Listening, Corporeality, Place and Presence

1. Introduction

This chapter considers the role of sound, and more specifically, listening, in creating a sense of presence (of 'being there') in 'places' recreated by virtual reality technologies. We first briefly review the treatment of sound in place and presence research. Here we give particular attention to the role of sound in inducing a sense of presence in virtual environments which immerse their users in representations of particular places. We then consider the phenomenology of listening, the nature of different types of listening and their application: listening is active, directed, intentional hearing and is not merely ego-centric, it is body-centric. A classification of modes of listening which draws on work in film studies, virtual reality and audiology is then proposed as a means of supporting the design of place-centric virtual environments in providing an effective aural experience. Finally we apply this to a case study of listening in real and simulated soundscapes and suggest directions for further applications of this work

1.1 Sound, sense of place and presence

Studies of sense of place and allied concepts such as spirit of place , place identity and place attachment are distributed across the literature of phenomenological and social geography (the classic work of Relph, 1976, Tuan 1977 among others) environmental psychology (e.g. Canter (1997), cultural and leisure studies (e.g. Jorgenson and Stedman, 2001, Haldrup and Larsen, 2006), and philosophy (from Aristotle to Bachelard, 1994 and Casey, 1997). Throughout, the material characteristics of the physical space are treated as intrinsic to sense of place, comprising not merely the natural or manmade landscape, but sights, smells and sounds. As Tuan observes, sound serves to delineate the dimensions of a place, to create a sense of its size and the relative distance of objects within it from the observer. Further, places have their own characteristic sounds, some of which serve to identify a particular sort of place – the conversation, keyboard clicks and printer noises of a busy office, the constant thrum of a busy road, the birdsong and leaf rustle of a wooded glade – others of which are unique to that place alone. This is vividly encapsulated in the US National Public Radio (NPR) network's solicitation of 'audio postcards': "… the sound should somehow be remarkable -- the rasping of 17-year cicadas so loud it drowns out conversation; the music of church bells in the medieval German city resonating with history and spirituality and celebration; the midnight creaking and snapping of birches in the Maine woods in January eerie and otherworldly. This is sound that is not just ambience. It's the audio equivalent of that four-color photo. It should really make listeners feel they were there.". (NPR, n.d.)

However a detailed treatment of sound and place is rare in the academic social science literature[i], perhaps because much recent work has focused heavily on non-material, socio-cultural meanings of place. For this we must turn to the presence research community, and as we shall see later, to film design. There are many possible definitions of sense of presence, some emphasizing the illusion of non-mediation in virtual environments, others the quality of being with others who are not physically present, but for our purposes here we intend the sense of 'being there' in an environment (including real environments as well as virtual environments, or the

location of a movie scene, or the setting of a chapter in a novel…) even when one is physically situated in another. (Insko, 2003, Witmer and Singer, 1998).

Sound has been explored largely as a contributory factor to sense of presence: sound is better present rather than absent (Gilkey and Weisenberger, 1995; Hendrix and Barfield, 1996); usually better spatialised than not (Hendrix and Barfield, 1996; Stanney *et al.*, 1998; Murray *et al.*, 2000; Bormann, 2005); and, generally, the more realistic (or perhaps plausible) the better. Sound is also used to suggest a location or event, rather than simply reproducing it, or to evoke a particular mood, as discussed, *inter alia* in Robertson *et al.* (1998) and Sheridan (2004), and in common with other forms of mediated experience such as (video) games or movies: Kubrick's use of Ligeti's requiem in 2001: A Space Odyssey successfully transported us to the depths of the solar system while John William's theme music for Jaws created an extraordinary sense of dread respectively. In explicitly place-related[1] presence research, recent sound-focussed work includes the BENOGO project (Serafin and Serafin, 2004; Turner and Turner, 2006), whose virtual, photo-realistic recreations of botanic hothouses, cityscapes and interior environments included equally realistic audio and EMMA, where the virtual 'Relaxation Island', intended for psychotherapeutic use, was set in a soundscape of mewing seabirds and lapping waves (Freeman *et al.*, 2004).

However, the exploration of the phenomenology of listening rather than sound in itself has received less attention by the presence community. References to listening do appear regularly in almost all the various questionnaire instruments developed to assess presence (Witmer and Singer, 1998; Dinh *et al.*, 1999; Nichols *et al.*, 2000; Baños *et al.*, 2000; Lessiter *et al.*, 2001, Schubert *et al.*, 2001 and Larsson et al, 2000) but listening is generally treated as an un-nuanced activity and probe items confined largely to the clarity, realism and localization of sounds. Those research reports which provide a more considered treatment of listening are Pressing (1997) who discusses sound and musical performance in virtual environments and observed the "familiarity, context, and developmental or physiological significance clearly have a substantial impact on sound reception", Storms and Zyda (2000) who treat prior listening experience as a nuisance variable; Brunart (2002) who investigates how near-field audio displays may, for instance, allow discrimination of urgent and non-urgent warnings, sounds nearer the head being interpreted as more urgent, or to increase the perceived intimacy of an experience by presenting sounds as apparently very close to the listener; and Cheung and Marsden (2002) who discusses the role of expectation in auditory experience. A more substantive consideration, however, of listening and hearing is provided by Murray *et al.* (2004) and we return to this below.

