With my own ears: the ethics of sound in non-fiction film and TV

Leo Murray
Murdoch University
l.murray@murdoch.edu.au

Abstract

Sound in non-fiction is routinely manipulated in order to accentuate, suppress or replace some aspects of sound to present a clearer, more intelligible, exciting, involving or dramatic representation to the audience. This manipulation, or potential for manipulation, indicates the need for a discussion on the ethical dimensions of such a practice. Ethical models across a range of disciplines such as journalism, documentary and history can inform the approach applied to sound. This paper examines some of the realities and complexities of the non-fiction soundtrack. Given its potential to mislead, misrepresent and misinform, some practical guidelines for the ethical use of sound in non-fiction are suggested.
In the formative years of cinema the distinction between fiction and non-fiction was not yet clear. As Nichols (1995) points out, the adoption of fiction techniques and narrative structure has been a part of documentary since the pioneering films of Flaherty’s 1922 film Nanook of the North (2010). Daniel Levi has described the practices of early cinema as allowing “for a lot of two-way traffic across a weak ontological frontier” (cited in Juhasz and Lerner 2006: 22). Contemporary sound practice in non-fiction adopts many of the same techniques used in dramatic production straddling this same ontological divide. From the earliest days of non-fiction production sound has been used to perform multiple simultaneous functions using multiple sonic elements. The amalgam of authentic synchronous recordings, supplementary sound effects, disembodied voiceover, narration and music allows sound to fulfil various narrative, dramaturgical, referential and grammatical roles (Nichols, 1983: 20; Balázs, 1985: 199; Cavalcanti, 1985: 102). In addition to the spoken words of its participants and narrator the soundtrack also seeks to legitimise and support the images, to supply and augment a sense of realism, to add dramatic weight and to steer an emotional interpretation by the audience. Leaving aside the wider questions about what constitutes documentary as a whole, we may then ask what, if any, ethical duty is there on the part of the sound practitioner in the production of non-fiction material?

Some of the earliest uses of sound in non-fiction film are examples of the addition of sound to image to present a reality which never existed in the first place. Newsreel footage which includes the sound of the crowd at a football match or the sound of a fighter plane diving in wartime would almost certainly have originally been produced without sound, and had the sound elements added some time afterwards – narration, music and sound effects. The British Pathé online archive contains a great many films which contain sound, but which in all likelihood would have been filmed silent with the sound added after the fact. The British Pathé newsreel film Time To Remember – Enough Of Everything... (British Pathé, 1917) includes voiceover, sound effects (explosions and
gunshots) and music though the film was made in 1917, long before sync sound was widespread. In *Air Battles Over Europe* (British Pathe, 1944) there are several sequences of air battles which would have presented enormous, if not insurmountable difficulty for the recording of sound yet which contain the sounds of aircraft along with music and narration. The commentary declares that “concealed in cameras built into the wings of their aircraft is the evidence of victory”. Such visual evidence is augmented by sound almost certainly added after the fact, though this is of course difficult to confirm, such is the nature of sound production.

This audiovisual newsreel seen today is presented as a newsreel of the time, but no indication is given as to whether it was made with synchronous sound originally and, if not, when the sound was added. Looking back to newsreels of the early twentieth century there is a tendency for the historical and archival material to gradually acquire authenticity as it ages. Though the soundtrack has clearly been added afterwards it retains its archival quality since that is how the newsreels came to be presented to their contemporary audiences, though the added sonic elements are, strictly speaking, fabrications in which the soundtrack is purported to be something it is not.

**Sound authenticity and realism**

The normal processes of filmmaking routinely rely on disguised artifice to create replicas of reality in which “[t]he high degree of artifice produces an ideal appearance of absence of artifice” (Wood, 2000). Shots are edited to make sense where there was little sense before. Things are edited out, and edited in. Shots are framed to highlight certain things and avoid others. Reactions are inserted where reactions were not strictly sequential or consequent. Each film element is “subsumed by the needs of the story” (Metz, 1974: 45). The soundtrack is recorded but also created after filming is complete. How can we know whether what we are seeing and hearing is authentic? Any viewer of a documentary may at some stage question the material presented to them: What happened before that cut? Did these two events really happen one after another? What is outside the frame?
Prior to the technology of recording all sounds were authentic: in R. Murray Shafer’s words “every sound was uncounterfeitable, unique” (Shafer, 1993: 90). Our idea of a faithful recording carries with it the idea that it is somehow unchanged from the reality, but how can there be one faithful recording when each point of view yields a different perspective? Is a faithful recording of a piano performance, for example, that recorded from the perspective of the pianist playing it, from that of the conductor, or from that of someone seated in the audience? Each perspective will sound different yet each can claim to be authentic. Realism in film is often in practice the illusion of realism. Somewhat ironically, real realism can itself sometimes sound unrealistic.

