

STILL RECORDING AFRICAN MUSIC IN THE FIELD

By

GERHARD ROUX

Abstract: Field sound recordings are an indispensable source of data for ethnomusicologists. However, to my knowledge there are no standards or guidelines of how this data should be captured and managed. With the progress made in machine learning, it has become vital to record data in a way that also supports the retrieval of information about the music. This article describes a model developed for field recordings that aims to aid an objective data gathering process. This model, developed through an action research process that spanned multiple field recording sessions from 2009–2015, include recording equipment, production processes, the gathering of metadata as well as intellectual property rights. The core principles identified in this research are that field recording systems should be designed to provide accurate feedback as a means of quality control and should capture and manage metadata without relying on secondary tools. The major findings are presented in the form of a checklist that can serve as a point of departure for ethnomusicologists making field recordings.

Keywords: Field recordings, data capture, audio production, music information retrieval, audio metadata, checklists.

Sixty-four years ago, Hugh Tracey (1955: 6) wrote an article about the “art of recording in Africa.” In it he expressed the wish: “I hope that those who have developed techniques of their own will contribute their ideas to this Journal, giving the benefit of their experience to others.” The aim of this article is to share a model developed for recording performances of Xhosa *uhadi* (gourd resonator bow) and *umrhubhe* (mouth bow), *isitolotolo* (jaw harp) and *umngqokolo* (overtone singing) as well as accompanied and unaccompanied vocal music ensembles in the Eastern and Western Cape provinces of South Africa.

When reading Tracey’s article where he describes the technical shortcomings of magnetic tape recording one cannot help imagine how well the modern tools at our disposal would have served his gigantic efforts. Not only do we now have access to more advanced recording equipment than in Tracey’s time, through machine-learning, we also have access to an unprecedented depth of analysis through automated processes (Kamath 2001: 1).

This article presents the techniques developed through an action research process that spanned multiple field recording sessions from 2009 to 2015. As a music technologist, I was contracted on various occasions to make field recordings for research purposes or was asked to assist ethnomusicologists in planning and executing their field recordings. Through multiple iterations, recording techniques were implemented and evaluated, and the findings of this work are presented here.

Even though there has been a technological revolution since Tracey wrote his article, it is remarkable how many of the principles he identified are still relevant. The primary difference, and core argument of this article, is that we have many more tools for objective measurement at our disposal than previously imagined. Think of it as the difference of a photograph of an object versus a three-dimensional (3D) scan of an object. The latter can provide much more information about the object and allow the exact reproduction of the object through 3D reproduction technologies. The major challenge shifts to an audio capturing process that serves this objective data gathering process. The validity of laboratory results depends on the rigour with which the measurements are taken. Since the future advancements in machine learning cannot be predicted, it is essential to take the kind of approaches in data gathering that can serve future research well.

In line with Tracey's wish the target audience for this article is ethnomusicologists, and a deliberate attempt has been made to limit the technical discussions to a level that is relevant to those who primarily make use of recordings as a tool to serve their research. In homage to Tracey's work, this article can be considered a pastiche of his 1955 article and therefore follows his structure and is written in a nearly similar narrative style.

Background

Recording technology has served ethnomusicology from its earliest existence. Almost a century ago von Hornbostel (1928: 34) remarked: "As material for study, phonograms are immensely superior to notations of melodies taken down from direct hearing; and it is inconceivable why again and again the inferior method should be used." Bartók said that "[t]he only really true notations are the soundtracks on the record itself" (Bartók and Lord 1951: 3). Initially, technology was a new way to document music, but as computational technology developed, it was realised that technology could also be applied to extract features through a process that Kassler (1966: 59) called Music Information Retrieval (MIR). Futrelle and Downie (2003: 121) define MIR as an:

[...] interdisciplinary research area encompassing computer science and information retrieval, musicology and music theory, audio engineering and digital signal processing, cognitive science, library science, publishing,

and law. Its agenda, roughly, is to develop ways of managing collections of musical material for preservation, access, research, and other uses.

The objective of MIR, according to Downie (2008: 247), “is the provision of a level of access to the world’s vast store of music on a level equal to, or exceeding, that currently being afforded by text-based search engines.” In the same way that it is almost impossible to recognise a criminal’s face on grainy security camera footage, MIR can be hampered if recordings do not possess adequate “resolution” that accurately represent as many features as possible. With the aim of “MIR-friendly” ethnomusicological field recordings, the good news is that the goalposts have not shifted since Tracey’s time: quality high fidelity recordings serve both traditional approaches and advanced feature extraction with the help of machines. A primary goal of field recordings remains high quality for preservation. The International Standards Organisation (ISO 9000: 2015) defines quality as the “degree to which a set of inherent characteristics fulfils requirement.” Therefore, it is up to ethnomusicologists to make recordings keeping their unique requirements in mind.

Recording equipment

One can agree with Tracey (1955: 6) that “[t]here are so many brands of recording equipment on the market that it would be invidious to single out any one make.” Digital recorders now replace the magnetic tape recorders of his time. A result of the film production sector that faces many of the same challenges as ethnomusicological field recordings, there is much equipment designed for use outdoors in less than optimal conditions. Mobile devices like lap-top computers, tablets and smartphones have also gained the capacity over the last decade to function as fully-fledged recording studios.

Portable multitrack recorders by various manufacturers provide the field recordist a tool to capture up to 64 discrete channels at resolutions of 24-bit / 96 kHz. Since these recorders have no moving parts, there is much less that can go wrong than with the portable tape recorders of the past. These machines also require less power. Some portable recorders record to multiple media at once, mostly in the form of Secure Digital (SD) cards or Solid State Hard Drives, to ensure data redundancy.

Many portable recorders are not only fully functioning workstations on their own but can also serve as an interface for a computer-based digital audio workstation (DAW). In many contexts, a field recording might be served better by a computer-based DAW, and in other cases, the flexibility and robustness of a portable recorder might serve the situation best.

