

The ‘Truth of Sound’: Exploring Immersive Location Sound Recording in Realist Filmmaking

DOI: <https://doi.org/10.33008/IJCMR.2020.01> | Issue 7 | October 2021

Steve Whitford

Abstract

This article focuses on the somewhat neglected (at least within scholarly circles) area of location-based sound recording, drawing much-needed critical attention to the intricacies and skills involved in location sound recording within realist filmmaking – both scripted and unscripted. Through my own practice-as-research, I aim to reimagine an ontological definition of location sound recording by proposing that a reinvigoration of the ‘realist’ genre can be achieved by connecting the storytelling skills of recording for single camera with the new opportunities afforded by immersive audio technologies – ambisonics here being a vital part of that development process. I demonstrate how use of such immersive audio technologies offer new creative opportunities for realist makers and audiences, based on the unique experience of geographical place and physical event that immersive audio delivers.

‘The potential of the ambisonics mic is limitless and we’re only just starting to see what content producers can really achieve with it now.’
– Rode, Australia, in an interview with the author, July 2019.

The art of location-based sound recording has been a neglected area of academic research. I seek to address this by drawing critical attention to the intricacies and skills involved in location sound recording within ‘realist’ filmmaking – both scripted and unscripted. I show how this art continues to be central to the creative process of production, in driving the narrative and shaping the text’s influence, within the profilmic space. I hypothesise that the realist sound recordist’s role has an authorial voice and a creative agency. I seek to reimagine to develop an ontological re-definition of location sound recording by proposing that a reinvigoration of the realist genre – unscripted in particular – can be achieved by connecting the story-telling skills in recording for single camera, with the new opportunities afforded by the emerging technologies of immersive field sound recording – ambisonics being a vital part of that development. I argue that deploying an ambisonic-centred location sound recording methodology, fused with the existing art of recording actuality sound, will offer new creative opportunities for realist makers and audiences, now presented with an exciting ability to experience a sense of the geographical place and physical event that immersive audio delivers.

Scholarly study of sound in film has so far focused primarily on music and post-production sound-design in fiction narrative cinema (Weis and Belton, 1985; Altman, 1992; Beck 2008; Sonnenschein, 2001). The function of sound in documentaries has been a relatively under-researched area in academia as well as being largely overlooked by film critics and often lacking the recognition it deserves within the industry. As veteran documentary-maker Roger Graef observed in an interview with the author recently, ‘Ah, sound – the Cinderella part of documentary filmmaking’ (2020).

Indeed, it is often the Director that is credited as the sole author of a film, and if critical discussion recognises the role of crew at all, it is usually around cinematography, but rarely sound. Outside of the realms of music and post-production, the role of the sound recordist in

realist film production, and its part in shaping the authorial voice and creative agency of the film text, has rarely been studied [1]. In part this is a result of the historical dominance of auteur theory, which tended to ascribe authorship to the individual vision of the film's Director.

Yet even when the collective authorship of the filmic text is recognised by scholars, the role of the sound recordist remains ambiguous. For example, in arguing for a collective approach to authorship, Paul Sellors has commented:

Is the sound recordist a member of a film's collective authorship? This is not so simple to determine. Some sound recordists will count as authors under a notion of collective filmic authorship while others will not. It will depend on the recordist's contribution to the filmic utterance... we need to understand this person's role in producing not just the material film, but also its utterance' (2007: 269).

Illustrating this authorial ambiguity, Sellors further observes that 'Auteurists have tried to explain a film's coherency by overvaluing the authorial control and artistic aptitude of an individual' (Sellors, 2007: 268). Gaut, as quoted by Sellors, argues that the authorial should in fact be 'multiply classified: by actors, cameramen, editors, composers, and so on' (ibid, 267). As Sellors summarises this perspective: 'Gaut, instead, looks at the function of a collective to get from individual contributions to a completed text' (ibid, 268).

This article ventures a position that seeks to draw critical attention to the intricacies and skills involved in the art of the location sound recordist and to render visible its 'utterance', to borrow Sellors' term; to show how this art continues to be central to the creative process of production in driving the narrative and shaping the text's influence. I focus specifically on the role of sound recording within the inter-related sub-genres of realist filmmaking: social realism (scripted) and observational documentary (ObsDoc), where, as I will seek to show, sound carries a significant indexical value to the film text's assertion about its relationship to reality. Although there are fundamental differences between Social Realist fiction films (scripted, using actors) and Observational Documentaries (unscripted, using social 'actors'), the two genres and their often-hybridised forms, share a similar approach to the depiction of the pro-filmic event – scripted or unscripted.

