INTERNATIONAL JOURNAL OF CREATIVE MEDIA RESEARCH

Vol.11, April 2025

Song and Soundscape – Immersion and Presence in Spatial Audio

DOI: [10pt black]

Philip Reeder, *University of Gloucestershire*

& Danielle Meunier, *Canadian College of Performing Arts*

Abstract

Returns is a creative practice research project, involving multichannel recordings of traditional folk songs performed in rural locations. This writing discusses how aspects of immersion and presence were considered over the making process alongside the dissemination method of head tracked headphone-based listening via Apple's spatial audio. Perceptual immersion, narrative immersion, and certain dimensions of presence are aligned with praxis, resulting in findings from across production processes and the playback system. We propose that these individual parts form a synergy to enhance immersive potential.

Soundcloud link: https://soundcloud.com/returns-678875839/sets/returns?si=28ffa7a2de584478abdfdb86c6e5f25b&utm_source=clipboard&utm_medium=text&utm_campaign=social_sharing%20

Introduction

The audio EP *Returns* is the outcome of producer Philip Reeder recording traditional folk songs, sung by Danielle Meunier at selected exterior landmarks in Cornwall, England (fig.1). The three pieces show an evolving approach to producing for headphone listening via what Apple terms spatial audio - an implementation of Dolby Atmos for head tracked headphone, binaural, and Dolby Atmos compatible systems.

As we have recorded performances of folk songs amongst physical soundscapes, *Returns* relates to various established practices such as soundscape production, phonography, and folk song recording. It could be seen as building on the authors previous work *Still Every Year They Went*, where Reeder recorded singer songwriter Johny Lamb aka 30lbs of Bone, singing traditional shanties on a fishing boat whilst at sea. In that project Reeder integrated

soundscape recordings with the performances by Lamb, heavily editing both the vocal, instruments, and surrounding material, as well as instrumentalising soundscape material to form pitch-based structures and rhythms that formed a significant part of the musical arrangement. That approach therefore involved the curation and microphone techniques of phonography, and embraced contemporary trends in electronic music production sound alongside aspects of reduced listening often rooted in acousmatic practice.

Returns is focussed on the question of how content and playback systems co-act to impact immersion and presence. In particular, it considers how the character of immersion and presence might depend on the performance, recording, editing, and listening processes involved. We will therefore begin with an overview of the pieces and locations, examine the practical workflow, investigate immersion and presence sit in relation to production decisions, and discuss aspects from the performer's perspective.

Song and Locations

The project emerged out of a series of brief experiments the artists conducted where a song was sung softly, and recorded in close proximity to the microphone. One example, when stranded in a rainstorm, took place in a vehicle with the aggressive weather conditions to the exterior. This and similar listening situations were the foundation for the series of recordings that eventually formed *Returns*, and our focus on immersion and presence. A number of songs were trialled. The chosen songs were amenable to rubato with significant pauses. They could also withstand the tensions of performances across multiple locations and sometimes major digital editing. The choice of recording sites was guided by locations and soundscapes the producers felt complemented the song and the performance style, with careful thought as to how the microphone could capture the unique contexts. For instance, the recording of Meunier singing 'What a Voice' (trad Norfolk folk song) is a single take, with barely perceptible biophony or geophony (Krause 1987), and yet reveals a subtle resonance. This natural reverb, inherent to the stoney acoustics of the historic, dilapidated, roofless engine house on Botallack cliffs, is the main complement to the voice. 'Song to the Seals' (1930) comprises two takes in a single location, where the voice combines with insect life, wind, and occasional passing cars on hills above Zennor. The final piece, 'A Stór Mo Chroí' (*Treasure of my Heart*), weaves together the remaining locations (see Figure 1) with their matching sung excerpts. In addition to exploring aspects of immersion and presence, the EP arcs from documentation of a performance through to

soundscape song collage reflecting the contemporary fiction of popular music production, thus building on Reeder's previous projects (Lamb and Reeder 2016).

