Chapter 2

Overview of Forensic Audio

The ability to record audio signals has been around for over a century, but only over the last four decades or so has recorded audio data been used in a forensic context. This chapter provides the background to the science and technology of forensic audio; in particular paying attention to the analysis of forensic audio material and providing an overview of the ways a recording may be tampered with. The chapter also outlines the importance of having the ability to detect a copy recording as part of a forensic examination, which forms the subject of this thesis.

2.1 Background

2.1.1 Brief History of Audio Recording

It was in 1898 that the first magnetic recorder patent was issued to Vladimar Poulson for a steel wire based recording system [8]. Various other wire based recorders followed, along with an assortment of steel tape based recorders. Paper and plastic metalised tapes were the next developments, allowing for the possibility of editing recorded material by cutting and splicing the recording medium [9]. A great deal of research and development went into improving tape formulations and backing materials for these open reel formats over the coming years.

The 1960s saw the introduction of the Compact Cassette developed by Philips. The Compact Cassette format was small and robust [10], these attributes, coupled with low production costs and modest sound quality opened a mass market in home recording. With manufacturers continually improving the technical specification of the recorders and tape formulations and with the introduction of noise reduction techniques, such as Dolby B [11]
and Dolby C [12], the Compact Cassette became a rival to the long established open reel recorder. The Compact Cassette could achieve frequency response, dynamics and distortion figures to match the much larger and clumsier open reel formats, losing out only in poorer wow and flutter specifications. The Compact Cassette is still around today, but is losing ground all the time to formats based on digital recording technology.

Digital audio recording has been a commercial proposition for about 30 years [13]. The early digital recorders utilised helical scan video recording systems to store the audio data. In this type of system, the audio waveforms were digitised using Pulse Code Modulation (PCM) techniques and processed to interface with the video recorder, taking advantage of the wide bandwidth and storage capacity of such systems [14 pp7-11].

The first domestic digital storage medium available to the consumer was the audio Compact Disc developed jointly by Philips and Sony in 1982 [14 ch8], but could only be used with pre recorded material. The Rotary Head Digital Audio Tape (RDAT) standard appeared in 1984. The RDAT is a high quality helically based digital recording system. Originally intended for the domestic market the RDAT or DAT as it became known, never found its place; however, it became a very widely used format in the professional and semi-professional audio fields [15].

The 1990s saw the introduction of two digital formats intended for the domestic market incorporating audio data compression techniques [16]. These were the Sony Mini-Disc [17] and the Philips Digital Compact Cassette (DCC) [18] recording systems. Both of these systems used perceptual encoding to reduce data storage requirements [19]. The DCC never amounted to anything in terms of sales and is now extinct [20]. The Mini-Disc also saw a slow start but is today a widely accepted consumer format. Some current models using high-density discs and high compression ratios are capable of storing up to 34 hours of stereo audio on to a single disc [21 p37].

By 1989 production had started of the Recordable Compact Disc (CDR) co-invented by Philips, Sony and Tayio Yuden. The late 90s saw the standards put forward for a Super Audio Compact Disc [22] and the Digital Versatile Disc (DVD) system, where a separate application format for audio only, known as DVD-A was introduced [23 ch6]. Both of these formats allow the storage of high-resolution digital audio data.

The past decade has also seen the introduction of solid state portable audio recording devices both for the professional and consumer markets. Today, in a multi-media society, digital audio can be stored on a variety of dedicated audio formats as well as video recording mediums and data formats designed for network transfer and use with
desk top computer systems.

2.1.2 Metropolitan Police Forensic Audio Laboratory
The Metropolitan Police Forensic Audio Laboratory forms part of the Operational and Technical Support Unit. Its primary function is to supply the Metropolitan Police and other associated police services with a forensic audio facility. The Tape Laboratory as it was originally known had its inception in the middle to late 1960s, housed in a building initially built as a police wireless receiving station [24].

Audio recordings and to a lesser extent video recordings were used to secure various forms of evidence and a need had arisen for controlled copies of this evidence to be produced for use by the investigating officers and the judicial system [25]. The wide ranging operational experience and specialised technical expertise built up by the laboratory over the coming years, was later used by the Home Office as a model to set up an equivalent laboratory to serve the needs of county police forces. The Home Office laboratory became operational in July 1979 [26]. The Metropolitan Police Tape Laboratory split into separate audio and video laboratories circa 1980 [25].