In the world of audio recording, Bennett (2012: 123) describes a common misconception that the equipment is the solution to all problems. While Tracey (1955: 6) recommended that “[o]ne must obtain the best apparatus one can in the given circumstances”, adequate equipment is much more affordable and obtainable than in Tracey’s time. Mobile devices in combination with suitable interfaces are capable of high-quality recording. There are various interfaces or docks available for both iOS and Android tablets and smartphones (*cf* Bartlett and Bartlett 2017: 255).

Laptop computers have become so powerful in recent years that they are capable of recording hundreds of tracks simultaneously. Although latency remains a challenge in audio and video processing (Kuroda and Nishitani 1998: 1203), interfaces with internal routing capabilities allow routing schemes with virtually no latency since the computer’s software processing is bypassed. With the recently adopted IEEE standard known as Audio Video Bridging/Time-sensitive Networking (AVB/TSN) audio interfaces and computers can now exchange data through high-speed synchronised networks (Garner and Ryu 2011: 140). Transmitting audio through computer networks instead of audio cables lends itself to more flexible recording setups than what was possible in the past.

Knopoff (2004: 178) said that “[o]ne aspect which has received some critical consideration by ethnomusicologists is the extent to which field workers and their recording equipment can be seen as an intrusive element in the performance context.” One way to limit this “intrusion” during field recordings is by using AVB/TSN audio interfaces which allow the recording technician a distance of up to 100 meters from the performance.

Weight and portability

Tracey (1955: 6) considered “weight and portability” as “major considerations for country recordings.” Even though most locations where I have made field recordings was accessible with a vehicle, with minor road construction as part of the journey, light and compact gear make transport easier. The biggest challenge is always to have equipment that is versatile enough to manage unplanned scenarios while being within the luggage weight allocation on most flights. By using microphone cables with a reduced outer diameter and employing AVB/TSN audio interfaces that allow the interface to be placed closer to the microphones requiring shorter cables, a considerable amount of weight can be reduced.

A good quality microphone stand weighs around three kilograms, and to reduce this one can use a camera flash stand that weighs about 500 grams. While camera equipment usually uses ¼ inch threading, there are adaptors available to convert the ¼ inch thread to both microphone thread standards used in Europe and North

America. Flexible microphone tripods can also be used to attach microphones to whatever objects are available. Furthermore, microphones can be hung from roofs or trees or stuck to objects with adhesive tape.

Loudspeaker playback

Tracey (1955: 6) said: “[f]or most recording purposes in Africa, a loudspeaker playback is essential or the performers will feel disappointed.” Having to play back a recording to every member of the ensemble with everybody taking turns to use my headphones was learnt the hard way. It is not only for the satisfaction of the musicians recorded to have playback speak-ers on hand but it can also serve as a feedback loop that facilitates quality control. Knopoff (2004: 179), while discussing field recordings, reported that “[v]ery occasionally in my experience, following a mistake in song order or some other perceived imperfection, a singer would request that the tape be rewound, in effect editing a recorded performance on the spot.” In view of data integrity, a recording with a “mistake”, from the viewpoint of the performers, cannot be considered accurate documentation of the work and the corrective feed-back given by the performers aids the capturing of accurate data.

Feedback from the performer not only applies to learning about obvious mistakes; in my experience, other shortcomings can also be pointed out by the people who knows the mu-sic the best. Examples of feedback would be: “we cannot hear X or Y which is supposed to be the dominant voice.” This principle of the creators being the first “users” of a still uncompleted product has also been extensively applied in software development (Ries 2011: 191) and business (Blank 2013: 66). Field recordists who do not know the music might think that they are making a good recording while missing out on an important attribute. By having playback, the musicians who are familiar with the music can point out if any important features or attributes are not translating to the recording. If ethnomusicologists are going to draw conclusions based on these recordings, it is clear how valuable feedback is to ensure the integrity of the audio data captured. The use of feedback from research subjects in ethnomusicological research is not new: Arom and Fürniss (1993: 9) made use of synthesisers with micro-tuning capabilities and relied on the feedback from subjects in experiments trying to map scale models.

On a practical level, there are products without much weight penalty from several manufacturers which allow play-back. With the widespread use of mobile devices for music listening, a new market has developed for small high-powered loudspeakers. Not only do many of these speakers have Bluetooth connectivity that saves weight on cables, they also have built-in rechargeable batteries. The accuracy of frequency reproduction and levels these speakers can produce in such a small form is remarkable.

One aspect that Tracey did not discuss was the listening systems for the recording technicians. While speakers can provide valuable feedback to performers, these cannot be used while recording since it needs a degree of separation that is not possible in the field. Recording technicians, therefore, rely on headphones to monitor recordings. The value of accurate monitoring cannot be overstated. Especially when one is recording Xhosa *uhadi* (gourd resonator bow) and *umrhubhe* (mouth bow) it is difficult to balance the string and the resonator in the recording, and with inadequate monitoring, this is nearly impossible. The issue is further complicated by the fact that the design of closed headphones the technician must use to achieve isolation has inherent shortcomings in frequency reproduction (Poldy 2012: 663). While Fayte (2008: 206) advocates that one should compensate for headphones' "qualities and quirks and [...] learn to adjust for those", I have always found it impossible to "translate" a deficient frequency response and this personal observation is backed by neurological evidence as reported by Sams (*et al.* 1993: 363) that finds auditory memory to last for about ten seconds. Rather than relying on faulty auditory memory, a better solution is to address the shortcomings of headphones by modifying the frequency response using corrective software that applies a filter curve based on measurements taken from the specific headphones used (see Figure 1).



Figure 1. Software that compensates for inadequate frequency response of headphones. Screenshot by Author.