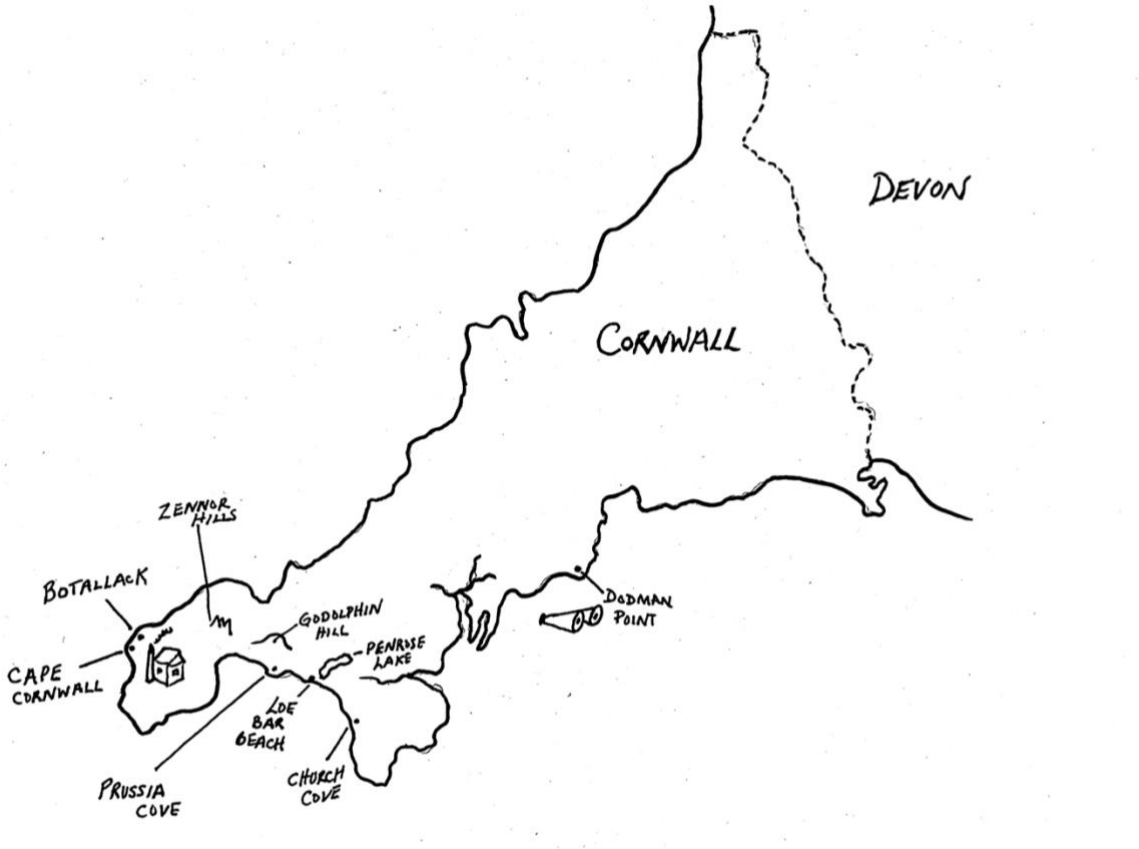

Figure 1: Recording map of *Returns* in Cornwall, UK, drawn by Danielle Meunier.

Workflow

Recording in exterior rural environments tends to create particular constraints for projects, that feed through into future production phases. Locations are sometimes remote, so equipment must be compact, easy to carry long distances, and suitable for the terrain and weather encountered. To record for multichannel formats beyond stereo, we employed an adaptable microphone setup with two points of capture. The first was a headset worn by the performer, equipped with a lavalier microphone (DPA 4066). This approach permitted few audible indications of the wider soundscape, and granted a close recording of the voice. While it is not used on 'What a Voice', it is employed for a few seconds on 'Song to the Seals', and is used exclusively on 'A Stór Mo Chroí' vocals. Although this proximate

microphone denies sonically remote context thus reducing acoustic bleed, intimate details are often exposed. For instance, when heard in isolation, it is possible to hear changes in timbre which were the result of acoustic reflections caused by waterproof clothing.