Today the Metropolitan Police Audio Laboratory is housed in a purpose built environment and supplies wide-ranging forensic audio support to the police and Criminal Justice System. The laboratory currently examines in excess of 1500 individual recordings a year. There is a wide range of analogue and digital recording formats capable of storing audio data that are regularly submitted to the laboratory for work to be carried out. These recordings originate from covert and overt police operations as well as recorded evidence supplied by witnesses, victims and suspects.

Enhancement and analysis/authenticity examinations form the two main services provided by the laboratory. Enhancement encompasses the electronic processing of speech signals that have been degraded prior to and during recording onto the recording medium [27]. The causes of this degradation or poor quality may be for a wide range of reasons, including: convolutional degradation due to the environment, masking by other acoustic or electronically generated noises, poor quality recording systems and placement of the pick up transducers. The enhancement process is carried out to improve the intelligibility of such recordings; this is in order to secure the best evidence, initially for use by investigating police officers and later for use by the courts. The enhancement procedures require the application of knowledge in time and frequency domain signal processing, acoustics, psychoacoustics and speech production. Additionally, skill and experience, as well as an extensive range of highly specialised digital processing equipment are required.
Most recordings that pass through the laboratory are enhanced, prior to a copying and reformatting process.

Forensic audio recording analysis may take many forms, for example gunshot analysis [30], [31] or extraction of dialled telephone numbers from telephone recordings, but the most common request is for forensic analysis concerning issues of integrity. Such work comes under the banner of authenticity analysis, and in a full examination may include: establishing that the recording has not been copied, examining for editing or other forms of tampering, matching a recording to recording machine, establishing chronology of recorded events.

2.2 The Authenticity Examination

Authenticity analysis is a complex forensic science, requiring a detailed auditory and instrumental examination of an audio recording. Proving the authenticity of a recording involves the verification of claims and statements associated with its content and history. The examination is carried out to determine if the recording is original and to explain the reasons for any anomalies identified; for example, discontinuities found in the recorded signal. In addition, examination can corroborate any alleged facts relating to the recording, recording system and recording environment. If possible, a witness statement is obtained from the person who made the recording containing answers to a number of questions set by the laboratory. This statement should identify exactly when, where and how the recording had been allegedly produced, using what equipment and manipulated by what methods. From the information contained within the statement, the laboratory then examines the recording and any associated recording equipment to correlate the physical data on the recording with the detailed account given by the witness. The objective of the authenticity analysis is to establish as far as possible that the recording is a true ‘acoustic representation’ of events made at a specific time and place.

In 1974 forensic audio was brought to the attention of the public and forensic science as a whole, due to an event in American history which became known as ‘Watergate’. One central issue of the ‘Watergate’ affair concerned an audio recording made at the White House in Washington DC;

“A tape recording of conversations held on June 20th 1972 in the Executive Office Building contains a section lasting 18 and one half minutes during which buzz sounds but no discernible speech sounds are heard.” [32]

An advisory panel was set up consisting of six leading scientists chosen to cover a range of technical and scientific disciplines relevant to the task of examining the recording and
associated recording equipment relating to the buzz sounds. The in-depth investigation took six months to complete and resulted in a number of conclusions:

“The erasing and recording operations which produced the buzz section were done directly on the Evidence tape... The erasures and buzz recordings were done in at least five, and perhaps as many as nine, separate contiguous segments... Erasure and recording in at least five places on the tape required hand operation of the keyboard controls on the Uher 5000 machine... Erased portions of the tape probably contained speech originally.” [32]

What this is saying, is that someone had deliberately erased an 18.5-minute section of speech from the recording under investigation! The findings from the report produced from this investigation were presented to the U.S District Court for the District of Columbia on January 15th 1974 - the rest is history [33].

Much of the forensic analysis work conducted on analogue recordings today, is still largely based on techniques used in the ‘Watergate’ investigation, [32], [34] [35, 36]. Since that time, many challenges to the integrity of tape recordings used by prosecuting authorities have been made by the defence teams representing the accused. These challenges usually take the form of claiming that malicious tampering has occurred to the recording. Articles have even appeared giving details of how to falsify recorded evidence using editing techniques [37].