2. Listening

In describing the auditory system, Gibson (1966) argues that the perception of sound involves listening, not just passive hearing. And it is to listening we now turn our attention. We begin with a consideration of listening with reference to our bodies before discussing the ways in which different forms of listening have been classified.

---

[1] Much VR-based presence work is ultimately about persuading people that they are somewhere they are not, we intend here research whose primary aim is to (re)create specific places, whether real or imaginary.

Sound is not merely (potential) sensory stimuli, a source of auditory information - it is information. Gibson also distinguishes between two functions of the auditory system:

- Extero-receptive – this collects information on the direction of the sound event, its orientation, the nature of the event giving rise to it;

- Proprio-receptive – this collects and processes sounds made by the individual (e.g. speaking, breathing, sniffing).

The listening system comprises the ears which collect sounds and through the movement of the heads collect information on their direction[2]. The auditory system processes information concerning the intensity, pitch, direction and duration of the sounds but this is not confined to the ears. The interpretation of auditory information is always with reference to the body. At a the simplest level, our corporeality allows us to locate sound – above, below, in front – behind and in doing so identifies the location of our bodies. While the ear may be the most obvious focus of our audition they are not the sole source of information.

Rodaway (2001:91) notes that the body has its own auditory presence, "both explicitly through the vocal chords and implicitly in the friction of its movement (internally and against the external environment) and, most importantly, its own biorhythms which allow us to measure the pattern of sounds (rhythm, pace, duration). Auditory perception is against this corporeal background and in reference to it." The auditory world is experienced as it surrounds us and as participants in it.

Auditory space is, however, quite different from visual space. Hull (1990) comments that "Sound places one with a world." Similarly Ihde (1976) observes that we are at the edge of visual space and consequently at a distance, in contrast we are also at the centre of the auditory experience. This is all neatly summarized by Carpenter (cited in Rodaway, *ibid*: 114) as follows, "Auditory space has no favoured focus. It is a sphere without fixed boundaries, space itself (soundfield) not space containing a thing. It is not pictorial space, boxed-in, but dynamic, always in flux, creating its own dimensions moment by moment. It has no fixed boundaries, it is indifferent to background. The eye focuses, pin-points, abstracts, locating each object in physical space against a background; the ear, however, favours sounds from any direction".

Listening is thus not only a matter of localising sounds (as discussed in much presence research) but also places us in space through its inherently corporeal nature. Listening, like its visual equivalent, looking, is active and intentional. A consequence of these active, involved and directed senses is that they have the attribute of "throwness". "Throwness" is a term introduced by the phenomenological philosopher Martin Heidegger (1927/1962) and used to describe the way in which moods are disclosed (become known)[3]. Dreyfus (1991: 174) helpfully notes that "Moods provide the background on the basis of which specific events can affect us". If I am in a bad-tempered mood, everything is annoying to me. Moods colour our view of the world and events and reveal how things are going with us - i.e. moods reflect things back to

---

[2] Many animals have the added luxury of being able to swivel their ears too.

[3] Winograd and Flores (1986) are usually credited with introducing a number of Heidegger's concepts including ready-to-hand, throwness and present-at-hand to the computing world and Zahorik and Jenison (1998) have briefly discussed the concept in the context of presence. It should be said that in consulting the source material (Heidegger's *Being and Time*) it is a little difficult to reconciled their interpretation with the original.

us. Things, whether we are seeing or listening to them, are encountered as attractive, appealing, boring, tiresome and a dozen other things and there is no way in which they can appear neutral. The "throwness" of listening compels us to have the world disclosed to us in a manner which reflects our moods. (The language of Heidegger may be difficult but the points he makes are invaluable.) In all, this is why listening is so compelling, and its interpretative function so central. As Menninger (1938), quoted in Walker (1995) asserts "Listening is a magnetic and strange thing, a creative force. The friends who listen to us are the ones we move toward, and we want to sit in their radius. When we are listened to, it creates us, makes us unfold and expand."