The second point of capture was a double mid-side (DMS or MSM) microphone array (Sennheiser MKH800 Twin and MKH800), spaced within a metre of Meunier on 'What a Voice', several metres for 'Song to the Seals', and up to 10 metres away for 'A Stór Mo Chroí' (fig. 2). This microphone technique provided the site context excluded by the headset, and was decoded to output a horizontal-plane surround sound. Productions were assembled in Logic X, utilising its spatial audio functionality.

For the performer, one further workflow was necessary. Songs were recorded across different days and locations, and to keep the performer in key we used a drone pitch. This was played back in an earbud worn by the performer, beyond the microphones' auditory reach. Other performance aspects were at the discretion of the performer.

Locations were often visited multiple times to capture the most engaging weather and environmental conditions, as this was an important aspect when working towards an immersive outcome.

Figure 2: Varying environmental conditions and proximities from the double mid-side microphone, at Botallack (left) and Dodman Point (right), in Cornwall. Photos by Philip Reeder.

Immersion

Drawing on review work by Agrawal and Simon et al. (2019, 6), we take immersion to be a "state of deep mental involvement in which…cognitive processes cause a shift in … attentional state such that one may experience dissociation from the awareness of the physical world". Whilst they acknowledge that measurable properties of a system play a significant role, they also bring together a number of researchers' findings into a tripartite, which feeds into a psychological state of immersion. These findings include perceptual immersion (McMahan 2003) where multisensory stimulation by the system is sufficient to prevent a transfer in attention; narrative immersion (Ryan 2003) when engrossed in the narrative; or being absorbed in a challenge (Ermi and Mäyrä 2005). As we are thinking initially about listening, we are most interested in the first two aspects - perceptual and narrative immersion.

Perceptual immersion suggests that shifts in attention can be prevented by 'blocking the external world and constraining the user's perception to the presented stimulus' (McMahan 2003, 77). This blocking has been noted specifically in headphone works for theatre by Klich (2017, 368-370) who affirms that "headphones enforce perceptual immersion" as they "exclude the exterior world". For audio on this project, the ability to exclude the exterior world is not just a system property, but also one that reflects decisions around content. The audio in 'A Stór Mo Chroí' was designed to have a high programmatic amplitude with broadband noise masking owing to the qualities of the source material (waves, wind, rain). In conjunction with the vocal heard via the lavalier microphone, the perceptual bandwidth for listening was intended to be largely occupied. In contrast, 'What a Voice' is prone to admitting exterior sounds, with a single vocal in a sonically desolate location. Although the voice is clearly portrayed, the absence of broadband noise masking could allow real-world sounds to impact the listening situation and dilute the potential for perceptual immersion.

In noting that the vocal in 'What a Voice' is prominent, we intended to touch upon two aspects of narrative immersion posited by Ryan. Given reasonable exclusion of the exterior world by the playback system as outlined above, we were interested in temporal immersion (Ryan 2003) where curiosity for what will follow, maintains attention. The other is spatial

immersion where there is an 'intimate relation to the setting' (Ryan 2003, 122) through an act of exploration, in this case through listening. As the voice is so exposed in 'What a Voice' through the microphone position and recording situation, the listener can attend to the narrative aspect of the lyrics, and the delivery by the performer. This production decision allows those possibilities for temporal immersion (i.e. what will happen to the protagonist?) to come to the fore. 'A Stór Mo Chroí' incorporates material more overtly suited to spatial immersion, involving recordings and performances from multiple sites and weather conditions. The varied soundscapes composed in the opening minutes enable exploration of contrasting and rapidly changing spatial scenes. The evolving approach to the soundscape in conjunction with the sung narrative, allows for a combination of the temporal and spatial aspects of narrative immersion.

Presenting environments recorded in surround sound does not automatically result in sonic outcomes that are superior to stereo, with stronger possibilities for immersion. Locations don't always have desirable sound emanating from every direction, and getting a microphone to locations that do exhibit such spatial potential can be prohibitively difficult. Marrying content that considers immersive attributes, with a corresponding immersive system, is thus important to create fertile conditions for immersive experiences. From a sense of absorption in the narrative and perceptual immersion, many of the multichannel recordings in *Returns* are gathered from positions that lend themselves to the use of spatial output systems. As an example, recordings were taken not just on the beach, but also whilst standing in the waves themselves (fig. 3).