An authenticity examination may be required for investigative reasons, chain of custody contradiction or more likely, as a result of a defence expert alleging that the recording has been tampered with. Allegations of tampering are more likely to be aimed at non-police made recordings. The reason for this is twofold:

- Police evidence would have been sealed at the earliest opportunity, logged in an exhibit book and stored under secure conditions. Its trail would be documented and could be followed from the point of its origin up to the point where it is presented at court. Non police made evidence is likely to have no provable antecedents before the exhibit came into police possession. This may have provided ample opportunity to tamper with the recording/s.
- Motives to tamper with recordings are often easy to find when dealing with evidence from concerned, involved or interested parties. It is more difficult to establish a motive for the police to tamper with a recording as their investigation should be based on impartiality and the penalties for tampering with evidence are severe. However, over the years the laboratory has had to analyse many police recordings for evidence of tampering after allegations had been made by
represents of the defendant.

Compared to analogue audio, digital audio appears to be a far greater challenge when trying to establish the integrity of a recording coming under forensic scrutiny. Generally, it is of widespread concern that digital audio recordings cannot be reliably authenticated. It is known that the defence teams connected with two unrelated criminal trials [38], [39], mounted sustained challenges that digital audio recordings cannot be authenticated.

2.2.1 Defining Authenticity Boundaries

What does the term authenticity or authentic mean when used in connection with the forensic examination of audio recordings either analogue or digital? The definition for an authentic recording as put forward by the Audio Engineering Society (AES), standard AES27-1996:

“A recording made simultaneously with the acoustic events it purports to have recorded, and in a manner fully and completely consistent with the methods of recording claimed by the party who produced the recording; a recording free from unexplained artefacts, alterations, additions, deletions, or edits.” [40]

The AES also state in standard AES43-2000 the following definition regarding the verification of authenticity:

“The forensic tape analyst shall examine the designated original recording along with utilizing the designated original recording device. The forensic tape analyst shall render findings that would scientifically evince that the designated original recording device recorded the designated original recording, and found no conclusive evidence of tampering, unauthorised editing, or any form of intentional deletions, material or otherwise, within the recorded content.” [41]

Finding a recording authentic shows that the events recorded are genuine and that the record can be relied upon.

The authenticity question can be split into two parts:

The first part seeks to establish that the recorded acoustic events are consistent with what is known or alleged about the incidents under investigation, including:

1. Establishing that voices heard on the recording are consistent with people allegedly present at the time of the recording.
2. Establishing that the date and time when the recording was produced matches
Chapter 2: Overview of Forensic Audio

that of the alleged date and time of the acoustic events.

3. Establishing that the recording was made at the location alleged.

External evidence in the form of chain of custody and witness testimony would assist in proving the recording’s authenticity regarding points 1 to 3. In the absence of external evidence the following may be considered. A phonetic/linguistic voice comparison could give credibility to point 1. Clearly, unless the recording contains some form of reliable date and time information as part of its data structure, point 2 can rarely be ascertained. By checking tape batch numbers it may however, be possible to show that the date when the recording had allegedly been produced was before the manufactured date of the tape. Unusual environmental cues may provide strong support for point 3.

The second part seeks to establish if the recording is original and in doing so eliminates the possibility of many forms of tampering having occurred. For the forensic examiner, the analysis to be conducted is often based on allegations of deliberate recording alteration. For this scenario the authenticity examination is mainly concerned with the following points:

1. Is the recording original or a copy?
2. Since the time of its production has the recording been edited or modified?
3. Has the recording been made on the alleged original recording machine?
4. Are the recording contents consistent with what is known or alleged?

There is a school of thought that a recording is not authentic and therefore should not be used as evidence if it is not in exactly the form it was at the time of its production and/or is not a complete record of events. Klepper states:

“It is our position then, that the tape expert must assume that any violation of the integrity of a recording taints all the information in that recording, and that a tape which has been recorded as a unit cannot be arbitrarily segmented to suit a particular argument.” [42]

Hollien takes a similar view:

“A tape recording must include a complete set of events, and nothing can have been added, deleted or changed at any time during the recording or subsequently.” [35 p163]

In this thesis the view will be taken that a recording is no less authentic having been damaged in some way, either physically or by electromagnetic means since the time of its production, providing that this has all been declared in witness statements relating to the antecedents of the tape or explained by forensic analysis. The witness declarations form
part of the provenance of the recording, where provenance goes hand in hand with authenticity. Of course the credibility of the recorded evidence might be called into question, especially if for example a large section of a potentially relevant part of the recording has been erased. However, credibility and authenticity are separate issues, the credibility of the recorded evidence is ultimately for the court to decide.