2.1 Classifying and supporting different modes of listening

Sound and listening has been the subject of much attention in the film community, and here Sonnenschein (2001), an authority on cinematic sound design, notes the active nature of listening, entailing filtering, selective focus, links to memory and the capacity to respond. Sonnenschein cites the film theorist's Chion's three listening modes: reduced, concerned only with the parameters of a sound and not its source or meaning; causal, where the listener is identifying the sound source; and semantic, relating to the spoken language or other symbolic codes (Chion, 1994). To this Sonnenschein adds a further mode, referential listening, which implies links to the emotional connotations and meaning of the sound as well as an awareness of its context. When constructing soundscapes for films, sound designers consciously use sounds in the manner suggested by such a classification. Within this, while dialogue, almost always takes precedence, effects and music often vie for prominence within a mix. Music usually is regarded as providing the emotional backdrop while effects supply almost all of the context. All three combine to form a soundtrack that "is communicable and valid but unanalyzable" (Doane, 1985), designed to elicit emotions, invoke feelings and set moods. No matter what the method of production for dialogue, music or effects, between them, as illustrated below, they support each of the three levels of listening proposed by Chion and extended by Sonnenschein.

Reduced listening Here we have the rumble of a 'star ship', the sound of tyres on the road to indicate movement from within a car, or general 'room tone' to indicate that the characters inhabit a larger space than is visible. All of these are termed 'atmos[pheres]' and are an essential part of any sound design. It is expected that they will become rapidly habituated, and they are typically introduced slightly above the mix and then reduced to sit back within it, as a way of emulating the way in which listeners experience and interpret real world environments. This 'selection' is essential as 'irrelevant' sounds have to be ignored in order to interpret what is either most useful or most interesting (Bordwell and Thompson, 1985). Most films adhere to the principle of 'starve the eye and feed the ear'. During the introduction to scenes or transitions between scenes 'atoms' plays an important part. One of the most useful applications of reduced listening for the designer is the ability to colour the sounds, this means affecting each sound's timbre so that the sound itself provides cues about the environment. A radio playing music with a lot of reverberation can suggest a large space - add the sound of gently lapping water and the listener thinks of an indoor swimming pool. Aesthetic cues are also put to use: the crackles and pops of a record player place the music within the scene, and therefore make it diegetic, with inherent narrative importance, but the intention of these additional sounds is not to be consciously heard but rather felt (Beauchamp, 2005). Effective use was made of this in *The Matrix* where rhythmic effects aid the perception of slow motion without the

shots appearing static. The sounds themselves are not important but their innate temporal dimension allows time to progress slowly while still retaining the audience's interest.

Causal listening Causal listening, which Bohme (2000) refers to as 'object orientated' and Metz (1985) calls the 'sound of what?' is by far the most important form of listening in film. Everything that can be seen on the screen that can make a sound potentially has to be heard, without imposing upon the dialogue. So the first thing any sound design team does is to list the sounds that are either essential or might enhance the film. This proves often to be of great advantage to film makers, as it can transform the artificial into the real: polystyrene rocks can have weight when rumbling towards the lead character, plywood doors become solid oak, and painted backdrops can appear animated. These sounds are commonly synchronised with the action and provide all of the physical cues about a sound generating object such as its mass, velocity, composition and vibration. The film director Robert Bresson states that he will replace an 'image with a sound whenever possible' as a 'sound always evokes an image; an image never evokes a sound'. A further example can be seen in *The Empire Strikes Back,* where the director Irving Kershner uses a pneumatic sound to suggest a door opening - in actuality this 'sleight of hand' was a single shot of a closed door cut straight to an open door without any intervening movement (Chion, 1994). Chion refers to 'synchresis' as being the bonding of a sound to a visible source. This allows Foley artists to replace human heads with water melons in horror films when a character's head explodes and baseball bats hitting leather baseball gloves stand in for human punches. This is a form of analogy, where isomorphic or iconic sounds can stand in for 'real-world' sounds, as long as they match the audience's perception of what the source's timbre and dynamics should be.

Semantic/referential listening Here is clearly where dialogue lies, but also music, especially themes such as those associated with the shark in *Jaws* and Darth Vader in *Star Wars*. But the sound designer also works in this area: as Smeagol strangles his friend to gain the ring in the *Lord of the Rings* trilogy, there are series of interrelated cries which express the character's emotional trauma as the ring calls out to him. Off screen sounds are routinely utilized in the horror genre to indicate that the yet unseen antagonist is approaching the heroes with murderous intent. An example from comedy is when the sound of a car crashing in *A Fish Called Wanda* indicates that an unseen character is yet again driving on the wrong side of the road and has caused another collision. Subtlety is also beneficial here: as a fridge door closes it can have the sound of glass vibrating indicating that it is full of bottles, which might mean that it is full of beer, or it might be hollow indicating that it is empty, or close with a solid thud, suggesting that it is full. In turn this can reinforce the sense that the character has a drinking or eating problem. Motifs are the most useful form of sense making sounds, making connections to 'story, plot and mood' (Burwell, 2003). Sound imagery usually takes the form of simile, hyperbole, metaphor, allegory, irony, paradox or vivification (Sonnenschein, 2001). With simile a clock ticking can suggest a heart beating regularly, implying that a character is trying to calm themselves before a confrontation. Hyperbole can take the form of the sound of fireworks when a champagne cork explodes to exaggerate the impression of a celebration. An example of a metaphor in sound design terms could be the groans of a nonexistent character as someone kneads bread, showing that they are taking out their anger on the bread, as a substitute for the unseen person. Allegory can take the form of the sound of chains clinking as a character walks, when there are no visible shackles, suggesting the