Power supply

Tracey (1955: 6) had to rely on “portable generators [or] a bank of batteries” for his power supply. While the South African power grid reaches deep into the rural areas of the Eastern Cape, there are still many areas without electricity. However, the power consumption of modern mobile electronic devices is so efficient that it is possible to do multiday recording sessions without access to mains electricity or generators. Furthermore, with the amount of sunshine in Africa, portable solar panels can be used to charge batteries. There are laptop backpacks available that have built-in solar panels. Low-cost power banks can extend the battery life of mobile devices or be used to charge devices. I have found that interfaces that can be bus-powered from the Universal Serial Bus (USB), FireWire or Thunderbolt ports of a laptop computer simplify the power supply requirements. Most portable recorders can be powered by AA-batteries which can also be recharged or replaced by new batteries. One may think of USB as a data transfer protocol, but it is also a handy power supply standard.

Microphones

Tracey (1955: 7) concluded that “[t]he question of suitable microphones and microphone characteristics must be left to individual taste and pocket.” Usually, recording technicians express a strong preference for certain brands or types of microphones for specific applications and through my research, I have also developed certain preferences. In what follows I present the microphone characteristics and techniques I find useful in field recordings.

For scientific documentation, accurate frequency response is one of the most important characteristics a microphone should possess. Microphones should stretch to at least 50 kHz, beyond the hearing spectrum of humans. While Oohashi (*et al.* 2000: 3548) found that hypersonic frequencies influence blood flow and brain activity, the reason for using microphones that can record up to 50 kHz has nothing to do with what happens supersonically, but with what occurs in the audible spectrum. According to Eargle (2005: 28), the resonance of capacitor microphones is typically in the 8–12 kHz region. This resonance is often the reason why capacitor microphones have such a “bright” and “present” sound. While this resonance is desirable in many recordings, it is especially troublesome when recording Xhosa uhadi and umrhubhe bow music, because in these instruments the higher part of the spectrum can easily mask the lower part of the spectrum (see Figure 2). By using microphones with an extended frequency response and the absence of selfresonance I have found it easier to achieve a balance between the higher and lower frequencies these instruments produce.

With regards to the operating principle, I prefer microphones that make use of radio frequency (RF) biasing since the lower electrical impedance of RF capacitor

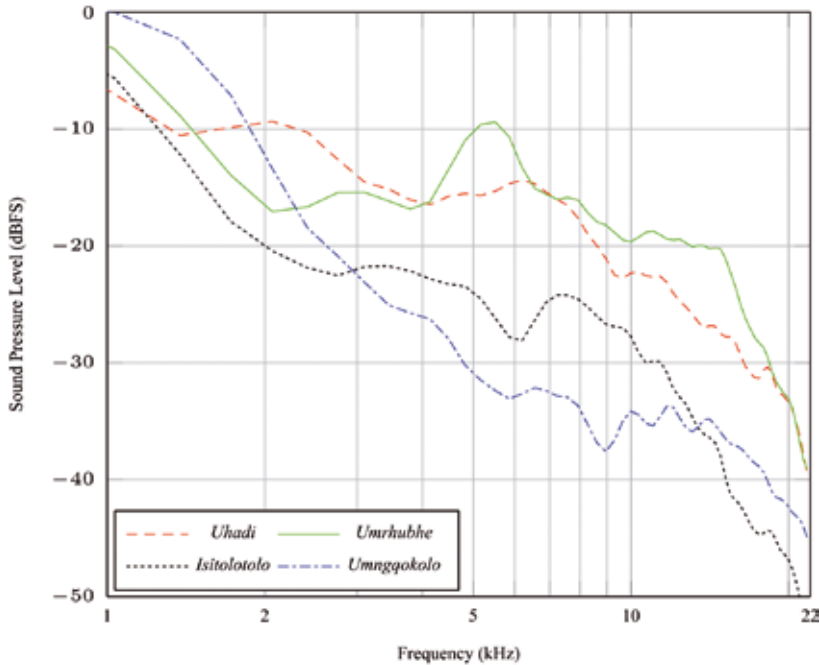


Figure 2. The long-term average spectra of the uhadi, umrhubhe, isitolotolo and umngqokolo above 1 kHz, measured under ideal conditions to show the strong overtone spectrum of the uhadi and umrhubhe. Experiment by the Author.

microphones makes them much less susceptible to short-circuits than microphones biased with direct current when used in high humidity (*cf* Ballou and Ciaudellie 2015: 512, Eargle 2005: 138). With recordings made outdoors or in buildings without climate control in the subtropical Eastern Cape, microphone short-circuits caused by the often-high humidity can cripple the process.

At the time Tracey (1955: 7) wrote his article he only had access to monophonic recording equipment; stereo became the standard in around 1968 when record companies decided to phase out monophonic productions (Fox 1968: 1). Today one can go beyond stereo with a few surround sound formats available (Gerzon 1973; Holman 2014). To accurately document a performance, recording in surround sound can produce an unmatched spatial reproduction. I try to record in surround where possible, and it is easy to produce a stereo version from a surround recording, but impossible to create accurate surround sound from a stereo recording. There are surround microphones available that are in a single enclosure, and even low-cost handheld recorders are capable of surround recording.

Performance context and production flow

Tracey (1955: 7) believed that “no control of the performers, either in time or space should be allowed”, but he did concede that recording technology placed limits on what could be practically recorded. The shortcomings of the vintage equipment necessitated that performances were shortened to fit the recording medium. Modern recording equipment enables one to interfere much less and to record in situations truer to the context. Tracey (*ibid.*:10) recognised the value of the context:

Women who sing pounding songs, and there are many such who do so every day of their lives, sing better if a pestle and mortar are on the spot and the clank of the pestle added to a voice a little short of breath makes a far better recording than the same song divorced from its occupation.

If Tracey’s advice of “no control” is followed the goal is to design flexible and agile systems that can handle the challenges of field recordings. One should also be careful of the common pitfall where technicians place a higher value on the technical than the performance fidelity. During a recording project with Ladysmith Black Mambazo in 2013, members of the ensemble told me this was their first recording, having recorded for decades, where the technician asked them how they would like to stand. Usually they are instructed.