In both cases, the aspiration is to use filmic devices to create for the viewer a sense of 'being there' [2] to minimise, if not illuminate, the inherent mediation of a reality created by the camera and sound recordist. Kuhn and Westwell et al define the profilmic space as 'The space created within the film frame as opposed to the space of the real world' or the world the lens sees. The 'authenticity' of the sound that is recorded in the pro-filmic event, or what I would term 'The Truth of Sound', is an essential component to achieve the aspiration in creating the sense of 'being there'.

Fred Wiseman, one of America's most prominent directors of documentary who, along with contemporaries, the Maysles brothers, Don Pennebaker and Richard Leacock, helped establish the American Direct Cinema tradition of the 1960s, in an interview with David Winn for The National Academy of Television Arts and Sciences, observed that: 'Observational cinema somehow seems to suggest that you just turn the camera on and let things happen in front of you, when in fact all aspects of movies are the result of thousands of choices' (emmyonline.org, 2014). It is these choices that define the observational documentary genre but Wiseman's comments also highlight the tension imposed by the

ambition of ObsDoc filmmakers to minimise the mediation of reality and so to aspire to present to the viewers a sense of ‘being there’ with Wiseman’s own comment that ‘the notion that cinema is the truth, or that anything is the truth is preposterous... Everything is subjective, and everything represents a choice.’ (ibid. 2014).

Those thousands of unscripted choices within the ObsDoc production process pose questions around the authorial voice, too, not only because of the inherent tension Wiseman identifies between a film grammar that seemingly presents to the viewer ‘reality’ unfolding ‘as it is’ and the constructive nature of filmmaking, but also around the particular production context of the genre. Observational documentary does not just exist in spectatorship; crucially, it also exists in the actual: in the physical event-space. This involves literally sharing ‘slices of life’ with protagonists and being part of unscripted events, thus requiring a ‘reactive’ approach relying in many ways on the relationships forged within the making process: inter-protagonist; inter-spatial and inter-makers.

I consider the inter-makers choreography, beyond the pro-filmic space, to a specific ‘Action-space’, demonstrating how the camera operator and the sound recordist perform collaborative yet individualised and autonomous roles. The makers’ choreography requires an equivalence of cross-craft empathy in facilitating the creation of each’s own independent narrative, opting to privilege accordingly – both contributing to the ‘audio-visual scenography’ which Chion defines as: ‘Everything in the conjunction of sounds and images that concerns the constructing of a fantasmatic diegetic ‘scenic space’ (2009: 469), with meaning deriving from ‘live’ juxtaposition – or some of the ‘thousands of choices’ that Wiseman identified.

The sound recordist’s agency centres around choices made in the event along with specific pre-emptive selections of audio equipment; chosen and deployed explicitly to gather ‘audio signs’ so as to contribute to ‘meaning’ and questions around the film’s text and its reception. Paul Sellors et al identify this authorial contribution as ‘utterance’ (ibid, 268), which he defines further as being the ‘collective authorship through theories of collective intentions’ (ibid, 268). Perhaps in a Venn diagram of ‘thinking’ (analysis) and ‘hands on’ (technical) elements, the ObsDoc location sound recordist’s utterance situates in that overlap.

That utterance clearly affirms the importance of sound’s indexical relation to the authenticity of the realist text, whether scripted or unscripted. Realist director Ken Loach, for example, observed that ‘the sound is true when it reflects the real experience of being in a location. ... if the sound is not true, then the whole authenticity of the film is undermined’ (Author interview, 2020). Similarly, Loach-collaborator (editor) Jonathan Morrison reflected that ‘the authentic sound that we get from [the recordist] is all important ... The Realism of the Sound. ... what we make is social realism, so the sound has to be real’ (Author interview, 2020). Loach-collaborator (sound recordist) Ray Beckett also refers to his approach to recording sound as ‘direct sound’, as ‘capturing the moment in front of the camera’ (Author interview, 2020), which is identical whether he works on documentary or fiction social realist film. It is an approach which is akin to what Robinson described as a commitment to ‘jishizhuyi’ [translated as ‘record-ism’], a kind of ‘on-the-spot realism’ [translated as ‘document-ism’] (Luku, 2002: 1-30). Robinson elaborates: ‘In the context of documentary practice, this entails the realisation of a spontaneous and unscripted quality that is a fundamental and defining characteristic distinguishing [jishizhuyi]’ (ibid., 1).