character's feeling of being imprisoned. Ironic sounds are the mainstay of comedy: the miniature pistol which produces the sound of a canon, while the much larger gun of the 'hero' creates merely a 'pop'. The contradiction of a paradox, where the sound of a child screaming can come from a car as it starts up, foretelling the car accident, this can be an effective device, especially as the scream was originally attributed to a malfunction in the car. The final form of imagery can be both sinister and comedic, an example is the sigh of a sofa as an obese character throws themselves onto it, with a further exhalation by the sofa when the character arises, as the object expresses relief at the removal of its burden.

Post-production mixing An experienced sound designer limits the number of key sounds that an audience has to interpret. This is achieved through judicious mixing, where after being introduced, a sound may be dropped in volume, or its frequency are altered so that there is no spectral overlap. But more often it is dropped, emulating habituation, as the sound has moved from providing information about object and event to effectively becoming the object of reduced listening. This form of mixing guides the audience through a series of key sounds so that they can make sense of the complex auditory environment they are inhabiting. Sound design in films thus has little to do with the presentation of reality, it is a technique, which if used successfully manipulates perceived reality, enhancing the narrative and contributing towards the suspension of disbelief and placing the viewer within the environment of the story. Designers are not concerned with recreating sounds: accuracy is confined to the genre of documentary, sound is used to extend the screen, highlighting truths and obscuring lies, through its use of synchrony and conventions. Sound is a 'rashomon' phenomenon, it exists only in the individual interpretations of its listeners (Altman, 1992), and if it is executed successfully, then audiences will remember visual elements that only existed aurally (Beauchamp, 2005).

In presence research, we find a similar framework underpinning the argument in Murray *et al.*, (2000) for the significance of sound in immersive virtual environments. Following Gilkey and Weisenberger (1995), Murray adopts Ramsdell's three levels of hearing as a framework to discuss an empirical study of induced hearing loss and its implications for VE design. Ramsdell's three levels are *social hearing* which concerns communication, *warning hearing* which relates to sounds that indicate something happening – the ring of the doorbell, or the boiling of the kettle and *primitive hearing*, relating to background sounds of which we are not normally consciously aware (Ramsdell, 1978). In Ramsdell's view, primitive hearing is essential for *psychological coupling*, the sense of active connection with the environment. Murray's data was obtained from studies of people with temporary, induced hearing loss walking around a university campus and attempting to communicate with others. The results indicate that support for both warning hearing and primitive hearing are necessary for a sense of presence in real environments, and by implication in their virtual counterparts, with the addition of social hearing in the case of environments where other people are present.

2.2 A three-level classification of listening

While several elements are common to the descriptions of listening summarized above, other useful points are unique to single accounts. Drawing, *inter alia*, on the work of Sonnenschein and Chion in film theory, Ramsdell in the study of deafness and Murray *et al.*, in presence research, and taking the corporeality of listening into consideration, we propose a three level account:

Pre-listening. This is pre-conscious and is prone to rapid habituation. The listener is not normally aware of sounds in this level of listening, but can comment on them if asked to do so ("there is a buzzing sound"). Sounds are not consciously attributed to their sources. Pre-listening is equivalent to Ramsdell's *primitive hearing*, and Chion's *reduced listening*. We also include such things as the startle reflex, alerting the body to potential threats – this is the most ancient from of listening phylogenetically.

Object-event listening: this might equally be described as *everyday listening* (Gaver, 1993), or *causal listening* (Chion) and would include Ramsdell's *warning hearing* (this is at a higher level than the startle reflex). In object-event listening we typically recognize and locate the sound source relative to our bodies and are able to judge the size and shapes of objects. ("The sound of a heavy door closing behind me", "the sound of a cat meowing a long way off").

The final category we describe as **sense-making listening**. This final category of listening is built upon object-event listening but involves making sense of the sound and often includes affective or autobiographical elements ("the sound of a door closing, which means it must be about 6.30pm as that's when my wife returns from work"). Similarly "the sound of a cat meowing" becomes – the sub-vocalized "he's not hungry again is he?" . The category is similar to Sonnenschein's *referential listening*. This is also the body in social context (*cf.* Merleau-Ponty's *intentional arc*, Merleau-Ponty, 1945*)*. Semantic listening to speech or other forms of audible communication would fall into this category, but are not of primary concern in this context.