A recording found original and un-tampered with would be considered to have its integrity intact. However, this does not necessarily mean that it is also an authentic recording as the acoustic events it contains may have been staged in some way.

The verification of the authenticity of an audio recording will be considered to be a recording that is an original, having a provenance based on documentary evidence in the form of chain of custody and witness statements that can be supported by acoustic and technical evidence from a forensic examination. Proving that the recording is original eliminates the possibility of many forms of deliberate editing having occurred. The significance of any corruption occurring to an original recording after it has been produced must be judged on its own merits, taking into account the cause of the violation, where it comes in the recording, how much information has been potentially lost and any other corroborative evidence.

2.2.2 The Significance of Copy Detection in an Authenticity Examination
Establishing that the recording is the original is of primary importance in a forensic examination [43]. Basic tampering such as stops/starts, pause actions or over recordings may be found on an original recording and are usually straight forward to detect. More sophisticated tampering means that the purported original is a copy recording. By showing that the recording is an original, the application of sophisticated editing and manipulation techniques can be ruled out. It follows that the authentication of a copy recording is not possible, even when there is no evidence to suggest tampering has occurred. Something that is audibly consistent with an edit may have an innocent explanation, but the lack of audible evidence of editing is not proof that editing has not taken place. Establishing the originality of the recording provides the proof that the recording has not been edited. Further, a phonetic/linguistic anomaly found in a passage of speech having no evidence of recorder manipulation may indicate the recording has been edited post production and such a recording could generate a difference of opinion between forensic audio experts in a court of law. The significance of scientifically establishing a recording’s original status should eliminate any such conflicts.
An examination of two simple scenarios, where establishing the originality of a recording can have highly significant but opposite implications follows:

- Several unexplained high frequency clicks are identified on a recording. It has not been possible to establish their source and they occur in sections where speech is not taking place. They are most likely innocent – the result of some electromagnetically induced switching noise from a piece of local electrical apparatus. However, laboratory tests and field tests failed to establish their origin. Of course, they could be the remnant of some form of intermediate to sophisticated editing process. By making a range of further tests, it has now been possible to prove that this recording is an original. The significance is that if it is an original recording, then by implication the clicks have to have an innocent explanation.

- An examination of a recording results in nothing untoward from its contents. It appears to be one continuous recording. However, from the examination it is established that it is a copied recording. So although there is no direct evidence the recording has been deliberately tampered with, the recording is in a form that could have been tampered with. If the person making the recording has produced a witness statement claiming the recording to be original, then clearly the recording has to be treated with extreme caution. Indeed, it would be unlikely that the recording would subsequently be allowed as evidence in court.

The rule under English law regarding the admissibility of original and copied recorded evidence, falls under the best-evidence rule which is defined for recordings in Archbold as:

“Recordings are admissible as evidence provided they are shown to be both original and authentic. Copies are inadmissible in the absence of (a) proper explanation as to why the originals are not available, and (b) proof of the complete accuracy of the copies.”[44]

The fact that copy recordings are only admissible under restricted conditions goes some way to providing a safeguard against edited recordings appearing as evidence.

2.2.3 Overview of Authenticity Tests

The authentication of a recording may be investigated by looking closely at a number of areas. The procedures to be described provide an insight into the tests applied in the analysis of a typical analogue recording and are not exhaustive:
• **Visual Inspection**: A visual inspection of the cassette shell and entire length of the tape takes place. This is to look for evidence of the cassette shell having been opened or the tape having been spliced in an editing process. Additionally, the tape is examined for tape or oxide damage, which might explain audible anomalies such as dropouts or transient like sounds when the recording is played back [34].

• **Critical listening**: The recordings are copied on to a high quality computer based sound processing system designed for the manipulation of audio data. In this environment, it is a simple process to listen to the material and almost instantly find and playback different parts of interest. The listening process is used to help identify the position of, and assess in detail, various events or anomalies that may require further instrumental analysis in addition to the routine tests. Many aural cues are listened for in the process, including: variation of sounds in the background, transient like sounds of any nature, audio dropouts, audio quality changes, unnatural speech or voice characteristics, acoustic quality of the recording environment, artefacts produced by the recorder such as switching transients, hum, tones or any forms of electronically generated noise [34].