Figure 3: 'Immersed' whilst recording waves in surround at Loe Bar. Photo by Danielle Meunier.

Similarly, the sound of rippling tidal water from a flooded peninsula, ensured that the delicate sounds would be captured in surround whilst Meunier sang (fig. 4).

Figure 4: Singing and recording at the end of a peninsula on Penrose Lake, Cornwall. Photos by Danielle Meunier and Philip Reeder.

Sometimes when speaking about such locations we discussed how to give the listener the feeling of 'being there'. This oriented our consideration of immersion towards presence as 'the participant's sense of "being there"' (Slater, Usoh, and Steed 1994, 2). As we acknowledged the limitation of field recordings 'to truly substitute for any given environment' (Lopez cited by Bailey, 2009, 246), the notion of presence allowed us to consider the degree to which combinations of system and content might generate a perceived absence of mediation.

Presence

A number of dimensions for presence, which have been usefully reviewed to align with interior acoustic music recordings (Kelly et al. 2020), were first compiled by Lombard and Ditton (1997). Some dimensions, like presence as immersion, overlap with concepts explored above, and some are of limited utility for this project such as quality of social interaction, the social impact of events in the environment, and the way users might "respond to the computer itself as an intelligent, social agent" (McMahan 2003, 73). Not

all factors have to be employed to result in a degree of presence (73), and when considering music played through head tracked headphone systems, we considered those aspects of presence that related to realism and interaction.

Referring back to Lombard and Ditton (1997), realism can be social (could it plausibly happen), and perceptual (even if it isn't socially plausible, is the technical fidelity convincing). In 'What a Voice' both social and perceptual realism are at play. We placed the DMS microphone within 1 metre of Meunier, aiming for a high degree of social realism so listeners can localise the voice moving when Meunier swayed whilst singing. The microphones are noted for their high-resolution and low noise floor (Robjohns 2000), so the recording captures Meunier's voice altering in tone, and the quiet reverb excited within the recorded space. What is heard on the recording is therefore also perceptually plausible. In this case, the transmission of plausible social clues fuses with convincing technical fidelity. The combination increases the potential for realism, which the use of head tracking in playback will further enhance owing to its superior localisation. Head tracking with headphones has been proven to lead to higher levels of localisation particularly in relation to the historically thorny issue of front and rear localisation (Stitt and Katz et al. 2016). Stitt and Katz et al. further note realism is substantially reduced without head tracking, because the sound source is anchored to a head-centric system as opposed to the external world. These issues are avoided by mixing for the head tracking allowed on the Apple Music store, amongst others.

In 'Song to the Seals', the DMS microphone was used as the primary source again, albeit at a greater distance and low to the ground. The mid-side microphone had been placed at an unusually low vantage point in an effort to increase the potential for spatial immersion. The recording portrayed insects loudly and with more detail than would typically be heard by a standing human, compromising social realism as a result. Furthermore, when the lyrics mention dreaming, we selected the lavalier microphone, and a shift in audio quality can be heard. Whilst the change is partially intended to extend listener engagement, it also alters the sonic characteristics of Meunier's recorded voice in an implausible manner. It creates an instantaneous increase in perceived proximity and clarity, and thus weakens the potential for social realism as this shift could not have happened in real time. Taken as a whole, the approaches on 'Song to the Seals' are examples where efforts to increase spatial immersion might have come at the expense of social realism, suggesting that the potential for inverse relationships between immersion and presence should be considered mid-production.