During my research, I identified two major challenges, namely the movement of performers and the dynamics of sound produced. For uhadi, umrhubhe and isitolotolo the performer is usually seated which makes the recording, from a setup perspective, relatively easy. It is the vocal ensembles that pose more challenges because in African music the sound and movement are closely connected. Kubik (1979: 227) captured this idea well with his remark that “African music is not sound alone.” Regarding dynamics, Tracey (1955: 10) faced the same challenges with sudden loud sounds, specifically whistles, and hoped that the technology of the future would be able to “cope with” these.

To record moving musicians, I rely on various strategies depending on the situation. Firstly, I try to capture as wide a sound field as possible by using surround microphone clusters. My preferred setup is the Optimal Cardioid Triangle Surround developed by Theile (2001: 29) because all the microphones can be mounted on a single microphone stand and can easily be moved as necessitated by the situation. The recording device can also be mounted on the microphone stand.

The second strategy is to design the recording system so that everything can be moved quickly. Recording technicians are used to being in control of the

recording environment but I had to shift my approach when doing recordings for ethnomusicological research because it requires a much more reflexive or flexible approach. While this was outside my comfort zone, our colleagues in the electronic news gathering (ENG) and documentary film location recording industries face similar challenges every day and have developed useful strategies and equipment to deal with the unexpected (cf Alten 2010: 237). The robust equipment and approaches followed by ENG are suitable for music recording, especially field recordings.

During a recording in Mkankato in the Eastern Cape where I did not know what to expect, this setup proved useful. The performers decided to leave the building, and it was useful to “follow” the performers. The only intervention was to disconnect the mains power, have the recording device automatically switched over to the internal backup batteries, and the recording continued without trouble.

To “spot” specific performers one can use interference tube microphones, also called “rifle” or “shotgun” microphones, mounted on a boom pole that is typically used to record actors on film sets. As a last resort, one can fit lavalier microphones and radio transmitter packs on individual performers but this method introduces as many new problems as it solves current ones.

Managing sudden loud sounds is much easier in the domain of digital recordings because most digital portable recording devices have built-in limiters that prevent dynamic peaks in the sound from clipping the recorded waveform (cf Elliot 2005: 217). While this limiting does prevent distortion, it also modifies the transient response of the original sound and therefore may be considered an imperfect representation of the performance. A better approach is to prevent clipping from occurring in the first place. If recording at a minimum depth of 24 bits, the theoretical dynamic range of 114 dB allows the technician to leave plenty of headroom. For additional protection against overloads, microphone signals can be split using a Y-cable or transformer splitter and recorded on a secondary channel at a much lower gain (cf Yewdall 2012: 79)

Documentation of metadata

To maintain the integrity of research data collected, it is of vital importance that the data is gathered in a systematic fashion. While the field recording can be considered the primary data, it needs to be appended with additional information to ensure its validity. Not only will accurate and rich metadata serve the primary research process, it can also serve legal, archival and heritage purposes. Seeger (1992: 345) made a convincing case for the legal obligations of ethnomusicologists doing field recordings and metadata can help to distribute performance royalties

where they are due. Seeger (1986: 261) also pointed out the importance of archives in ethnomusicological research. The better the quality of the metadata, the easier re-recordings can exist as part of structured collections. Lastly, from a heritage point of view, Clayton (1999: 86) argues that “recordings can (and indeed should) be returned to the societies from which they originated, for which they serve as valuable cultural documents.” One will also rely on metadata to facilitate this “return” to the communities of origin.

I often receive field recordings from ethnomusicologists who request noise reduction or address other technical shortcomings like distortion or phase cancellations. In the process of attempting to solve these technical challenges, it helps to know what recording signal chains were used, how the musicians were positioned and where the microphones were placed. In many cases, it is impossible to provide this information, not even something as simple as a photograph taken with a mobile phone is available to aid the process. It is shock-ing that during album productions in recording studios more metadata is collected than in many scientific research processes. On the topic of gathering metadata, the field recording model presented here was inspired by an approach from the discipline of software engineering where the goal is, according to Ambler (2012: 164), to have documentation that is “just barely good enough.” This means that we possess metadata that answers all the “what”, “where”, “when” and “how” questions without having spent more resources on creating the metadata than the actual data. Rüping (2005: 4) found that there comes a point where the amount of documentation starts to hamper the usefulness thereof; too much information might make it difficult to find the relevant information.

Tracey’s observation that people rarely capture enough metadata still stands and therefore one of the goals of this research was to develop methodologies that automatically document as much metadata as possible. Luckily the broadcast and film production sectors have developed tools and methodologies, although not commonly used by the music recording industry, that can serve ethnomusicologists very well in gathering metadata.

The Microsoft WAVE file format, the most common format for recording and storing uncompressed audio files, has been extended by the European Broadcasting Union (2010) to allow various forms of metadata to be embedded in the recording itself. This metadata can include time and date stamps, details about the creator, location, as well as context-specific notes. Many portable digital recorders can embed metadata. Many even allow you to use a smartphone as a Bluetooth input device. Computer-based digital audio workstations are lagging in this functionality, but most can embed, or at least display, some basic metadata. Third-party software

may also provide advanced editing of audio file headers. These include multi-platform open source applications such as BWF MetaEdit, the free applications, Wave Agent and Metadigger and commercial applications such as Soundminer.

The primary method of exchanging information between musicians and the recording technician is through verbal communication. Record all these conversations using an omnidirectional lavalier microphone plugged directly into a smartphone. Afterwards this recording is imported into the primary recording device and serves as a document of all the verbal interaction surrounding the recording process. For example, while the musicians are listening back to a take, the secondary recording documents their comments.