There are interesting parallels here with the three-layer formulation of presence and self to be found in Riva *et al.*'s (2004) three-layer model of presence. This draws on Damasio's theorisation (Damasio, 1999) of three aspects of self, the pre-conscious proto-self and two aspects of the conscious self, the core self and the extended self. The three layers of this model adapted in relation to presence are proposed to be:

- *proto presence* – the embodied aspect of presence relating to the differentiation of the self from the world;

- *core presence* – a process of selective attention to perceptual stimuli, supporting the discrimination of external reality from the contents of one's consciousness, dreams or memories;

- *extended presence* - which serves to assess the relationship and significance of events in the world in the context of the memories and so forth which make up the autobiographical self.

However, our treatment of listening emphasises an aspect of being-in-the-world that Riva *et al.*/Damasio neglect – that of intentionality. We now turn to the empirical investigation.

3. The empirical study

The work reported here is part of a larger study examining the relationship between sound, sense of presence and sense of place in real and artificial soundscapes, of which other aspects have been reported elsewhere (Turner *et al.*, 2003) Our hypothesis in this instance was that the 'throwness' of listening is such that all three forms of listening would be evident even in conditions where the soundscape is

clearly artificial. The part of the study discussed in this chapter had two conditions with 10 participants assigned to each:

1. Participants physically present in the Jack Kilby Computing Centre at Napier University (JKCC) (fig. 1.) The JKCC is a 500 seat, 24/7, very busy, open-access computing facility.

2. Participants located in a different room and seated at a table among in the midst of eight speakers and four sub basses reproducing the JKCC soundscape. See figure 2.

The soundscape itself was recorded in the JKCC. An eight channel system was used. Omni directional tie-clip microphones, using suspension mounts, were placed in an ellipse, at 1.3m in height (the approximate head-height of a seated listener). Spacing of microphones was chosen to correspond to the loudspeaker positioning during reproduction, so time delays would match. A thirty minute recording was made at 96kHz, 24 bit. This higher recording rate enabled a wider range of harmonics and a greater dynamic range to be recorded than is possible with normal CD rates (44.1kHz, 16 bit). This provides a theoretical frequency range of 20 Hz – 48 kHz and 0 – 148 dB, compared to 20 Hz – 22.05 kHz and 96 dB when using CD rates. The higher settings allow the recording of the more subtle aspects of the soundfield associated with the reverberation of the room, which is typically contained within the higher complex harmonics. The increase in the dynamic range also captured the quieter sounds, which are normally lost in the noise floor of recording equipment. It also allowed the accurate capture of transient peaks without the requirement for compression. The result of this is that the increase in rates more accurately reflects the experience of actually being in the environment, as a listener's hearing capabilities are typically well beyond that associated with any commercially available recording system.

The reproduction system employed a compact loudspeaker in place of each microphone and four sub bass monitors in order to extend the frequency range for lower frequencies. This was located in a quiet room and the positioning of each compact monitor matched exactly the original microphone positions. The system was calibrated to sound pressure levels during the recording. Participants listened to a continuous 15 minute extract from the recording, whilst sitting on a height adjustable chair which raised their ears to height of 1.3m, which allowed accurate alignment with the centre of the loudspeakers.

Figure 1: 500 seat JKCC

3.1 Participants and procedure

Twenty participants were invited to participate in the study and were randomly assigned to one of the two conditions. The participants varied with respect to their age, sex and background. All participants took part in the study on a voluntary basis and all had a high command of spoken English.

In both conditions participants were seated and requested to verbalise what they heard, as they heard it, during the fifteen minute session. They were informed that the session would last for about fifteen minutes and that they could ask any questions afterwards. They were also told that they could end the experiment at any point. Verbalisation was recorded using a microphone attached to the participant's collar. In condition 2 (artificial soundscape) participants were not informed of the location where the recording had been made. A questionnaire intended to probe sense of presence – adapted from the widely used Slater-Usoh-Steed instrument (Slater *et al.*, 1994) was administered at the end of the session.
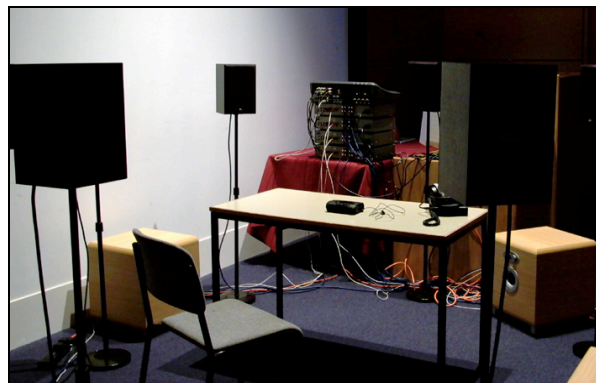

Figure 2: recreating the JKCC

3.2 Analysis and Results

Transcribed verbalisations were coded for one of the three forms of listening by the second and third authors, a sample of coding being cross-checked. There was evidence of all three forms of listening in both real and artificial conditions. Table 1 shows the number of participants making each type of verbalisation. All participants verbalised instances of object-event listening. In the real soundscape, the verbalisations of one of the ten participants were limited to object-event listening and a further four verbalised object-event and sense-making listening only. In the artificial condition, three participants verbalised object-event and sense-making listening only. There was further variation between individual participants in the *proportion* of verbalisations for the different forms of listening.