• **Continuous time-domain level plots**: Various time-plotted-against-amplitude charts are produced of parts or all of the tape under examination [45]. The plots are usually produced by chart recorders of the thermal array type. In their simplest form these charts show amplitude of the recording on the vertical axis using linear, rms or logarithmic scaling, against time on the horizontal axis. Depending on the objective of the analysis, the plot may have various weightings applied, such as A weighting or CCIR 468 noise weighting [46].

• **Time-domain analysis**: Anomalies found during a recording would be investigated in detail by zooming in on the characteristics of the relevant section of the signal’s waveform in both time and linear amplitude and hard copy printouts of the results obtained. For analogue recordings, this type of analysis would typically be applied to replay transients [47]. Replay transients are detected during replay of the recording, but are the result of the recording process. Transients are characteristic of the make and model of recorder used, and of the way in which the recorder has been operated. Transient analysis is a powerful tool used to assist in matching recording to recorder, detecting copy recordings and establishing the chronology of different recordings found on the same tape. They occur whenever the recorder is turned on into a recording mode or off from a recording mode. As an example, fig
2.1 shows a waveform left on a recording after the recorder has been turned to off. It consists of a transient relating to the record head and a further transient occurring around 500ms later that relates to the erase head.

- **Spectral analysis**: Sections of the recording are analysed in the frequency-domain usually as an averaged power spectral estimate. This form of analysis will be used to identify characteristics about the recorder’s frequency response or to identify discrete tones within the recorded material that could be characteristic of a particular recording machine. Such discrete tones, including induced mains power-line frequency, may be useful in showing that a particular recording machine has produced a particular recording or in detecting copied material [48], [49]. Grigoras, has proposed a method which uses induced power line frequency signals to establish the date and the time at which a digital audio recording had been produced [3].

- **Spatial/Time-domain analysis of erase head marks**: Analysis techniques have been developed that exploit the fact that the erase head width of an analogue recorder is wider than the record head width of the same recorder. By using specialised narrow width playback heads it is possible to analyse the waveforms found in this ‘difference region’, fig 2.2. When erase head switching waveforms are found in the difference region due to stopping and starting of the recorder, it can be concluded that they must be original and not copied, as copied erase head transients cannot exist outside the region written over by the record head [50]. Additionally, this technique can be coupled with time-domain analysis of the waveforms left by the recording head and can be used to show how a recording has been stopped, started, over recorded or copied. Elements of this type of analysis can help in establishing that a recording is consistent with having been produced on a particular recording machine.
Chapter 2: Overview of Forensic Audio

- **Microscopy**: Placing a ferro-fluid type material, consisting of micron-sized iron particles in a liquid suspension, onto a small section of the recording under analysis, produces a bitter pattern. The pattern represents the changes in magnetised levels at the surface of the tape, which can be viewed under a low power microscope [51]. An analogy of the school physics experiment of visualising the magnetic field emanating from the poles of a bar magnet by placing iron filings on to a sheet of paper placed over the magnet helps to understand the principle. The relationship between the recording head and erase head positions identified in these visible markings can be significant in showing that a recording is consistent with an original and having been produced on a particular recorder. In general, ferro fluids have poor resolution, are messy and can result in erasure of high frequency information over the region the fluid has been applied. Recently, materials with magneto optical properties, in the form of a thin film of lutetium bismuth iron garnet, formed over a glass substrate with a gold reflective layer, have become available from the Research Institute for Material Science and Technology in Russia [52]. The optical properties of the material change in response to the magnetisation on the surface of the recording under analysis. The image is viewed via a polarising microscope. The method allows high resolution images of the magnetic structure from the recorded data to be produced, having none of the disadvantages of the ferro-fluid approach [53]. As an example of this technique, an image of a stereo recording produced on a Compact Cassette recorder is shown in fig 2.3, where the two recorded tracks containing a 1kHz recorded signal can be clearly seen.
2.2.4 Forensic Voice Comparison

Forensic voice comparison or ‘speaker recognition’ started in the 1960s when a need arose for the judicial system to establish the identity of recorded voices. Speaker recognition is based on attributing a sample of a speech recording to a person using its acoustic, phonetic and linguistic features. Interestingly, two different approaches to forensic voice comparison work developed independently on either side of the Atlantic. These two methods were a ‘voiceprint’ based technique in the USA and auditory-phonetic based techniques in the UK [54].

