedure[306].
- Voice data is often poorly preserved, leading to tampering or loss of evidentiary value.

✓ **Overdependence on Foreign Technology**

- Most advanced voice recognition software is developed abroad.
- Raises concerns about:
  - Data sovereignty,
  - Access control,
  - Backdoor vulnerabilities[307].

RECOMMENDATIONS

➢ **Establish Uniform National Forensic Standards**

- Formulate SOPs for voice analysis under BIS or DFSS, aligned with international protocols (e.g., ISO/IEC 19794-13 for voice biometrics).

➢ **Amend Legal Frameworks to Include AI & Biometrics**

- Update **IT Act, 2000** and **BSA Act, 2023** to explicitly address:
  - AI-generated evidence,
  - Chain of custody for digital voiceprints,
  - Admissibility rules for algorithm-based recognition[308].

➢ **Ensure Privacy Safeguards and Informed Consent**

- Introduce data protection regulations specifically for biometric voice data (under proposed Digital Personal Data Protection Act) with:
  - Consent frameworks,
  - Retention limits,
  - Right to explanation in automated decisions[309].

➢ **Invest in Anti-Deepfake Forensics**

- Fund R&D in Indian forensic labs to develop tools for:
  - Detecting synthetic speech,
  - Verifying voice integrity,
  - Authenticating metadata[310].

➢ **Capacity Building for Law Enforcement & Judiciary**

- Conduct regular certified training on:

[303] Supreme Court e-Committee, Report on Use of AI in Indian Judiciary, 2021.
[304] Justice K.S. Puttaswamy v. Union of India (2017) 10 SCC 1 – Recognized right to privacy as fundamental under Article 21.
[305] National Cyber Security Coordinator Office, Advisory on Deepfake Voice Technology, 2023.
[306] Bureau of Police Research and Development (BPR&D), Handbook on Cyber and Forensic Investigation, 2022.
[307] Ministry of Electronics and IT (MeitY), Cyber Security Policy Draft, 2023.
[308] ISO/IEC 19794-13:2021 – Biometric Data Interchange Formats — Voice Data.
[309] Suggestions from Law Commission of India, Report No. 276, *Electronic Evidence*, 2021.
[310] Digital Personal Data Protection Act, 2023 (yet to be fully enforced); provisions on biometric data.

- Voice evidence collection,
- Preservation techniques,
- Legal procedures for admissibility[311].

➢ **Promote Indigenous Development of Voice Tech**

- Encourage Indian start-ups and research institutions (e.g., IITs, IIITs) to build voice recognition tools under Digital India & Make in India initiatives[312].

6. CONCLUSION

The integration of **ICT tools, particularly voice recognition systems,** into the Indian criminal justice system signifies a transformative shift in the way evidence is gathered, analysed, and presented. Voice evidence, once seen as supplementary is now emerging as a central pillar in a wide range of investigations, from cybercrime and terrorism to corruption and organized crime.

Judicial precedents, especially from the Supreme Court, have recognized the admissibility of voice recordings, provided they meet the standards of **relevance, authenticity, clarity, and proper certification. Forensic science has responded with increasingly sophisticated methods such as spectrographic analysis, biometric matching, and AI-powered voiceprint comparison.** At the same time, these technologies must be deployed carefully to uphold the **constitutional principles of privacy, fair trial, and non-discrimination.**

Despite notable advancements, significant **challenges** remain: the absence of clear legal provisions for AI-generated voice evidence, inconsistent standards across forensic labs, risks from deepfake technology, and limited technical capacity within law enforcement. Without addressing these concerns, the risk of **miscarriage of justice** due to unreliable or manipulated voice evidence remains real.

Therefore, it is essential to **reform and modernize legal, technical, and institutional frameworks**. This includes adopting **uniform forensic standards**, ensuring **data protection**, and investing in **indigenous voice technologies** that are transparent and legally accountable. Only with such safeguards can India fully harness the potential of voice recognition systems while preserving **due process, evidentiary integrity, and public trust.**

7. REFERRENCES

1. R.M. Malkani v. State of Maharashtra, AIR 1973 SC 157; (1973) 1 SCC 471.
2. Ziyauddin Burhanuddin Bukhari v. Brijmohan Ramdass Mehra, (1976) 2 SCC 17.
3. Ram Singh v. Col. Ram Singh, AIR 1986 SC 3; (1985) Supp SCC 611.
4. Yusufalli Esmail Nagree v. State of Maharashtra, AIR 1968 SC 147.
5. Tukaram S. Dighole v. Manikrao Shivaji Kokate, (2010) 4 SCC 329.
6. Selvi v. State of Karnataka, (2010) 7 SCC 263.
7. Justice K.S. Puttaswamy v. Union of India, (2017) 10 SCC 1.
8. State of Punjab v. Baldev Singh, (1999) 6 SCC 172.

STATUTES AND OFFICIAL GUIDELINES

9. Bharatiya Sakshya Adhiniyam, 2023.
10. Information Technology Act, 2000 (with 2008 amendments).
11. Indian Telegraph Act, 1885.
12. Digital Personal Data Protection Act, 2023 (notified).
13. UIDAI (Unique Identification Authority of India), Voice Biometric Framework, 2021.
14. Ministry of Home Affairs, DFSS, Forensic Audio and Video Analysis Manual, 2022.

[311] Indian Cybercrime Coordination Centre (I4C), MHA, R&D Division Reports, 2022–2024.
[312] Judicial Academy Curriculum on ICT Evidence (National Judicial Academy, Bhopal), 2022.

15. NCRB, Annual Report on Cybercrime Trends, 2021.
16. Bureau of Police Research and Development (BPR&D), Handbook on Cyber and Forensic Investigation, 2022.
17. National Cyber Security Coordinator, Advisory on Deepfake Voice Threats, 2023.
18. National Judicial Academy, Curriculum on ICT Evidence, 2022.
19. Supreme Court e-Committee, Model Guidelines on Presentation of Electronic Evidence, 2022.
20. Indian Cybercrime Coordination Centre (I4C), MHA, R&D Division Reports, 2022–2024.

STANDARDS AND TECHNICAL DOCUMENTS

21. ISO/IEC 19795 – Biometric Performance Testing and Reporting.
22. ISO/IEC 19794-13:2021 – Biometric Data Interchange Formats: Voice Data.
23. BIS (Bureau of Indian Standards), Draft Guidelines for Biometric Systems Testing, 2021.

ACADEMIC AND RESEARCH SOURCES

24. Verma, S.K., Forensic Science and Criminology, 2015.
25. Trivedi, H.M., "Forensic Linguistics and Profiling," Journal of Indian Law & Technology, 2019.
26. Law Commission of India, Report No. 276, Electronic Evidence, 2021.
27. MeitY (Ministry of Electronics & IT), Cybersecurity Policy Draft, 2023.
28. Startup India Report, Biometric Innovation in India, Ministry of Commerce, 2021.