In 'A Stór Mo Chroí' we move further from social realism, involving multiple sung takes in obviously different locations. This impossibility was alluded to by Harrison (1999) about a series of footsteps in his piece *Unsound Objects* when he notes 'all are independently possible and aurally believable, but they are not actually possible simultaneously in the physical domain of "real life"'. We opted to coalesce our takes by selecting the lavalier microphone for the vocal, aiming to maintain consistency of perceptual realism at the cost of social realism. Whilst the vocal's acoustic is sufficiently consistent, at certain points in the track we refer to the incongruence between changing environment and vocal constancy, by using a convolution reverb that takes an exogenous impulse response (Reeder 2013), in this case an excerpt from 'What a Voice'. This imbues the vocal with additional and socially unlikely tones. Whilst this production decision might actively erode social realism, perceptual realism could be maintained owing to the technical fidelity involved in this process.

Extending from Lombard and Ditton's concepts above, experiences of presence can also consider the importance of being able to interact with the environment. The use of head tracked spatial audio over headphones does provide a degree to which listeners can 'do there' as well as 'be there'. When conducting participatory research into 'doing' as an aspect of presence, Slater and Steed (1998) noted 'When they are told to move - turn their head…they frequently have an observable "aha!" type experience indicating a transition from low to high presence' (2). Whilst our listeners are not able to interact to the same bodily degree as in virtual reality, there is a degree of exploration and interaction through spatial re-orientation of the mix available to the listener.

The Performer's Perspective

Whilst discussions thus far have focussed upon recording, editing and presenting the work, observations also arose from the performer's perspective. The concept of flow has been compared to being immersed in a challenge (Agrawal et al. 2019), albeit with caveats around flow being absolute and immersion being a spectrum. Nakamura and Csikszentmihalyi (2009) assert that the state of flow, "is one of dynamic equilibrium". Citing Berlyn (1960) and Hunt (1965), they go on to explain:

> The balance is fragile. Entering flow depends on establishing a balance of perceived action capacities and action opportunities. If challenges exceed skills, one first

> becomes vigilant and then anxious; if skills exceed challenges, one first relaxes and eventually becomes bored. (196)

There were a number of challenges to achieving 'balance'. Meunier's unobscured ear was open to the unmediated biophony and geophony, and the acoustic feedback from the surrounding structures, amongst other sensory inputs from the physical environment (e.g. visual, kinaesthetic, olfactory). These unmediated elements had to be negotiated, alongside the mediated pitch in her earbud, her awareness of the microphones, and the implications of being recorded.

Meunier's awareness of the microphones and recording process had an impact on the dynamic equilibrium (Nakamura and Csikszentmihalyi 2009). Early developmental recordings used only the DMS microphone placed at a distance. After several takes it was observed that although the performer was singing in relative tune, when takes were compared the tuning was not precise. To resolve this issue a drone tone was introduced to help centre all takes in the same tonality. A headset lavalier microphone was also added. Meunier noticed that once the headset and drone tone were introduced, she felt freed up to use a greater range of expression knowing the more subtle vocal textures would be captured by the lavalier. Because Meunier's vocal origins lie in classical training, the perception of one microphone sitting at a distance encouraged the use of fully projected vocals, and an increased anxiety that the subtler vocal colours would not be received. The end result of both microphones was less worry about vocal production, greater freedom to respond with embodied emotion and expression, and ultimately increased levels of absorption in the act of performance.

Arriving at this final microphone configuration also led to a renewed focus on the original impetus for the project – the physical impossibility of hearing an intimate, proximal vocal, situated within the often sonically boisterous geophony and biophony of an environment. Aspects of paraverbal expression, such as the audible sigh in 'What a Voice' would not be heard by a listener unless in very close proximity. In 'Song to the Seals' the singer half whispers, half sings the question 'Was it a dream, were all asleep?'. This lends an intimacy that cannot be experienced by a listener positioned at a greater distance from the singer. The final held note in 'Song to the Seals', also offers a level of proximal intimacy. To achieve this delicate straight tone, the singer implemented a delicate balance of breath pressure and vocal fold contact to achieve a sound that sits on the edge of

phonation. By sustaining the note beyond learnt expectations the listener is suspended in time awaiting the final denouement in the form of gentle vibrato to complete the phrase.