Fidelity in terms of sound most frequently relates to the idea of its truthfulness or accuracy. For Bordwell and Thompson (1985: 283), for example, fidelity “refers to the extent to which sound is faithful to the source as we conceive it”. For others the term fidelity contains some notion of quality. Ultimately the best judges of fidelity or authenticity are the filmmakers and participants themselves. This illustrates the curious notion of authenticity when discussing nonfiction and nonfiction sound in particular. Does authenticity in this context mean not changing the recording? Can it mean removing the parts that make it appear inauthentic, or combining different elements to recreate a more representative version of (sound) events?

**Sound in sport, news and current affairs**

Television sport may seem an unlikely genre to examine for its authentic use of sound, but part of the difficulty in analysing the nonfiction soundtrack is the difficulty for anyone other than those engaged in the practice of determining whether the sound that is used is in fact real at all, much less unmediated. Television sport is perhaps best explained as a composite model of television genres suggested by Gary Whannel (1992: 56) lying somewhere between drama, journalism and light entertainment. Bill Whiston described his role as a television sport sound supervisor in terms more readily associated with dramatic production: “We try and enhance the experience. We tread the middle road between what’s real and what’s unreal” (Andrews, 2012). A judgement is made by the
production team on how best to present the coverage. A different standard may well be applied if the production is viewed as drama or entertainment as opposed to a more journalistic and objective version of the actual events.

Consider the perspective of a sound producer working on factual productions such as sport, natural history documentary, live concert or political convention. Each production could equally be treated as drama, journalism or light entertainment. Adopting a dramatic approach would give the license to augment sounds, or to highlight certain sounds in order to create a perspective which encouraged a more immersive experience. A journalistic approach may tend to a representation which encouraged balance and fairness though it may sometimes tend towards over-simplification. A light entertainment approach may emphasise the personalities of the presenters as well as the participants and may downplay context and the seriousness of the event for those engaged in it.

In television sports sound is frequently fabricated or augmented in some way because of the simple logistical problem of recording sound over a wide geographical area, or because of noise pollution during the event itself. In horseracing, for example, the track may be several kilometres long and multiple cameras are required to produce images at all parts of the course. There is no direct microphone equivalent of a zoom lens, yet the sound that routinely accompanies these images is relatively consistent, appears realistic and therefore matches our expectation. This would either suggest that somehow a microphone is moving alongside the galloping horses in order to maintain a consistent sound perspective, or that one or more microphones is actually attached to one or more of the horses or riders. The sound of horseracing is frequently a library sound effect (such as the sound of buffalos stampeding), particularly for long shots where there is no microphone nearby (Andrews, 2012).

Rather than approaching the analysis and discussion of television sport as an unmediated depiction of events as they appear to be it is useful to treat it from the perspective of the production team whose job it is to create the production. In Designing Sound and Silence (2006) Maasø describes the
process of “taming wild sound” of televised sporting events, comparing the sound components of television broadcast from the early 1990s to the present day. Part of the change concerns the addition of artificial sounds as well as natural ones. This augmentation happens in a number of ways. As early as the 1992 Olympics digital samplers were employed to play, as required, pre-recorded sounds of skiers (Maasø, 2006: 25-26), which would otherwise be impossible. The sounds of the skiers could be recorded in practice sessions which could then be played back on demand in synchronisation with live images where crowd or other noises such as PA announcements or vehicles might render the live sounds of skiing unclear or inaudible, or for parts of the course where microphones could not pick up sound. “Sampling thus provided sounds to on-screen events where silence was the alternative in earlier broadcasting” (Maasø, 2006: 26). By prerecording and some sonic sleight of hand, the accompanying sound can then be provided to give a cleaner, or more exact version of the sound which is expected but which could not be achieved with completely authentic means.

The use of sound in a journalistic context can also highlight the difficulty in pinning down the proper use of authentic sound. James Batcho (2012) highlighted the problem of sound representation for news reporting in what has come to be known as The Dean Scream (CNN, 2004). The television coverage of the live event contrasted markedly with the lived experience of the audience who were present. Whilst the recording is faithful to one of the speaker’s words, the television requirement to highlight intelligibility over faithfulness is misleading through its omission of context. The perspective of this sound recording, though entirely authentic in one sense, when used in combination with the similarly exclusive perspective of the camera provides a misleading (and, in this case, damaging) representation that did not correspond with the experience of those individuals present at the event (Murray, 2010: 134).