The ‘voiceprint’ better known today as a sound spectrogram was initially developed in 1942 by Bell System Communication Services [55]. The spectrogram is a plot of the speech signal shown in the form of time on the X-axis, frequency on the Y axis and amplitude level represented by greyscale or colour. Initially these were cumbersome analogue devices [55], but today they are realised digitally and are based on processors using FFT algorithms [56]. An early comprehensive study of speech spectrograms was first published in 1947 and used as an aid for linguists interpreting speech sound patterns [57]. By the early 60s Kersta had developed a method of voice identification based on the speech spectrogram for forensic applications. He reported this technique as yielding high levels of correct identification [58]. However, in general the scientific community was sceptical of spectrographic speaker identification used for legal purposes [59], [60]. Staff from the Metropolitan Police attended a voice identification course run by Kersta in 1974 and returned unconvinced of the claims made for his techniques [61].

The road to admissibility of voice identification evidence in the USA has not been straightforward. A history of landmark cases can be found at [62]. The FBI has made extensive use of the spectrogram for forensic voice identification work [63]. However,
Chapter 2: Overview of Forensic Audio

their view on voice spectrograms is cautious, "Spectrographic voice comparison is a relatively accurate, but not positive technique for comparing an unknown voice sample with known verbatim voice exemplars" [64].

The ethics of speaker identification has often been in hot debate and these issues are discussed at [65]. Over the years the UK has adopted techniques for forensic voice identification using auditory, phonetic and acoustic analysis including sound spectrograms. However, until recently, there have been a small number of forensic experts in the UK that only use an auditory/phonetic approach. In the case of Regina v Robb [66] an expert applied a purely auditory/phonetic analysis to a voice comparison which produced a positive result. An appeal was made against the subsequent conviction, based on the expert using only an auditory/phonetic approach with no supporting acoustic analysis such as a sound spectrogram but the court of appeal upheld the original conviction.

A good introduction and background to the techniques applied in forensic voice comparison work can be found at [35 ch9-11], [62], [67-70].

2.2.5 Matching a Recording to a Location

Establishing that a recording has been produced at a particular location is based on attempting to recreate the original recording conditions from which the recording under analysis has allegedly been produced. It is only usually possible to indicate a consistency with the recording having been made at the location.

The comparison is based on modifications of the acoustic signals as they travel through the air to the recorder. For example, rooms with large hard surfaces such as uncovered floors and solid walls will be reverberant having long decay times as the sound bounces from the hard surfaces with little absorption. Rooms with carpets soft furnishings and curtains will have much shorter reverberation times, as these materials absorb greater amounts of the sound energy. In any particular room the behaviour of sound is unique and highly complex, because of the absorption and reflection properties of all the shapes within it. The situation is further complicated as sound sources in different locations within the same room will experience different modifications before they are picked up by a microphone from a fixed location [71].

From a forensic identification perspective the more unusual the sets of recorded acoustic signals are the better. Additive noises such as aircraft noise, air conditioning noise, traffic or train noise, or even the sound of a clock ticking, may help in establishing where a particular recording had been produced.
Overall, an audio recording can be thought of as a linear acoustic observation of a scene delayed in time, consisting of a number of events that are convolved with the acoustic space and recording equipment.

2.3 Tampering in Digital Times

A long departed member of the Metropolitan Police Audio Laboratory was once asked by a Judge during a trial “Is it possible that Mr X has edited this recording?” to which the reply was “Your Honour, it is possible for man to land on the moon, however I would not expect my neighbour to do it using the contents of his garden shed” [25]. A somewhat cheeky analogy perhaps, but back in the 1970s this would have been a way of informing the court of the great difficulties that would have to be overcome in successfully producing an edited recording that was undetectable by forensic examination.

The following three decades have however changed all that, with major advances in portable audio recorder technology and sophisticated computer systems now commonplace in the home. It is possible to download freeware or shareware audio processing software from the Internet, capable of carrying out editing and sound manipulation that 20 years ago would have only been achievable in a professional recording studio. Broeders reports:

“Overall the prospects for this particular branch of forensic audio (integrity and authenticity examination) is not bright. The increasingly widespread availability of relatively inexpensive digital sound processing equipment and its ease of operation makes certain types of manipulation comparatively easy to perform. If done competently, such manipulation may leave no traces and might therefore be impossible to detect from an engineering point of view.”[72]

Clearly the ability to alter a recording in a professional manner has never been easier, however there are still limitations to the exercise.