The impact of biophony and geophony was also a factor in the performer's sense of immersion. The semi-enclosed 20m^2 location for 'What a Voice' offered sonic, visual and kinaesthetic seclusion, which for Meunier facilitated physical relaxation and sensorial stillness that shaped aspects of her performance. The moor-top and coastal locations with the presence of waves, insects, high levels of wind in 'Song to the Seals' and 'A Stór Mo Chroí', are in sharp contrast to the seclusion, but both relate to what Stock (2011) refers to as direct kinaesthetic participation:

> Senses other than sight and sound are heightened, such as: smell and touch (deliberate or accidental); bodily sensations like wind, rain, heat on the body; changes in texture of surfaces; and importantly, direct kinaesthetic participation through a physical exploration and experience of site. (4)

These sensory inputs shape the performance and align with narrative immersion with respect to exploration of space. Meunier was absorbed in the immediate environment and the challenge of her performance, and dissociated from the larger environment outside her occupied vicinity.

Conclusion

The *Returns EP* is an interdisciplinary project that combines vocal performance and soundscape recording with various production decisions directed towards producing music for a head tracked headphone system. This article examines how content and playback systems might come together to impact immersion and presence, adopting a maker's perspective throughout.

We noted that perceptual immersion is focussed on how a system can limit access to the external world, and this extended to inform content variables, such as the frequency occupancy and amplitude in 'A Stór Mo Chroí'. Conversely, lower levels of perceptual immersion might positively impact narrative immersion in sparsely populated audio content, as in 'What a Voice'.

The search for locations appropriate to fuller spatial representation, led to an investigation into how dimensions of presence such as perceptual and social realism complemented one another at the recording stage. This was augmented by acknowledging that a head tracked

headphone system that improves the sound localisation of Meunier's movement can further enhance the quality of the perceptual and social realism.

Our discussion also touched on possible trade-offs in 'Song to the Seals', regarding how perceptual and social realism are not always positively correlated, and taken further, how immersion and presence might sometimes have an inverse relationship. We have examined performance strategies that aimed to keep the performer in a state of flow by addressing various technical challenges, and noted how the performer's immersion in the sites and performances allowed her to act as an interpretive conduit of the environmental elements.

When considering the project as a whole the notion of positive synergy seems apt, where one decision might magnify the effect of related decisions. With particular readings of immersion and presence as a foundation, our creative practice research has uncovered the importance of an underlying intent to work towards an immersive experience, maintaining a sensitivity to likely synergies from across the production process and playback systems.

References

Agrawal, Sarvesh & Simon, Adèle & Bech, Søren & Bærentsen, Klaus & Forchhammer, Søren. 2019. "Defining Immersion: Literature Review and Implications for Research on Immersive Audiovisual Experiences." *Conference: Audio Engineering Society 147*. https://www.researchgate.net/publication/336414398_Defining_Immersion_Literature_Review_and_Implications_for_Research_on_Immersive_Audiovisual_Experiences

Bailey, Thomas Bey William. 2012. *Micro Bionic Radical Electronic Music & Sound Art in the 21st Century*. 2nd ed. Belsona Books.

Bantock, G. 1930. 'Song to the Seals'

Dibben, N. 2013. "The Intimate Singing Voice: Auditory Spatial Perception and Emotion in PopRecordings" in *Electrified Voices: medial, socio-historical and cultural aspects of voice transfer*. Dimitri Zakharine & Nils Meise eds. 107-122. V&R Unipress

Ermi, L. and Mäyrä , F. 2005. 'Fundamental Components of the Gameplay Experience: Analysing Immersion,' in *Worlds in Play: Selected papers of the 2005 Digital Games Association Second International Conference*, edited by Suzanne de Castell and Jennifer Jenson: 37-53. https://homepages.tuni.fi/frans.mayra/gameplay_experience.pdf

Harrison, Jonty. 1999. “Imaginary Space.” Spaces In The Imagination Australasian Computer Music Conference - Keynote Address. eContact 3, no 2. https://econtact.ca/3_2/ACMConference.htm