Authenticity is not, then, the only problematic area for sound in non-fiction. Along with the actuality of the recording, perspective and context play a pivotal role in determining whether the
representation can be said to be truthful. Where a photograph shows an image bounded by a rectangular frame, by definition material exists outside that frame. Sound is a less obviously exclusively defined representation to the audience. Similarly picture edits in television and film, though not consciously noted by the audience, indicate the places where a particular sequence begins and ends. By comparison sound editing (layering, replacement, augmentation, addition) is designed to be inaudible and undetectable and makes scrutiny or analysis for the audience a difficult if not impossible task.

Many of the choices in sound practice are concerned with clarifying particular sounds so that they are interpreted in a certain way to convey a sense of realism rather than represent actuality. From the perspective of the audience it is reasonable to expect that the sound which accompanies television images, when presented as factual programming, is what it appears to be. Our determination of whether something is truthful or real depends on our interpretation of what we are presented with. In the absence of conflicting information we will seek out an economical interpretation and accept it as being real. From the perspective of the program makers the aim may be slightly different: a soundtrack which matches the images. To present a realistic, clear, intelligible and engaging program may involve some augmentation or other manipulation of the soundtrack which, though not entirely authentic, does not seek to mislead the audience. Without foreknowledge the audience would never be aware that the sound they hear is not a truthful, authentic or original sound. Through synchronisation with the image it has become the real sound.

The truth-telling role of sound

When filmmakers are dealing with documentary material there is an ethical duty to adhere to a degree of authenticity; however, differences exist over how to define authenticity in this context. Brian Winston argues convincingly that the focus of documentary ethics should be squarely on the relationship between documentary maker and participant as opposed to “an amorphous ‘truth-telling’ responsibility to the audience” (Winston, 2000: 5). Whilst it is hard to disagree with
Winston’s primary concern, it would be a mistake to dismiss this secondary truth-telling responsibility of non-fiction. Adopting a quasi-legal determination of a victimless manipulation in a documentary eschews the potential for a variety of other forms of misrepresentation, and the rewriting of history.

Positivist traditions of documentary filmmaking – scientific, objective, fact-finding, and observational – do little to address the problems inherent in documentary filmmaking (Williamson, 2006: 83). The objectivity, reality and truth are pursued, yet we know that what we see as a final product is rarely completely objective, real or truthful. Camera choices, sound recording choices, picture editing, sound editing and music choices each seek to manipulate the way the audience will receive the text. Their purpose is to manipulate the audience. This manipulation, through contrivance, selection and expressivity, serves the purpose of telling the story in a way that makes it as believable as possible.

Authentic sound, which can be defined here as ‘sound recordings of what they purport to be’ can be considered a benchmark for authentic non-fiction sound, whilst at the same time acknowledging that ‘authenticity’ is in itself a slippery term since a recording of any particular event can be radically different depending on the perspective or placement of the microphone, or the sound’s context. It is the choices in sound recording, editing and mixing that largely determine whether that sound representation taken as authentic is actually of the thing it purports to be. Augmentations and manipulations inevitably occur, as we have seen. It is only those who have actually done the work in creating the finished result who know whether it is a truthful representation, a more readily believable facsimile or a dramatic interpretation of the actual events.

**Understanding reality and truth**

Our acceptance or questioning of the reality and truthfulness of particular representations is dependent on our ability to recognise them as representations of reality; realism as opposed to actuality. Where sonic representations are not recognised as mediated signs our critical faculties are
to some extent bypassed and instead we accept the representational code as reality. Reality and truth were important subjects for C.S. Peirce, the founder of the philosophical school of pragmatism, since we can only determine reality and truthfulness from the signs we are given. For Peirce, the concept of abduction, being one of three modes of reasoning concerned with possible inferences, as opposed to induction (probable inferences) and deduction (necessary inferences), is pivotal to understanding how we make sense of the world (Peirce, Hartshorne and Weiss, 1960: 1.65-68). Abduction is the process fundamentally important to determining what one believes to be real or true (Peirce, 1976: 4:37-38):

Abduction is reasoning which professes to be such that in case there is any ascertainable truth concerning the matter in hand, the general method of this reasoning, though not necessarily each special application of it, must eventually approximate to the truth... there is only a relative preference between different abductions; and the ground of such preference must be economical. That is to say, the better abduction is the one which is likely to lead to the truth with the lesser expenditure of time, vitality, etc.