Technological filmmaking advancements have been an historic enabler of content innovation, transforming how makers have utilised, developed and deployed those innovations to explore

new opportunities in developing genre-specific film languages. For example, the change from 35mm film cameras to 16mm film cameras; separate Sound (Nagra); Timecode; zoom lenses; radio mics and so on. As Leacock observed after shooting documentary on 35mm film cameras:

This experience gave me a goal with clearly defined standards. I needed a camera that I could hand hold, that would run on battery power; that was silent, you can't film a symphony orchestra rehearsing with a noisy camera; a recorder as portable as the camera, battery powered, with no cable connecting it to the camera, that would give us quality sound; synchronous, not just with one camera but with all cameras. What we call in physics, a general solution (Leacock, 1993).

Indeed, as Barnouw in Robinson recognises of Fred Wiseman: 'This tradition emerged in the wake of specific technological developments – most obviously the disaggregation of camera, microphones and tape recorder, enabling synchronised sound shoots for the first time' (Robinson, 1993: 11).

It sometimes goes unnoticed that as well as editing and de facto directing, Fred Wiseman was, and also still is, the sound recordist on his films. This technological enabling process continues to evolve storytelling possibilities and choices, and to open new markets, requiring a continued evolution of those defined 'soft skills' underpinning the Recordist's utterance. The investigation of the role and creative agency of the sound recordist becomes even more complex yet relevant in the currently transforming landscape of film production, with the recent emergence of consumer accessible VR and 360-degree immersive technologies and their vibrant, cross-platform experimentation. The aspiration for immersion, interactivity and viscerality, in other words creating a sense of 'being there', is central to these new technologies, and they afford the potential to enhance this experience for the viewers in ways that previous technologies could not. Many filmmakers are experimenting with these new technologies within the documentary form, which stem from a similar aspiration to that of the ObsDocs genre – to put the viewer 'within' the film space, or, to create a sense that '...there's no separation between the audience watching the film and the events in the film' (Wiseman in Atkins., 1976: 43). But these new experiments, as before, focus predominantly on the visual, relying largely on 360-degree cameras and XR visual designs to create a sense of visceral immersion.

It is therefore key to explore the potential contribution and effect of an immersive location sound recording methodology on the prevailing classic single camera, '2D' filming methodology, and to understand how this might contribute to the reinvigoration or reimagining of the realist/ObsDoc genre in an age of immersive media. The hypothesis that guides this is that this positioning would widen the understanding of how visceral immersion can be achieved, specifically suggesting that ambisonic audio would contribute to this.

So, what is ambisonics? Robjohns in 'Sound On Sound' explains that: 'Ambisonics was conceived in the late 1960s as a complete recording and reproduction system capable of recreating accurate three-dimensional sound stages...' (Robjohns, 2001). Ambisonic records and reproduces 360-degree immersive sound from a single microphone source, giving four channels of audio recorded in the field. Software in post-production converts these channels so it is possible to recreate the effect of any conventional microphone polar pattern, pointed in any direction within the 360-degree audio soundscape. Furthermore, as ambisonics is 'speaker agnostic', a mix can then be transcoded to any transmission/consumption format,

from mono to full 360-degree immersive stereo, with height information (see Binaural later). Although flawed, an analogy for ambisonics is of an ‘audio lens’ which can be zoomed, focused, panned and tilted to fine-tune the overall sound pick-up, post event, meaning that some audio ‘focus’ decisions can be made later – software can then steer a ‘virtual hyper-cardioid mic’ towards a sound source, offering more options in post. Google has recently adopted ambisonics as the audio format of choice for VR (virtual reality) and audio companies are now marketing ambisonic-capable location microphones and recorders.