2.3.1 The Problems of Editing Speech in the Forensic Situation

Sophisticated digital editing and sound processing software can be a powerful tool when attempting to make changes to recorded material. Manipulation of audio evidence in a way that would suit the forger’s case, is however, still extremely difficult and full of potential pitfalls for the unwitting forger. For a competent attempt at tampering with audio evidence, not only is knowledge of recording, editing and signal manipulation techniques a
requirement, but also knowledge of general acoustics and forensic audio. Additionally and most importantly, enough recorded material needs to be available in order to be able to manipulate it in the way required. The problems encountered when editing a speech recording will now be discussed.

In a natural conversation spoken words are not discrete, the last syllable of one word flows in a seamless fashion into the first syllable of the next word and so on; gaps in time only appear at points of natural pauses in the speech flow. Taking a word from one point of a recording and placing it at another, would require skill and some luck in overcoming the problem of the uninterrupted speech flow. Often, two people will talk at the same time, over these sections of recording editing is not possible. Speech consists of complex sets of interacting parameters [73] and the place in the dialogue that an edited word is to be inserted should match in terms of the intonation and inflection, if that word is to sound natural and in context with its neighbours. Background correlated noises that are found on the recording are likely to result in audible discontinuities at the point of an edit. Speech that has been edited together from more than one recording, may result in the identification of a number of ‘tell-tale’ signs established from a forensic analysis; these include: level changes, background noise changes and convolutional changes caused by different recording equipment or recording environments.

Even with sophisticated digital assistance, successful editing is therefore not as straightforward as at first it might appear. It would be a fairly simple process to perhaps remove complete sections of a recording or take a sentence from one section and insert it somewhere else in the same recording. In terms of taking a recording and editing lots of little sections to produce a false set of events and to do it in a way that is not detectable, would make this at best a highly unlikely event and at worst impossible.

It is worth noting that the defence can “test the evidence” and may make allegations against a recording to be used for prosecution purposes with little or nothing to back up their claims. In the case of Regina v Sian & Sian [74] the defence persisted in making allegations of editing even though a report from their own forensic expert said this had not taken place.

2.3.2 Editing Audio Recordings

Audio editing may be categorised into three levels: basic, intermediate and sophisticated.
Chapter 2: Overview of Forensic Audio

Editing/Tampering at a Basic Level

Basic editing or tampering can be described as manipulation and modification directly to the original recording either during or after the recording had taken place, and may include the following actions:

- Editing the recording “on the fly”, achieved by stopping or starting the recorder while the recording was taking place.
- Erasure of recorded sections after the recording had taken place.
- Over recording new material on top of part of the original material.
- Physically removing, moving or adding sections to the original recorded material by cutting and splicing the tape.

Using the forensic analysis techniques previously described, modifications to the original recording produced by these methods are usually straightforward to detect.

Editing/Tampering Intermediate Level

Intermediate level editing/tampering may involve copying sections from one or more original recordings to a new recording. In order to limit the possibility of finding identifying factors originating from more than one recording machine, the final edited copy would ideally be produced on the recorder used to record the original recording. Copying a recording that has been edited using cutting and splicing techniques will eliminate the physical evidence of the splices. Editing/tampering using methods involving playback machine to copy recorder is usually detectable.

Editing/Tampering High Level

When using more sophisticated methods of tampering, the original recording would be copied on to a computer hard drive via suitable audio interface equipment such as a soundcard. The stored audio material would be manipulated using propriety editing and sound processing software. The final edited version would then be copied from the computer audio interface back to a recorder used to make the copy. One set up is shown in fig 2.4, where an analogue recording would use appropriate analogue to digital (ADC) and digital to analogue (DAC) converters forming the analogue interface. A digital recording may be able to transfer in and out of the editing system via the analogue interface or by a purely digital signal path. For both methods the original recording, playback and copying processes could all be carried out using the same recording machine.
Chapter 2: Overview of Forensic Audio

The quality of an audio signal passing in and out of an uncompressed digital audio recorder is dependent on the ADC and the DAC conversion processes and is independent of the recording medium. The digital recording medium will be a transparent carrier of the signal and copying the audio signal in the digital domain should, in a properly designed and working system, produce the copied signal bit for bit the same as the original. Copying the audio data through an analogue interface would result in the modification of the original recording by a further DAC and ADC process.