The mechanism that makes it possible to keep two separate recordings synchronised is called Linear Time Code (LTC), a synchronisation protocol developed by the Society of Motion Picture and Television Engineers (SMPTE) to synchronise multiple cameras and audio recording devices on film sets. Since most portable digital recorders and computer-based workstations can be configured to synchronise to SMPTE LTC, it can serve as a pre-server of time. If synchronised to time-of-day and one records a performance at two seconds after three in the afternoon, it would be placed at 15:00:02 in the timeline of the recording device and embedded in the file. All conversations before and after the recording that were recorded on the secondary device can be imported into the session and reviewed in context. Having embedded time information can also help to distinguish between an elected pitch ex-pression and a vocalist becoming tired or an instrument going out of tune after hours of play.

Timecode can be generated by timecode generators, hardware devices created for the purpose, or through software. The accuracy of hardware generators can be improved if the source of time can be derived from Global Positioning System (GPS) satellites, for which time is provided by atomic clocks.¹ Software LTC generators can be made more accurate by synchronising computer clocks to NTP servers by using a smartphone.

Another challenge ethnomusicologists face is to separate similar instruments in ensembles when transcribing polyphonic sources. Obtaining source separation may be considered when capturing metadata since these separate sources never exist on their own; the mu-sical performance is an emergent product of the interaction between them. Nevertheless, in many research contexts, it is important to know

¹ Although GPS, in this case, is not used for its primary purpose, namely providing the location, logging accurate location data can also serve as important metadata since many field recordings do not take place in formal settlements that can be found on a map.

who plays or sings. Arom (2004: 104) achieved this source separation through overdubbing. This is where each member of the ensemble is recorded separately by playing along to a recording of the ensemble played back through headphones. Arom noted that this method has its shortcomings because musicians do not play the same when playing on their own. This challenge can be solved with piezoelectric transducers, technology used in phonograph pickups since 1935 (*cf* Jaffe, Cook & Jaffe 1971: 272). Since piezoelectric transducers only pick up the vibrations of the instrument itself and not from the air like conventional microphones, one can achieve perfect isolation while an ensemble is performing together. Piezoelectric sensors do not accurately translate the air resonance of instruments; therefore, it is not wise to rely on them for accurate sound reproduction, but it is invaluable as a source of metadata to assist with source separation. These piezoelectric transducers are recorded onto separate tracks, and afterwards the ethnomusicologist can listen to each source separately.

In the absence of formal metadata standards for ethnomusicological field recordings, it is up to ethnomusicologists to capture enough metadata to serve primary and future research. The processing of crime scenes can serve as a guideline: Law enforcement relies on every documentation method available to capture crime scene data. These include photos, videos, sketched diagrams, written descriptions, samples of matter, foot and fingerprints and even three-dimensional scans. Since nobody knows where the investigation will lead and what additional information will come to light, as much useful information as possible is documented. The same should be said for ethnomusicological research. Since one cannot predict what findings might emerge or in what direction the research will lead, it is crucial to design a metadata capture scheme as a vital part of any field recording process. A useful point of departure when designing a metadata scheme is to familiarise oneself with archival standards as developed by the International Association of Sound and Audiovisual Archives (2017) and the European Broadcasting Union (2008).

Findings

A useful way to present the findings of this research, in addition to the methodology shared earlier, is to include the checklist that was the primary outcome of this study (see Figure 3). The aviation industry has used checklists since 1935 (Myers 2016: 1). Since then checklists are used by various other disciplines such as medicine (Hales *et al.* 2008: 22), software development (Brykczynski 1999: 82) and qualitative research (Barbour 2001: 1115). Checklists help one to make fewer errors since they provide structure in a complicated task where one is often focusing on the wrong aspect at the wrong time.

FIELD RECORDING CHECKLIST

METADATA SETUP

Session notes Open
 Audio log Rolling
 Intellectual Property Rights Composition
 Performance
 Recording
 Recording Process Metadata Where?
 How?
 What?

DAW / RECORDER SETUP

Computer date & time NTP / Checked
 Computer sleep Disabled
 Master clock Verified
 Sample rate All Devices \geq 48 kHz
 Bit depth \geq 24
 Frame rate 23.98 / 24 / 25 / 30 / NA
 I/O setup All Hardware Mapped
 Buffers High
 Storage space / allocation Adequate
 Set auto backups Maximum
 Channel names Checked
 Groups Assigned
 Playlists Created
 Timecode Incoming / Generating
 Quit non-essential apps All closed

POWER SUPPLY

Incoming grid / Solar Receiving
 Devices internal batteries Charged
 Spare batteries Charged
 UPS OK
 Interruption test Survived

LINES & ROUTING

Scratch test (Echo mic settings) Done
 Talkback OK
 Stage listenback Routed
 Monitoring Calibrated

Project Briefing

Timeline Agreed upon
 Process Explained
 Communications protocol Discussed

MICROPHONE SETUP

Gain Staging

Preamplifier gain On target
 Anticipate clipping Adequate Headroom
 Gain settings Saved / Logged

Sound Check

Spatial impression OK
 Stereo impression OK
 Transparency OK
 Sound balance OK
 Timbre OK
 Freedom from noise and distortions OK

AFTER FIRST PASS

Musician/s Approval Approved
 All tracks recorded Yes
 Isolated tracks Solo'ed
 Verify clip names & Numbers Checked
 Verify playlists Checked
 Check disk allocation Checked
 Bleed None / Under control
 Plot setup Documented

END OF SESSION

Roomtone / Ambience Recorded / NA
 Delay Recorded / NA
 Local backups Synchronised
 Off-site backups Physical / Cloud
 Update session notes Completed
 Time log Updated

Figure 3. A Field Recording Checklist designed by the Author to help researchers capture research data accurately.

The checklist developed from my research is an approximation and not without flaws. George Box (1987: 74) said that “all models are wrong; the practical question is how wrong do they have to be to not be useful.” This checklist assumes that every field recording will have different challenges and is therefore purposely vague. It does, however, focus the user’s attention on the non-negotiables of scientific data capture in field recordings and I would like to emphasise a few.

By starting with the metadata capture, the checklist invites good practice. For example, it is a much more scientifically rigorous approach to record files with the correct time-stamp and a unique and descriptive file name than to rename hundreds of files called “Audio_X-Y-Z” and update timestamps.