In terms of defining a field recording methodology, the ambisonic microphone movement around ‘action’ is not the classic reactive mono shotgun ‘point at action/speech’ mode: it can be moved to allow action to take place around it meaning that some audio ‘focus’ decisions can be made later where software can steer a ‘virtual shotgun mic’ towards a sound source. Crucially, in an interview with the author, in 2019, long-standing practitioner of location ambisonic recordings for international theatre sound design, John Leonard, advises: ‘...if you’re a distance away from the person talking, you can zoom-in [in post] ... but it’s like having a hyper cardioid pattern that’s too far away...’. Microphone placement and choreography within the pro filmic event space are therefore, still fundamental: new technology, with established skills. In the same way that other technical innovations have effected developing languages adding choice, so too have ambisonics.

This 360-degree audio ‘action’ recording provides an improved sense of space and place, bringing the location sound to bear – perfect for the visceral and authentic aspiration to put the ObsDocs viewer ‘there’. So how might the prevailing ‘2D’ shooting methodologies change for ambisonic-centred location sound recording, foregrounding ‘extreme naturalness’ or ‘being there’, within the two main filmmaking scenarios?

The first is separate sound: individual camera and sound operators, with a classic single camera narrative methodology, and typified by a ‘multi mono’ approach: i.e., radio mics, shotgun mic, placement mics – all augmented by a series of location-specific ambisonic atmosphere/place recordings. An ambisonic approach could utilise the ambisonic microphone as the main ‘action’ microphone, augmented with mono sources, such as radio mics.

The second is sound on camera: a single operator methodology. This approach utilises an on-camera mono microphone which effectively ‘looks’ wherever the lens is pointing. So, to pick-up ‘on mic’ sound, the camera has to point at the source otherwise it is ‘off mic’, or to use radio microphones but with a resulting increase in complexity for the single person operator. With ambisonic recording, the camera is liberated from needing to ‘aim’ at the sound source, and can now concentrate on shooting for the lens, thereby facilitating a more fluid camera response, now no longer dependent on inherent restrictions within the ‘single operator’ methodology.

With both separate sound and sound on camera methodologies, there are profound aesthetic and practical questions that arise, all of which impact on opportunities to examine and enhance the development of the form. Although the location audio can be embellished at the post-production stage, what remains crucial is the bridge between viewer and event space: being able to experience through one of the senses, an un-mediated ‘reality’. As Chesler summarised of Wiseman’s field sound recording strategies:

Ambient sound, typically picked up through an omnidirectional microphone, captures the whole of a sonic environment without privileging a specific sound source in a

scene. These ambiances defy logics of listening practice as all sounds within a space are captured within a 360-degree area.

Leonard comments on his methodological approach to recording in this format and makes a crucial observation: ‘Ambisonics gives me surround which is what I want, but it doesn't give me surround in such a way that it's distracting, which is also what I want... what it does have is extreme naturalness.’ Loach, too, observes that ‘...if it's about truth, and truth in the sense of authenticity, then you have to observe the natural rules of sound – of the experience of being there in terms of the sound... (Author interview, 2020).

So, then, how might realist/scripted and realist/ObsDoc filming methodologies change for ambisonic-centred location sound recording, in ways that foreground the truth of ‘extreme naturalness’ and the authenticity of ‘being there’? Under such a mode of recording, makers, protagonists and viewers are all placed in a common sound space – but might this be too visceral a viewing experience for some? In the *profilmic* space, the audio's equivalent of the camera's ‘lens’ is the microphone pick-up pattern. 360-degree location audio can facilitate audience engagement in a newly-defined profilmic space paradigm – not just what is visually in front of, but crucially, now *around* the lens. Might a shifting of received priorities require a commensurate academic re-definition of the term ‘profilmic’? Or indeed, a new additional classification – for example, the ‘extra-profilmic event-space’ now describing the new recorded 360-degrees situation-specific world?

For instance, for what Chion categorises as the audio-viewer, location 360-audio brings choice to aurally focus on sound elements happening outside of the ‘profilmic event’, and then to be able to select and interpret from their own ‘point of audition’ (ibid. 2009: 485), that being the spatial position from where we hear a sound. How, then, do the storytellers deal with an audience's ability to process audio information (sub)consciously from out of vision and from the world which Schaeffer defines as the ‘acousmatic’ (1966: 91): for example, choosing their own points of audition according to distance; clarity; dynamic; trajectory; movement; power, and so on. As such, one might argue that realist/ObsDoc storytelling space has now become authentically immersive, effectively contributing to the audio-visual scenography of what Chion identifies as an ‘in-the-wings effect’, with sound being located in “‘absolute offscreen” space ... to create the impression that the screen has a contiguous space’ (ibid. 478).