	Real soundscape	Artificial soundscape
Pre-listening	5	3
Object-event listening	10	10
Sense-making listening	9	10

Table 1: Number of participants verbalising each type of listening in the real and artificial soundscapes

The following are representative examples of verbalizations of each type of listening:

Pre-listening

"Slight peeping noise" (Real soundscape)

"General background hum" (Real soundscape)

"Full of ambient noises" (Real soundscape)

"Lots of banging noises" (Real soundscape)

"General quite low noise" (Artificial soundscape)

"Continual clicking" (Artificial soundscape)

"I can hear tapping" (Artificial soundscape)

"The elements of the background noise, they are changed from being such a low drone to being a higher pitched noise" (Artificial soundscape)

Object-event listening

"But mostly it's typing and keyboards" (Real soundscape)

"Constant sound of paper being binned" (Real soundscape)

"There's a mobile phone somewhere and some laughing coming from back there somewhere" (Real soundscape)

"Could be photocopier sounds as well" (Real soundscape)

"I can hear people walking about now" (Artificial soundscape)

"Sounds like someone opening curtains or blinds" (Artificial soundscape)

"Must be walking downstairs… walking along a wooden floor and then onto a carpeted floor" (Artificial soundscape)

"Someone scrunching up a bit of paper and then throwing it away" (Artificial soundscape)

Sense-making listening

"And a lot of coughing, as if everyone has a cold at the same time" (Real soundscape)

"A girlie laughter – someone reasonably attracted to someone else" (Real soundscape)

"Still talking to a pal... and the noise is annoying behind me... still making noise" (Real soundscape)

"The beeping of the computer getting terribly upset" (Real soundscape)

"It sounds like I am sitting in an office somewhere – some high-ceilinged office" (Artificial soundscape)

"It's really quite annoying actually" (Artificial soundscape)

"Must be a place where everybody is very unwell... coughing and sneezing keeps going on." (Artificial soundscape)

"Sounds like people typing and just working in an office, I think" (Artificial soundscape)

Both conditions also generated a number of verbalisations which placed the listener corporeally in the soundscape.

"I'm next to the stairs, so you can hear the people walking up and down as well" (Real soundscape)

"More steps behind me" (Real soundscape)

"Laughing on my left" (Real soundscape)

"Somebody using a mouse on my left" (Real soundscape)

"Just a few people behind me slightly to my left" (Artificial soundscape)

"More hammering away on my right" (Artificial soundscape)

"Someone exhaling… a low voice whispering directly behind me." (Artificial soundscape)

"More tapping to my left again." (Artificial soundscape)

3.3 Discussion

The results suggest that the listening experience in this very high fidelity but artificial soundscape was very similar to the experience in its real-world counterpart. All three forms of listening were evident in both conditions, participants in the (very evidently) artificial condition still being able to make sense of the sounds, relate them to their bodies, their own experience and affective state and the perceived intentions and affect of others. The nature of the comments suggests that to some degree, these participants found themselves in a real place, even though this was not always identified as that where the soundscape was captured. It is also striking that the 'throwness' of listening compels people to interpret the soundscape about them despite the artificiality of the setting.

Instances of pre-listening are comparatively uncommon in both conditions. This is not surprising: (i) we were asking people to comment on something which is normally pre-conscious and (ii) the act of complying with a request to describe what they heard is likely to have prompted ascription of a source to the sound – object-event listening. We are considering how pre-listening might be captured more effectively. Although almost all participants verbalized instances of sense-making listening, there are also relatively few instances of this mode, but (by inspection) no systematic difference between conditions. There is also a notable difference between individuals in the relative proportions of object-event listening and sense-making listening, and in ego-centric localization of sounds. It is unclear how far this reflects differences in listening behaviour, verbal adeptness, expressive style, spatial ability or other cognitive characteristics. We have observed this effect before in analyzing free-form verbal responses in a similar context (Turner *et al.*, 2003): disentangling the 'nuisance' effect of individual differences remains an issue for place and presence research. In this instance a within-subjects comparison between conditions would have been a possibility, but this would have in turn raised issues around the effects of familiarisation. More interestingly, the data from the presence questionnaire from the participants in the artificial condition suggests that a relationship may exist between forms of listening and perceived presence – we would hypothesise that sense-making betokens a greater degree of intentionality, engagement and hence, conceivably presence – but with only 10 participants a further study is necessary before this can be more than a weakly indicative finding. Such further work would be facilitated by the development of a quantitative 'listening scale' which would complement qualitative data collection.