In Peirce’s view, when presented with a reasonable representation which satisfies the criteria for lived experience, and which is therefore an economical explanation, and which is given in the absence of an alternative, we interpret it as reality. For example, when provided with an image and sound which satisfy the criteria to be both synchronous and plausible we accept the sound/image as reality. Many of the choices in sound practice are concerned with clarifying particular sound-signs so that they are interpreted in a certain way to convey a sense of realism rather than represent reality. It is reasonable to believe that the visible object is the one which makes the sound, though we know this is often not the case. Without foreknowledge we would never be aware that the sound was not a truthful, original sound. Through synchronisation it has become the real sound. Our determination of whether something is truthful or real depends on our interpretation of what we are presented with. In the absence of conflicting information we will seek out an economical interpretation and accept it as being real.
As an audience we would hope that what we are seeing is somehow a truthful representation of real events, and if what we are seeing is a recreation, we would expect some explicit reference to the fact that it is not truthful. Where television sequences are edited together to create a narrative from non-consecutive sequences there is a good chance of detection. Writing about the “Crowngate Affair”, in which a documentary promo sequence on the Queen was edited to suggest that she was annoyed and stormed out of a photo shoot, BBC television producer Julian Mercer points out that “[a]ny medium in which hours of raw material are boiled down to minutes of consumption inevitably requires some manipulation. The challenge is to ensure the process engages but doesn't deceive its audience” (Mercer, 2011). To the participants who witnessed the events first hand there was an obvious mismatch with the edited sequence. For the audience, the fact that the sequence is edited from different shots also calls attention to the fact that the sequence is a construction, created from elements which are not necessarily contiguous. If the edited sequence contained only sonic manipulation (such as the addition of an off-screen comment from earlier in the day) it is unlikely that the manipulation would be as easy to detect.

Audience expectation plays a role in determining the elements chosen for use in a soundtrack. We do not usually notice artifice in the soundtrack when we are presented with what appears to be a reasonable representation which matches what we are seeing. Further, we can become aware - and even distrustful - of the soundtrack when it presents a real and authentic synchronous if it does not meet with our preconceptions of what things should sound like (Murray, 2010: 133). The ethical question this raises is whether truthful representation might be stretched sufficiently far to enable some small misrepresentation to better tell a larger truth. How much artistic license is too much in the production of a non-fiction soundtrack?

*A need for guidelines?*

In the specific area of film, *The Production Code* (MPPDA, 1930-1967) attempted to instil an ethical framework into the previously laissez faire film industry, inadvertently underscoring the power of
film, unlike other media, to present a convincing replica of reality. The first general principle of the code was that “No picture shall be produced that will lower the moral standards of those who see it. Hence the sympathy of the audience should never be thrown to the side of crime, wrongdoing, evil or sin” (MPPDA, 1930-1967). The section titled “Reasons Supporting the Preamble of the Code” outlines the moral obligations on the part of filmmakers, particularly because of the power of film: “The reaction of a reader to a book depends largely on the keenness of the reader; the reaction to a film depends on the vividness of the presentation” (Section III D c).

In non-fiction there is a difficulty in drawing attention to sonic elements which are non-real or inauthentic in some way. Even where there may be an intention to make the audience aware of such manipulations it is often not practicable for sound, especially where the real and the non-real (faked/recreated/augmented) coexist. The BBC Producers Guidelines (2005), for example, state that reconstructions “… be labelled as reconstructions. If unlabelled they should be differentiated in some way from the visual style of the rest of the programme such as using slow motion or black and white images in a consistent and repeated way”. Yet there is no advice on sonic reconstructions, on if or how they should be signposted. Greater diligence is required in declaring manipulations in the image than manipulations to sound.

We may also look to other areas of non-fiction production for guidance on ethics. Many journalistic bodies have developed practical ethical guidelines. The American Society of Magazine Editors’ (ASME) ethical codes outline basic principles about the readers’ entitlement to fairness and accuracy, the relationship of trust, the difference between editorial and marketing messages and the protection of editorial integrity (ASME, 2013). The Consultative Group of International and Regional Organisations of Journalists (CGIROJ) noted “the important role which information and communication play in the contemporary world, both in national and international spheres, with a growing social responsibility being placed upon the mass media and journalists”, and set out ten principles of professional ethics, the first two of which was “The people’s right to true information”
and “The journalist’s dedication to objective reality” (CGIROJ, 1983). Recognising that objective reality is sometimes a slippery concept, the principle is given some context:

...with due deployment of the creative capacity of the journalist, so that the public is provided with adequate material to facilitate the formation of an accurate and comprehensive picture of the world in which the origin, nature and essence of events, processes and state of affairs are understood as objectively as possible (CGIROJ, 1983).