Camera ‘coverage’ is therefore profoundly impacted. Would the immersive location audio principle within the ‘single person, single camera’ acquisition set-up benefit from a re-discovering of the ‘fixed, prime lens’ aesthetic, meaning that camera movement itself does the ‘zooming’ and not the lens? As Loach observes, zoom lenses in realist operations distort the vision-to-audio ‘perspective of the sound, so the wide shot doesn't have closeup sound on it. ... In general, it just devalues the truth of the sound because if you're a long way away from someone, you don't hear what they are saying’ (Author interview, 2020).

Perhaps, then, a new visual methodology is required, one that is analogous to the audio's ‘natural’ 360-degree coverage, and where the audio ‘frame’ now matches the visual frame. Such a methodology would serve to promote a ‘naturalising’ perspective, and thereby contribute to the visceral experience of the immersive audio-visual content within the profilmic event space. Chion identifies the audio-visual scenography as being further broadened still ‘through the use of entrances to, and exits from, the auditory field’ (ibid. 2009: 469).

But how might the immersive location audio principle then affect film narrative language? For example, a person enters through a door but is not in shot. In a 2-dimensional audio world, the door sound would intrude as it would appear unexplained ‘on top’ of the diegetic audio, but in a 3-dimensional immersive audio world, the audio-viewer (sub)consciously ‘auditions’ the sound of the door and rationalises accordingly, placing the sound within a natural, experiential ‘world mix’. The camera, now no longer needing to explain this out-of-vision sound with a cut-away of the door or a reframe, is liberated and can act as a purely pictorial storyteller.

Relating to this point, Weis and Belton observe that ‘what the soundtrack seeks to duplicate is the sound of the image, not that of the world’ (Johnson, 1985: 4). They describe a post-production response normalising ‘natural’ as a ‘construct’, based around a typical ‘scripted’ filmmaking methodology. But, again, in realist filmmaking, questions immediately arise around authenticity: ‘The idea of an actor recreating a performance in a studio, a dead studio, is completely against capturing the truth of the moment ... There’s heightened sound effects ... they’re there to get an effect. But you get the effect at the expense of truth’ (Loach, author interview, 2020). In effect, ambisonics already centralise ‘extreme naturalness’, and so perhaps the above Weis and Belton quote should be re-worked into something that presents an academic re-framing which moves away from sound as construct towards sound as truth: ‘What the soundtrack seeks to duplicate is the sound of the world, *not* that of the image’.

Both of these filmmaking scenarios open up a set of questions around agency, authorship, and performance. Who is now performing the filming – the camera? The sound recordist? The director? Or even a new role? Does a ‘fusion’ ethic better fit an emerging model around new audiences, new platforms, and new methods of consumption? Does this more liberated methodology now require a new response in terms of skillsets? Should the terms ‘camera person’, ‘sound recordist’ and ‘director’ now to be merged and re-titled, perhaps as ‘content acquisition artist’, or ‘maker’, or something else?

Equally, what is the effect here on the agency of the sound recordist, who is now consciously assessing, augmenting and recording a 360-degree environment, and thus telling their ‘story’ in a new, developing language? With the coincidental liberation of the camera as described, would a resulting shift in hierarchical-based assumptions of authorship/agency then provide a definitive response to Sellors’ earlier cited question: ‘Is the sound recordist a member of a film’s collective authorship’?

In any case, the ‘new role’ – be it content acquisition artist, maker or otherwise – foregrounds sound storytelling skills but now with a required empathy with the 360-sound world of the extra-profilmic event-space. This would include an added understanding of what is and is not achievable on location and in post-production, an ability to understand how ‘post-mic-steering’ will work, as well as having new visual storytelling skills – now filming for audio, perhaps? Conceptually speaking, should this approach be best described and understood as ‘sound on camera’ or ‘camera on sound’, or simply as ‘audio visual’?