Although copying through the digital interface seems on the surface to be ideal, equipment compatibility, sampling rate and synchronisation problems or even lack of digital input/output facilities, may make the analogue interface the more likely option for the would be forger. A properly produced copy recording made from a digital master via an analogue interface would from an auditory analysis be undetectable, certainly if the original is not available for comparison, as is usually the case for a forensic examination.

One other major influential factor that may affect the choice of transfer method, is that of copyright protection incorporated into consumer digital audio recorders, known as the Serial Copy Management System (SCMS) [75]. SCMS forms part of the consumer digital interface bus data structure, described by standard IEC60958-3 [76]. An in-depth description of the significant elements of the SCMS system pertinent to a forensic investigation is provided by the author at [4]. This system has been designed to stop consumers producing digital copies of copyright material beyond a single generation. SCMS provides a set of parameters that may be useful in the forensic examination of a consumer based digital audio recording. An original recording produced on a recorder submitted for analysis would show by its SCMS status that the recorded material is copy protected and original and therefore a copy recording, purporting to be an original will be simple to detect by its SCMS status. Even if a forger is aware of this, the options

![Diagram of computer editing system](image)

**Fig 2.4: Computer based editing system.**
available to be able to produce a digital copy recording having this SCMS status is extremely limited [4].

It may be concluded that copy recordings of digital media are more likely to be carried out via an analogue interface.

### 2.3.3 Analogue Verses Digital Recording: The Forensic Perspective

Most forensic type analogue recordings are produced on cheap low budget recording machines that are not precision pieces of engineering. As such, wide tolerances exist in mechanical and electrical characteristics, producing variance in the measured parameters between machines of the same make and model, which is very useful from a forensic perspective. Digital recorders have a very high specification and tight tolerances even on small portable recorders. Head switching transients produced by analogue recorders used to match recorder to recordings, detect copies and determine chronology of recorded events, do not exist for the digital recorder. Further, compared to its analogue counterpart, uncompressed digital recordings have larger dynamic range, effectively non existent wow and flutter, flat frequency response over the pass-band, no signal degradation for digital transferred copies and imperceptible degradation for analogue transferred copies.

These factors combine to produce a very difficult task for the forensic examiner attempting to establish the authenticity or integrity of a digital recording.

### 2.4 Summary

Establishing the authenticity of a recording can be achieved by examining a range of corroborating evidence regarding both technical and non-technical issues. Technical examination of a recording concerns matters relating to the recording’s integrity. A wide variety of analysis techniques may be used to establish the integrity of a recording and are based on the peculiarities of the recording system, taking into consideration the forensic aspects of the recording.

In general, digital technology assists forensic audio engineering, by providing real time signal processing and speech enhancement systems that could not be realised in the analogue domain. Further, higher quality smaller recording formats are now realised in the digital domain allowing more scope in the production of covert recordings. Along with these benefits for the recording and enhancement side of forensic audio engineering, issues concerning authenticity analysis have been disadvantaged by the availability of cheap
easily accessible tools that can tamper with recorded material in a way that may not be audibly or instrumentally detectable. Audio recordings produced on digital formats potentially offer further advantages to a forger by making a technical analysis significantly more difficult. However, an audio recording may be considered a record of overlapping events that are difficult to disentangle by the forger and this provides opportunities for the forensic engineer.

When carrying out a technical analysis concerning the integrity of a recording, the most important and perhaps the most difficult part is to establish that it is original. Methods used for analogue analysis such as switching transient examination are no longer valid for the digital recorder. The digital recording can be copied using digital or analogue interfacing. A digital transfer can potentially produce a bit for bit copy and be identical to the original. An analogue transfer will produce small changes that will not be audibly detectable without reference to the original recording.

In conclusion, copy recordings of digital media are more likely to be carried out via an analogue interface. Detection of copied digital recordings produced by the analogue interface forms the subject of the research to be reported in subsequent chapters. As in the case for analogue recordings, it is anticipated that over time a raft of techniques will be developed for digital recording integrity analysis. This is the beginning.