Brian Winston (2000: 118) describes some of the practical difficulties which beset working journalists:

Ethics in general, and ethical systems in particular, tend to be restrictive in free expression terms. They also tend to be individualistic and posit a free person facing choices as the norm. Like other workers, journalists for the most part are not free. They are responsible to their employers as well as to their readership/audience and to the participants, informants or contributors who interact with them; and these responsibilities can be at odds with each other. The owners need profit and, sometimes, a platform; the consumers information and/or titillation; the participants privacy or, sometimes, a platform.

Replacing ‘journalist’ with ‘sound practitioner’ in these two examples goes some way to describing the position of those working in non-fiction sound. In what is primarily a field which concerns reality there are pressures to present the work in the form that suits the employer’s wishes, or in a way more pleasing or dramatic to the audience, or in a way that casts the participants in a particular light.

Whether examining documentary films, television sports, news reports or dramatic reconstructions of historical events, there is a potential in the manner of sound representation for manipulation, misrepresentation or misinformation. The audience is largely unaware of the techniques of sound manipulation and uninformed of its potential to misrepresent. Sound, perhaps more than any other
area of media, has the potential to hide its tracks to such an extent that it is impossible to determine how something was recorded, whether it was actually live, and whether it was authentic.

Towards an ethics of sound

We have seen that sound in non-fiction is routinely manipulated to accentuate, suppress or replace some aspects of sound in order to present a clearer, more intelligible, exciting, involving or dramatic representation to the audience. Particularly in non-fiction productions there is an ethical dimension to the soundtrack and the choices that govern its construction. Given that the manipulation to achieve the desired end is both deliberate and fiendishly difficult to detect there is then a supplementary duty of sound producers to consider the impact of their work and whether it has the potential to mislead or misrepresent.

Where guidelines exist they are rarely suited to sound in particular, or recognise the differences between sound and visual representation in practice. In the absence of any formal code, what is offered here is intended as a starting point for a potential ethical code of practice for sound producers based on four impacts:

• On the audience – is there the potential to mislead?
• On the participants – is there the potential to misrepresent any of the participants?
• On accuracy – is there the potential to over-simplify or to misrepresent?
• On its becoming an archive – is there the likelihood that the material will become a de facto historical archive which will be taken (rightly or wrongly) as accurate?

The elements of the soundtrack through which the audience understands the text can be manipulated at every stage: whether in recording through microphone placement, re-recording and pre-recording; or through editing and mixing to remove, augment or replace sounds. All sound recording is to some extent mediated, and the mediation, whether consciously designed or unintentional, suggests that there is some onus on the practitioner to act in an ethical manner. In
moving from a view of the creation of the soundtrack from passive registration to active design we must recognise its potential to mislead. There is a duty of the practitioner to work in a manner consistent with a high ethical standard of professional practice which goes beyond the production of simple entertainment, and which takes account of context and of accuracy and does not take undue advantage of the difficulty for the audience in detecting the manipulation. In non-fiction productions there is an implication that the material being presented is authentic and that the sounds which accompany the images are what they purport to be. Even where authentic sound is used it requires attention to ensure that it does not mislead or misrepresent. If, for whatever reason, authentic sound is augmented, suppressed, or replaced, the impact of the mediation should be examined to ensure that the result does not deceive the participants or audience, or compromise the material’s accuracy or its legitimacy as archival material.

References


Batcho, James, 2012, ‘Revisiting the Howard Dean Scream: Sound Exclusivity in Broadcast News.’

Media Culture Society, vol. 34 no. 8, pp. 999-1012.


British Pathé, 1917, Time to Remember - Enough of Everything (Reel 4); British Pathé.


CNN, 2004. ‘Did Press Sink Howard Dean?; Was Media Too Timid in Challenging Bush’s Pre-War Claims?’ [Television Transcript], in Reliable Sources: CNN.

Flaherty, Robert Joseph, 2010, ‘Nanook of the North’, distributed by Gryphon Entertainment, South Melbourne, Victoria,