Although Michel Chion was writing about post-constructed soundscapes, location-recorded ambisonics similarly furthers the audio-viewer’s ‘choice’ principal and adds to the visceral nature of the audio-visual scenography that he describes. This article has aimed to consider the potential creative role of ambisonic-centred location sound recording within in the realist filmmaking genre, and reflect on how this emerging sound technology might work to centralise multiagency and multi-authorial arts, while still aspiring to immerse the audio-

viewer in a position closer to the reality that is being observed. If such a position were to be achieved, then there would be ‘no separation between the audience watching the film and the events in the film’ (Chion, 1976: 43). This would altogether continue the evolution of an audience’s potential ability to consume realist film, but now in a multi-platform, multi-screen, immersive world. Crucially, it facilitates the rediscovery of realist/ObsDoc single ‘sound/camera’ storytelling skills. As I have argued, ambisonics can contribute towards the reinvigoration of a whole new realist/ObsDocs format for makers, now no longer tied to the framing of -hour specials on terrestrial television with their speculative and high cost-bases, but instead now a reimagined version of, maybe, Instagram-length micro-docs. Might such social media-based micro-realist docs be consumed by audio-viewers on their mobile devices, while, let’s say, travelling on the proverbial (and actual) Clapham Omnibus, now fully immersed within a 360-degree audio film space, and now truly experiencing ‘no separation between the audience watching the film and the events in the film’ (ibid. 1976: 43)?

References

- Altman, R. (Ed.). (1992). *Sound theory, sound practice*. Routledge, New York
- Atkins, T. (Ed.). (1976). *Frederick Wiseman*. New York: Monarch.
- Beck, J. (Ed.). (2008). *Lowering the boom: critical studies in film sound*. University of Illinois Press.
- Beckett, R. (2000). Online interview the author. 17 Aug 2020.
- Chesler, G. (2012). ‘Truth in the mix: Frederick Wiseman’s construction of the observational microphone’. In: *Frederick Wiseman, Kino des Sozialen* (Ed. Eva Hohenberger, Vorwerk 8: 139-155.
- Chion, M. (2009). *Film and Sound Art*. (Gorbman, C. Trans.). Columbia University Press.
- Gaut, B. (1997). *Film Authorship and Collaboration*. In R. Allen, & M. Smith (Eds.), *Film Theory and Philosophy*. (pp. 149-172). Clarendon Press.
- Graef, R. (2020). Online interview with author: 22 Sep, 2020.
- Kuhn, A & Westwell, G. (2012). *A Dictionary of Film Studies*. Oxford University Press.
- Leacock; Richard. (1993). *A Revolution in Documentary Film Making as seen by a Participant*.
- Loach, K. (2000). Online interview the author. 19 Aug 2020.
- Morrison, J. (2000). Online interview the author. 31 Aug 2020.
- Robinson, L. (2007). *Contingency and Event in China’s New Documentary Film Movement*. Nottingham EPrints.
- Robjohns; H. (2001). *Surround Sound Explained: Part 3*. Ambisonics. SOS Publications Group.
- Schaeffer, P. (1966). *Traité des objets musicaux*. SEUIL.
- Sellors, P.C. (2007). ‘Collective Authorship in Film’, *The Journal of Aesthetics and Art Criticism* 65(3). Available at: www.jstor.org/stable/4622239.
- Sonnenschein, D. (2001). *Sound design: the expressive power of music, voice, and sound effects in cinema*. Michael Wiese Productions.
- Weis, E & Belton, J. (Eds.). (1985). *Film Sound: Theory and Practice*. Columbia University Press.
- Winn, D. (2014). 31st Annual News And Documentary Emmy Awards Lifetime Achievement Honoree: Frederick Wiseman. The National Academy Of Television Arts And Sciences. Available at: http://emmyonline.org/news/news_31st_interview_wiseman.html

Notes

[1] An exception is Chesler's illuminating analysis of the work of sound in Fred Wiseman's observational documentaries. See Chesler, G. (2012). 'Truth in the mix: Frederick Wiseman's construction of the observational microphone'. In: *Frederick Wiseman, Kino des Sozialen* (Ed. Eva Hohenberger, Vorwerk 8: 139-155.

[2] A defining term used by the veteran documentary filmmaker Richard Leacock. See 'A Search for the Feeling of Being There', *Memoirs of Richard Leacock* (1997). Available at: <https://mf.media.mit.edu/courses/2006/mas845/readings/files/RLFeeling%20of%20Being%20There.pdf>