The checklist guides the person making the recording to consider the issues like intellectual property rights. Three types of intellectual property rights have to be secured to dis-tribute a field recording. These are the rights of the composer, performers and recording technician. The right to use the composition has to be obtained by either directly obtaining the permission of the composer or by applying for the mechanical license through the publisher or royalty collection agency. Unlike composers, performers do not earn royalties on the broadcast or public playback of recordings. However, a recording cannot be exploited in any way that includes research without the express permission of the performer. In commercial recordings, the session musicians usually contractually waive the performance rights in exchange for remuneration for their services. If the researcher is not making the recordings themselves, the phonographic rights, the right of the recording itself, has to be secured. In commercial work, the phonographic rights are typically transferred to a record label who re-munerated the recording technician for the work.

From an ethical perspective, it is crucial to explain to the musicians the intellectual property licensing options available. Most ethnomusicological field recordings do not have commercial potential, but that should not be ruled out. Alan Lomax’s field recording of James T Carter and the prisoners performing “Po Lazarus” that was made at the Parchman Penitentiary in Mississippi was later used in the Cohen brothers’ film, “O Brother Where Art Thou” (2000) and formed part of the commercially successful album of the soundtrack (Ferris 2007: 136). With licensing options such as Creative Commons, it is easier to set up intellectual property ownership that is more beneficial for all parties involved than the traditional copyright models. With a Creative Commons license, one can have the composer, performers and recorder retain their rights while allowing for certain uses. For example, a Creative Commons Attribution-NonCommercial-No Derivatives license allows the work to be distributed and used for research but prevents any commercial use or adaptation of the work. In this way, a recording may be hosted in an open-access database for

research purposes and should a filmmaker wish to use the recording commercially, the composer and artists will have the option to monetise the work.

Many items of the checklist also deal with ensuring audio fidelity. Perfect fidelity does not exist: the moment a microphone is placed in a sound field, that sound field is destroyed. However, there is a big difference between an attack of a loud sound that is presented accurately over its full envelope and the same sound being limited or clipped by the recording device. While all qualities of a musical signal can be objectively measured, it is not possible to do so during a recording. Therefore, recordists make use of subjective measurement in real-time. The European Broadcasting Union (EBU) created a technical standard, *Assessment methods for the subjective evaluation of the quality of sound programme material – Music* (Tech 3286), that serves as a decision-making framework to assist technicians in evaluating audio. The field recording checklist lists the six subjective parameters that a recordist may verify. These are spatial impression, stereo impression, transparency, sound balance, timbre and freedom from noise and distortions.

Forms of data exchange such as feedback and communication also receive attention in the document. From a technical perspective, feedback, in the form of audio monitoring, is crucial for the technician. A feedback loop from the musicians is of vital importance since it serves as a verification of the data captured; the musicians will be the best people to affirm whether the recording is an accurate representation of the intention of the musical work.

Lastly, the checklist assists in logging the “how?” of the recording: who was sitting where, where the microphones were placed and how the environment influenced the sound recorded? There are numerous examples where I assured myself that I would remember the setup. However, opening the recording a few weeks later it is remarkable how many details one may not recall. This is not only bad for one’s research but also frustrates the future us-ability of the data by others.

Future research

Considering that ethnomusicologists presently must design their own field recording methodology, there is a lot of scope for what Topp Fargion (2009: 75) called “a more responsible, applied discipline and [...] best practice for making recordings that can be more easily disseminated.” The value of standardised processes, according to Watkins (2009: 7), is that procedures can be reused and shared with other practitioners in the field. Therefore, there is a need for future research to develop a standard practice for field recordings and ontologies of musical features specific to African music. The approach suggested in this article is an effort to lay the foundation for an agreed upon standard.

The objective of the so-called “semantic web”, proposed by Berners-Lee, Hendler and Lassila (2002: 25), is to turn the data on the internet, intended to be seen and heard by humans, into data that is useful for computers as well. To achieve this, Berners-Lee (2006: n.p.) proposed a model called the *Five-star Open Data Plan*. This model may serve as a starting point for standardised approaches to field recording capturing, storage and sharing. The first goal of this model is to make data available on the internet in any format. The second aim is to “make it available as structured data”, for example, as a MIDI file rather than a scan of a transcription. Thirdly data should be made available “in a non-proprietary open format.” The fourth aim is to assign objects with unique identifiers that can be used as a point of reference, and lastly, data should be linked to other data “to provide context.”

Obviously, the three types of intellectual property rights concerning field recordings, composition, performer and phonographic, should be secured before distributing any recording. Furthermore, the proper ethical practice should be followed. Even if all the intellectual property rights were rightfully secured, it should be disclosed to musicians where the recordings will be used and written consent must be obtained from them.

Conclusion

Through the sharing of information through simple interactions, advanced findings can emerge. One sees an example of this type of emergence in nature: while no termite possesses the knowledge to construct a mound, through the phenomenon called swarm intelligence, impressive structures arise through the simple interactions of many individual termites (*cf* Garnier, Gautrais and Theraulaz 2007: 3). To be part of the “swarm” and to reap the benefits of the “swarm-intelligence”, ethnomusicologists ideally need to make field recordings that are compatible with the data of others. Researchers who are failing to do so run the risk of isolating themselves in a world where machine learning will play an ever-increasing role in analysing music.

Lastly, there is also an element of social responsibility in making recordings and capturing metadata of high quality: it serves the musicians. Tracey (1955: 11) remarked:

Anything which will add to the quantity and quality of African recordings [...] will give recognition to the talents of proficient African musicians throughout the continent.

Through the application of the findings presented here, ethnomusicologists will mitigate some of the common challenges faced in making field recordings. The core principles identified in this research are that field recording systems should be

designed to provide accurate feedback as a means of quality control and should capture and manage metadata. Quality data captured through the methodologies suggested in this article will enhance the research potential of recordings. No researcher ever wished for fewer data points or reduced accuracy. Lastly, with the technological evolution not yet at its peak, the tools and processes of the future and the creative solutions devised by field recordists of the future promise ever increasing potential for research applications of field recordings.