4. Conclusions

This work has argued that, as listening locates our corporeal selves in the world, and the 'throwness' of listening is so compelling, a consideration of listening is a core

element in presence research, and in particular, where the aim is to create a sense of presence in a specific place. We have drawn on previous work to define a three level classification of listening and applied this to an empirical study of real and artificial soundscapes, where evidence of all three forms of listening in each condition was identified. Further work is required to identify the relationships between dominant modes of listening, presence and sense of place, to operationalise the three-fold descriptions as a measurement scale and to explore how best to support all three forms of listening in virtual 'places'. The empirical work reported here used a high-fidelity recording, but the use of sound design techniques from film media have intriguing potential. Here comparative trials are likely to be fruitful. Finally, on a theoretical plane, we argue that the treatment of listening as an *intentional* phenomenon advanced in this chapter contributes to the more general understanding of presence and self.

5. Acknowledgements

6. References

Altman, R. (1992) The Material Heterogeneity of Recorded Sound. In R. Altman (Ed.) *Sound Theory/Sound Practice*, New York: Routledge, 15-31.

Bachelard, G. (1994) *The poetics of space*. Boston: Beacon Press.

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpina, C., and Alcaniz, M. (2000) Presence and reality judgment in virtual environments: A unitary construct? *CyberPsychology and Behaviour*, **3**, 327-335.

Beauchamp, R. (2005) *Designing Sound for Animation*. Oxford: Focal Press.

Bohme, G. (2000). Acoustic Atmospheres. *Soundscape*, Spring 2000, 14-18.

Bordwell, D. and Thompson, K. (1985) Fundamental Aesthetics of Sound in the Cinema. In E. Weis, E. and J. Belton, J. (Eds.), *Film Sound: Theory and Practice* New York: Columbia University Press, 181-199)

Bormann, K. (2005) Presence and the Utility of Audio Spatialization. *Presence: Teleoperators & Virtual Environments*, **14(3)**, 278-297

Brunart, D. S. (2002) Near-Field Virtual Audio Displays. *Presence: Teleoperators & Virtual Environments*, **11(1)**, 93-106.

Burwell, C. (2003) Composing for the Coen Brothers. In L. Sider,, D. Freeman and J. Sider, (Eds.), *Soundscape: The School of Sound Lectures 1998 – 2001*, London: Wallflower Press, 195-208.

Canter, D. (1997). The facets of place. In G. T. Moore and R. W. Marans, (Eds.), *Advances in Environment, Behavior, and Design, Vol. 4: Toward the Integration of Theory, Methods, Research, and Utilization*. New York: Plenum, 109-147.

Casey, E.S. (1997). *The Fate of Place*. Univ. of California Press, Berkeley & Los Angeles

Chion, M. (1994) *Audio-Vision: Sound on Screen*. New York: Columbia University Press.

Chueng, P., Marsden, P. (2002) Designing Auditory Spaces to Support Sense of Place: The Role of Expectation. Position paper for *The Role of Place in On-line Communities* Workshop, CSCW2002, New Orleans.

Coyne, R. and Parker, M. (2008) Voice and Space. In *The Exploration of Space, Spatiality and Technology*, P. Turner and S. Turner (Eds.), IGI Global Press.

Damasio, A. (1999). *The feeling of what happens: body, emotion and the making of consciousness*. San Diego: Harcourt Brace.

Dinh, H. Q., Walker, N., Song, C., Kobayashi, A., & Hodges L.F. (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. *Proceedings of the IEEE Virtual Reality 1999*, 222-228.

Doane, M. A. (1985) Ideology and Practice of Sound Editing and Mixing. In E. Weis and J. Belton (Eds.), *Film Sound: Theory and Practice* , New York: Columbia University Press, 54-62.

Dreyfus, H. L. (1991) *Being-in-the-world: A Commentary on Heidegger's Being and Time, Division 1*. Cambridge, MA: MIT Press.

Freeman J, Lessiter J, Keogh E, Bond FW, & Chapman K. (2004). Relaxation Island: virtual, and really relaxing. *7th Annual International Workshop on Presence, PRESENCE 2004*, at Polytechnic University of Valencia, Valencia, Spain, 13-14-15 October 2004.

Gaver, W. W. (1993) How do we hear in the world? Explorations of Ecological Acoustics. *Ecological Psychology*, **5(1)**, 1-29.

Gibson, J. J. (1966) *The Senses Considered As Perceptual Systems*. Boston: Houghton Mifflin.

Gilkey, R. H., & Weisenberger, J. M. (1995). The sense of presence for the suddenly deafened adult. *Presence: Teleoperators and Virtual Environments*, **4(4)**, 357-363

Haldrup, M. and Larsen, J. (2006) Material Cultures of Tourism, *Leisure Studies*, 25:3, 275 - 289

Heidegger, M. (1927/1962) *Being and Time*. (Translated by J. Macquarrie & E. Robinson) New York: Harper Collins.