Kelly, Jack; Woszczyk, Wieslaw; King, Richard. 2020. “Are you there? : A Literature Review of Presence for Immersive Music Reproduction.” AES Convention 149 (October): 10389. https://www.aes.org/e-lib/browse.cfm?elib=20926

Klich, Rosemary. 2017. “Amplifying Sensory Spaces: The In- and Out-Puts of Headphone Theatre.” *Contemporary Theatre Review* 27, no. 3 (July): 366-378. DOI:10.1080/10486801.2017.1343247

Krause B . 1987. “Bioacoustics, habitat ambience in ecological balance,” *Whole Earth Review* 57: 14–18

Linklater, K. 2009 “The Alchemy of Breathing” in *Breath in Action: The Art of Breath in Vocal and Holistic Practice,* edited by Jane Boston and Rena Cook, 101-107. London: Jessica Kingsley

Lombard, Matthew and Ditton, Theresa. 1997. “At the Heart of it All: The Concept of Presence.” *Journal of Computer-Mediated Communication* 3, no. 2 (September) https://doi.org/10.1111/j.1083-6101.1997.tb00072.x.

McMahan, Alison. 2003. “Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Game” in *Video Game Theory Reader*, edited by Mark J.P. Wolf and Bernard Perron, 67-86. New York: Routledge, Taylor & Francis.

Nakamura, J. and Csikszentmihalyi, M. 2009. “Flow Theory and Research” in *The Oxford Handbook of Positive Psychology*, edited by Lopez, S.J, and Snyder, C. R., 195-206. Oxford: Oxford University Press

Normandeau, R. and Ledoux, D. 2018. “An Immersive Approach to 3D-Spatialized Music Composition: Tools and Pilot Survey.” *Audio Mostly ‘18 Conference: Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*. ACM Digital Library. https://dl.acm.org/doi/10.1145/3243274.3243300

O’Higgins, B. 1882-1949. ‘A Stór Mo Chroí’

Ong, W. 1982. *Orality and Literacy.* London and New York: Routledge

Reeder, Philip and Lamb, Johny. 2016. “Still Every Year They Went: Studio Performance at Sea.” *Performance Research* 21, no. 2 (May): 79-80. https://doi.org/10.1080/13528165.2016.1162513

Reeder, Philip. 2013. “Inter-Piece Sampling and Convolution: Portfolio of 5.1 Acousmatic and Electronica Compositions, Interactive Diagrams and Text.” PhD Thesis. University of the Arts London and Falmouth University. https://ualresearchonline.arts.ac.uk/id/eprint/8759/

Robjohns, Hugh. 2000. “Sennheiser MKH800: Extended-bandwidth Studio Microphone.” Sound on Sound. Accessed April 26, 2023. https://www.soundonsound.com/reviews/sennheiser-mkh800

Ryan M. L. 2003. *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media.* New York: The Johns Hopkins University Press.

Slater, M., Usoh, M., and Steed, A. 1994. “Depth of Presence in Virtual Environments.” Presence: Teleoperators & Virtual Environments, 3(2): 130–144

Stock, C. 2011. “Creating new narratives through shared time and space: Per-former/audience connection in multi-site dance events” in *Time Together: Viewing and Reviewing Contempo-rary Dance Practice: Refereed Proceedings of the World Dance Alliance Global Summit 2010.* In Caldwell, L (Ed.). 1-14. United States: World Dance Alliance-Americas

Truax, Barry. 2008. “Soundscape Composition as Global Music: Electroacoustic music as soundscape.” *Organised Sound* 13, no. 2 (July):103-109. https://doi.org/10.1017/S1355771808000149

Webber, Stefan; Weibel, David and Mast, Fred. 2021. “How to Get There When You Are There Already? Defining Presence in Virtual Reality and the Importance of Perceived Realism.” *Frontiers in Psychology* 12 (May). https://www.frontiersin.org/articles/10.3389/fpsyg.2021.628298/full

What a Voice, no date

Wilkie, Fiona. 2002. “Mapping the Terrain: a Survey of Site-Specific Performance in Britain,” *New Theatre Quarterly*, 18, no. 2: 140-160