References

Alten, Stanley

2010 *Audio in Media*, Ninth Edition. Boston, MA: Wadsworth.

Arom, Simha, and Susanne Fürniss

1993 "An Interactive Experimental Method for the Determination of Musical Scales in Oral Cultures: Application to the Vocal Music of the Aka Pygmies of Central Africa." *Contemporary Music Review* 9 (1-2): 7–12.

Arom, Simha

2004 *African Polyphony and Polyhythm: Musical Structure and Methodology*. Cambridge: Cambridge University Press.

Barbour Rosaline S.

2001 "Checklists for Improving Rigour in Qualitative Research: A case of the Tail Wagging the Dog?" *British Medical* 322 (7294): 1115–1117.

Ballou, Glen, Joe Ciaudelli and Volker Schmitt

2015 "Microphones." In *Handbook for Sound Engineers*, Fifth Edition, Glen Ballou, ed. 597–702. Burlington: Focal Press.

Bartlett, Bruce and Jenny Bartlett

2016 *Practical Recording Techniques: The Step-by-step Approach to Professional Audio Recording*, Seventh Edition. New York: Routledge.

Bartók, Béla and Albert B. Lord

1951 *Yugoslav folk music*, Vol. 1. Albany: State University of New York Press.

Bennett, Samantha

2012 "Revisiting the 'Double Production Industry': Advertising, Consumption and 'Technoporn' surrounding the Music Technology Press." In *Music, Business and Law: Essays on Contemporary Trends in the Music Industry*, Antti-Ville Kärjä, Lee Marshall and Johannes Brusila, eds. 117-145, Helsinki: IASPM Norden & Turku, International Institute for Popular Culture.

Berners-Lee, Tim, James Hendler, and Ora Lassila

2001 "The Semantic web." *Scientific American* 284 (5): 28–37.

Berners-Lee, Tim

2006 [Online] "Linked Data." Available at <https://www.w3.org/DesignIssues/LinkedData.html> [Accessed on 24 March 2017].

Blank, Steve

2013 "Why the Lean Start-up Changes Everything." *Harvard Business Review* 91 (5): 63–75.

Bramer, Max

2016 *Principles of Data Mining, Third Edition*. London: Springer.

Box, George and Norman Draper

1987 *Empirical Model-Building and Response Surfaces*. New York: Wiley.

Brykczynski, Bill

1999 "A Survey of Software Inspection Checklists." *SIGSOFT Softw. Eng. Notes* 24 (1): 82–89.

Cano, Pedro, Eloi Batlle, Ton Kalker, and Jaap Haitsma

2005 "A Review of Audio Fingerprinting." *The Journal of VISI Signal Processing Systems for Signal, Image and Video Technology* 41 (3): 271–284.

Clayton, Martin

1999 "A. H. Fox Strangways and The Music of Hindostan: Revisiting Historical Field Recordings." *Journal of the Royal Musical Association* 124 (1): 86–118.

Downie, J. Stephen

2003 "Music Information Retrieval." *Annual Review of Information Science and Technology* 37 (1): 295–340.

2008 "The Music Information Retrieval Evaluation Exchange (2005–2007): A Window into Music Information Retrieval Research." *Acoustical Science and Technology* 29 (4): 247–255.

Durey, Adriane Swaim, and Mark A. Clements

2002 "Features for Melody Spotting using Hidden Markov Models." *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*. II-1765–II-1768.

Eargle, John

2012 *The Microphone Book: From Mono to Stereo to Surround, a Guide to Microphone Design and Application*. Burlington: Focal Press.

Elliott, Simon T.

2005 "Sound Devices 722 Digital Audio Recorder." *Bioacoustics* 15 (2): 217–219.

European Broadcasting Union (EBU)

1997 [Technical Standard] *Assessment Methods for the Subjective Evaluation of the Quality of Sound Programme Material – Music*, EBU Tech 3286. Geneva, Switzerland: EBU.

2008 [Recommendation] *Digitisation of Programme Material in Audio Archives*. EBU R 105-2008. Geneva, Switzerland: EBU.

2011 [Standard] *Specification of the Broadcast Wave Format (BWF): A Format for Audio Data Files in Broadcasting. Version 2.0*, EBU-TECH 3285 v2. EBU: Geneva, Switzerland.

Fayte, Buster

2008 "The Complete Home Music Recording Starter Kit: Create Quality Home Recordings on a Budget!" Indianapolis: Que Publishing.

Fox, Hank

1968 "Stereo Rattles Stations—Mfrs. Strangle Monaural: Phasing Out to Choke Supply." *Billboard Magazine* 74 (1): 1,8.

Futrelle, Joe, and J. Stephen Downie

2003 "Interdisciplinary Research Issues in Music Information Retrieval: ISMIR 2000–2002." *Journal of New Music Research* (32) 2: 121–131.

Garner, Geoffrey M. and Hyunsurk Ryu

2011 "Synchronization of Audio/Video Bridging Networks using IEEE 802.1 AS." *IEEE Communications Magazine* 49 (2): 140–147.

Gerzon, Michael A.

1973 "Periphony: With-height Sound Reproduction." *Journal of the Audio Engineering Society* 21 (1): 2–10.

Hales, Brigitte, Marius Terblanche, Robert Fowler, and William Sibbald

2008 "Development of Medical Checklists for Improved Quality of Patient Care." *International Journal for Quality in Health Care* 20 (1): 22–30.

Holman, Tomlinson

2014 *Surround Sound: Up and Running, Second Edition*. Burlington: Focal Press.

Huber, David M.

2007 *The MIDI Manual: A Practical Guide to MIDI in the Project Studio, Third Edition*. Burlington: Focal Press.