Hendrix, C., & Barfield, W. (1996). The sense of presence within auditory environments. *Presence: Teleoperators and Virtual Environments*, **5(3)**, 290-301

Hull, J. (1990) *Touching the Rock: An Experience of Blindness*. London: SPCK Publishing.

Ihde, D. (1976) *Listening and the Voice*. Athens: Ohio University Press.

Jorgensen, B. S. and Stedman, R. C. (2001) Sense of place as an attitude: Lakeshore owners attitudes towards their properties. *Journal of Environmental Psychology*, **21**, 233-248.

Larsson, P., Västfjäll, D., and Kleiner, M. (2001) The actor-observer effect in virtual reality presentations. *CyberPsychology and Behavior, ***4**, 239-246.

Lessiter, J., Freeman, J., Keogh, E. and Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments*, **10**, 282-298.

Merleau-Ponty (1945) *Phénomènologie de la perception* , Paris: Gallimard

Metz, C. (1985) Aural Objects. In E. Weis and J. Belton (Eds.), *Film Sound: Theory and Practice*, New York: Columbia University Press, 154-161.

Murray, C.D., Arnold, P. and Thornton, B. (2000). Presence Accompanying Induced Hearing Loss: Implications for Immersive Virtual Environments. *Presence: Teleoperators & Virtual Environments*, **9(2)**, 137-148.

N.P.R. (n.d.) http://www.npr.org/about/pitch/postcards.html (last accessed 17/08/07)

National Geographic (n.d). http://www.nationalgeographic.com/xpeditions/activities/10/gasounds.html (last accessed 17/08/07)

Nichols, S., Haldane, C., & Wilson, J. R. (2000) Measurement of presence and its consequences in virtual environments. *International Journal of Human Computer Studies*, **52**, 471-491

Pressing, J. (1997) Some perspectives on performed sound and music in virtual environments. *Presence: Teleoperators & Virtual Environments*, **6(4)**, 482-502.

Ramsdell, D. A. (1978). The psychology of the hard-of-hearing and deafened adult. In H. Davis & S. R. Silverman (Eds.) *Hearing and Deafness* (4th ed). New York: Holt, Rinehart & Winston, 499-510

Relph, E. (1976) *Place and Placelessness*, London: Pion Books

Riva, G., Waterworth, J. A. and Waterworth E. L. (2004) The Layers Of Presence: A Bio-Cultural Approach To Understanding Presence In Natural And Mediated Environments. *CyberPsychology & Behavior, ***7(4)**, 405-419.

Robertson, J., de Quincey, A., Stapleford, T. and Wiggins, G. (1998) Real-Time Music Generation for a Virtual Environment. *Proc. ECAI 98 Workshop on AI/ALife and Entertainment*.

Rodaway, P. (2001) *Sensuous Geographies*. London: Routledge.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, **10**, 266-281.

Serafin, G. and Serafin, S., (2004) Sound Design to Enhance Presence in Photorealistic Virtual Reality, *Proceedings of the 2004 International Conference on Auditory Display*, Sidney, Australia, July 6-9.

Sheridan, T.B. (2004) Musings on Music Making and Listening: Supervisory Control and Virtual Reality. *Proc. IEEE*, **92(4)**, 601-605

Sonnenschein, D. (2001) *Sound Design: The Expressive Power of Music, Voice and Sound Effects in Cinema*. Studio City, CA: Michael Wise Productions.

Stanney, K.M., Mourant, R.R. and Kennedy, R.S. (1998) Human factors issues in virtual environments. *Presence: Teleoperators & Virtual Environments*, **7(4)**, 327-351

Storms, R. L. and Zyda, M. J. (2000) Interactions in Perceived Quality of Auditory-Visual Displays. *Presence: Teleoperators & Virtual Environments*, **9(6)**, 557-580

Tuan, Y.-F. (1977) *Space and Place*. Minneapolis: University of Minnesota Press.

Turner, P., McGregor, I., Turner, S. and Carroll, F. (2003) Using Soundscapes to Create A Sense of Place. *Proc. Int. Conference on Auditory Display*, Boston, MA, USA, 6-9 July 2003.

Turner, S., Turner, P., Carroll, F., O'Neill, S., Benyon, D., McCall, R. and Smyth, M. (2003). Re-creating the Botanics: towards a sense of place in virtual environments. *Proc. Of UK Conference on Environmental Psychology*, Aberdeen.

Winograd, T. and F. Flores (1986) *Understanding Computers and Cognition: A New Foundation for Design*. Norwood NJ, Ablex.

Witmer, B. G., & Singer, M. J. (1998) Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, **7**, 225-240.

Zahorik, P. and Jenison, R. L. (1998) Presence as Being-in-the-World. *Presence: Teleoperators and Virtual Environments*, **7(1)**, 78-89.

---

[i] Coyne and Parker's discussion of voice and space in this volume is a notable exception (Coyne and Parker, 2008).