International Association of Sound and Audiovisual Archives (IASA)

2017 IASA-TC 03 "The Safeguarding of the Audiovisual Heritage: Ethics, Principles and Preservation Strategy." London, UK: International Association of Sound and Audiovisual Archives.

International Standards Organisation (ISO)

2015 9000:2015 "Quality Management Systems: Fundamentals and Vocabulary." Geneva, Switzerland: International Organization for Standardization.

Jaffe, Bernard, William R. Cook Jr. and Hans Jaffe

1971 *Piezoelectric Ceramics*. London: Academic Press.

Juang, Biing Hwang and Laurence R. Rabiner

1991 "Hidden Markov Models for Speech Recognition." *Technometrics* 33 (3): 251–272.

Kamath, Chandrika

2001 "On Mining Scientific Datasets." In *Data Mining for Scientific and Engineering Applications*, Vol. 2. Robert L. Grossman, Chandrika Kamath, Philip Kegelmeyer, Vipin Kumar and Raju Namburu, eds. 15–22. Dordrecht: Kluwer Academic Publishers.

Kassler, Michael

1966 "Toward Musical Information Retrieval." *Perspectives of New Music* 4 (2): 59–67.

Klapuri, Anssi P., Antti J. Eronen, and Jaakko T. Astola

2006 "Analysis of the Meter of Acoustic Musical Signals." *IEEE Transactions on Audio, Speech, and Language Processing* 14 (1): 342–355.

Knopoff, Steven

2004 "Intrusions and Delusions: Considering the Impact of Recording Technology on the Subject Matter of Ethnomusicological Research." In *Music Research: New Directions for a New Century*, Michael Ewans, Rosalind Halton, and John A. Phillips, eds. 177–186. Buckinghamshire: Cambridge Scholars Press.

Kubik, Gerhard

1979 "Pattern Perception and Recognition in African Music." In *The Performing Arts: Music and Dance*, John Blacking, Joann W. Kealiinohomoku, eds. 221–49. The Hague: Mouton Publishers.

Kuroda, Ichiro and Takao Nishitani

1998 "Multimedia processors." *Proceedings of the IEEE* 86 (6): 1203–1221.

Kyriakakis, Chris, Panagiotis Tsakalides and Tomlinson Holman

1999 "Surrounded by Sound." *IEEE Signal processing magazine* 16 (1): 55–66.

Landau, Carolyn, and Janet Topp Fargion

2012 "We're all Archivists Now: Towards a More Equitable Ethnomusicology." *Ethnomusicology Forum* 21 (2): 125–140.

Leskovec, Jure, Anand Rajaraman and Jeffrey David Ullman

2014 *Mining of Massive Datasets*, Second Edition. Cambridge: Cambridge University Press.

Logan, Beth

2000 "Mel Frequency Cepstral Coefficients for Music Modeling." *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*. 1–11.

McAfee, Andrew and Erik Brynjolfsson

2012 "Big Data: The Management Revolution." *Harvard Business Review* 90 (10): 61–67.

Myers, Paul

2016 “Commercial Aircraft Electronic Checklists: Benefits and Challenges.” *Inter-national Journal of Aviation, Aeronautics, and Aerospace* 3 (1): 1–10.

Oohashi, Tsutomu, Emi Nishina, Manabu Honda et al.

2000 “Inaudible High-frequency Sounds Affect Brain Activity: Hypersonic Effect.” *Journal of Neurophysiology* 83 (6): 3548–3558.

Poldy, Carl A.

2012 “Headphones.” In *Loudspeaker and Headphone Handbook*. Third Edition. John Borwick, ed. 585–692. Oxford: Focal Press.

Ries, Eric

2011 *The Lean Startup*. New York: Crown Business.

Rüping, Andreas

2005 *Agile Documentation: A Pattern Guide to Producing Lightweight Documents for Software Projects*. New York: John Wiley & Sons.

Sams, Mikko, Riitta Hari, Josi Rif, and Jukka Knuutila

1993 “The Human Auditory Sensory Memory Trace Persists about 10 sec: Neuromagnetic Evidence.” *Journal of Cognitive Neuroscience* 5 (3): 363–370.

Saunders, John

1996 “Real-time Discrimination of Broadcast Speech / Music.” *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’96)*. 993–996.

Schneider, A.

2001 “Sound, Pitch, and Scale: From ‘Tone Measurements’ to Sonological Analysis in Ethnomusicology.” *Ethnomusicology* 45 (3): 489–519.

Schroeder, Manfred R. and Bishnu S. Atal

1985 “Code-excited linear prediction (CELP): High-quality Speech at very Low bit rates.” *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’85)*. 937–940.

Seeger, Anthony

1986 “The Role of Sound Archives in Ethnomusicology Today.” *Ethnomusicology* 30 (2): 261–276

1992 “Ethnomusicology and Music Law.” *Ethnomusicology* 36 (3): 345–359.

Theile, Günther

2001 “Natural 5.1 Music Recording based on Psychoacoustic Principals.” In *Audio Engineering Society 19th International Conference: Surround Sound-Techniques, Technology, and Perception*. 1–45. Germany: Schloss Elmau.

Topp Fargion, Janet

2009 “For My Own Research Purposes?: Examining Ethnomusicology Field Methods for a Sustainable Music.” *The World of Music* 51 (1): 75–93.

Tracey, Hugh

1955. "Recording African Music in the Field." *African Music* 1 (2): 6–11.

Tzanetakis, George, and Perry Cook

2002 "Musical genre classification of audio signals." *IEEE Transactions on Speech and Audio Processing* 10 (5): 293–302.

Von Hornbostel, Erich Moritz

1928. *African Negro Music*. *Africa* 1 (1): 30–62.

Watkins, John

2009 *Agile testing: How to Succeed in an Extreme Testing Environment*. Cambridge: Cambridge University Press.

Yewdall, David

2012 *The Practical Art of Motion Picture Sound, Fourth Edition*. Waltham: Focal Press.
