INVESTIGATING SPATIAL MUSIC QUALIA THROUGH TISSUE CONDUCTION

Peter Lennox
University of Derby, UK
P.Lennox@derby.ac.uk

Ian McKenzie
University of derby, UK
I.McKenzie1@derby.ac.uk

ABSTRACT

A spatial array of vibro-mechanical transducers for bone-and-tissue conduction has been used to convey spatial ambisonic soundscape and spatial musical material. One hundred volunteers have undergone a five-minute listening experiences, then have described the experience in their own words, on paper, in an unstructured elicitation exercise. The responses have been aggregated to elicit common emergent descriptive themes, which were then mapped against each other to identify to what extent the experience was valuable, enjoyable and informative, and what qualia were available through this technique. There appear to some substantive differences between this way of experiencing music and spatial sound, and other modes of listening. Notably, the haptic component of the experience appears potentially informative and enjoyable. We conclude that development of similar techniques may have implications for augmented perception, particularly in respect of quality of life (QoL) in cases of conductive hearing loss.

1. INTRODUCTION

This paper describes progress in investigating the experienced properties of spatial music delivered through an apparatus featuring multiple transducer locations situated on the cranium. As the listener's ears are unoccluded, residual air-conduction hearing is unaffected. This technique is a non-invasive augmentation of existing sensory capabilities; it can be efficacious in respect of the conductive component of hearing loss, but not sensorineural components.

The goal is to characterise, and subsequently quantify dimensions of perceptual experience in relation to music listening in this manner. Is the experience meaningful, in what ways, and in what ways might it differ from other music listening modes?

The aim of providing the spatial aspect of music follows the general trend in current music production and reproduction toward enhanced spatial attributes such as image focus [1], localizability, motion, spaciousness [2], and ensemble depth [3]. There may be substantive distinctions between the desirable and feasible spatial attributes of musical experience vs. those for spatial environment listening. Nevertheless, many of the experiential attributes that may be available through multi-speaker systems would not be feasible in conventional tissue or bone conduction techniques (discussed in section 4). Hence those with some degree of bilateral or unilateral hearing impairment do not have ready access to the kinds of musical experience available to listeners with unimpaired hearing.

Whilst communications difficulties in hearing impairment are receiving increasing attention, the quality-of-life (QoL) implications of music deprivation have received less. Assistive technologies for speech comprehension do not currently adapt well to music listening. [4]

In the first stage, to avoid a reductionist approach and obviate the need for trained listeners, we adopted an unstructured elicitation methodology (for discussion, see [5], [6]**)** whereby the prototype apparatus (discussed in section 5) was demonstrated over several days in various venues. Subjects were self-selecting, were given no instructions as to what to listen for and briefly recorded their initial impressions on paper after listening for approximately 5 minutes; some volunteered to repeat the experience on subsequent days. We observed variations in volunteers' responses (discussed in section 6). We did not aggregate data on known hearing impairments.

2. SENSORY AUGMENTATION FOR AUDIO MATERIAL

Approximately 5% of the World's population, that is, 360 million people, suffer from "disabling hearing loss" [7] and the proportion of over-65s rises to about 33% [8]. 13.4% of geriatric patients have significant conductive components to their hearing loss [9].

A substantial proportion of the population are subject to 'music deprivation' and inasmuch as music listening contributes to people's sense of wellbeing or "Quality of Life" (QoL), this deprivation may have significant and long-term health and wellbeing consequences. Assistive technologies implementing sensory augmentation could ameliorate the effects of lack of ready access to music, the experiential attributes of music listening can be reinstated and tangible benefits might accrue.

We distinguish sensory augmentation from sensory substitution in that the aim is to extend perception, not to substitute. However, augmentation might itself be augmented with some elements of substitution, and so the concepts overlap. Multimodal presentations of certain classes of information might provide richer experiences. Vibrotactile stimuli can be used to enhance perceived low frequency content, emphasize transients and steering of spatial auditory perception [10]. Philosophically, we think in terms of 'information channels' rather than direct sensory equivalence.

Such multimodal interactions will be subjects of future investigative work.

3. SPATIAL MUSIC

There has, in the late 20th and early 21st centuries, been burgeoning interest in spatial, surround or 3-dimensional music. The subject can be discussed in engineering, aesthetic and perceptual terms [11][12][13]. The underlying principles are that spatial (as against "non-spatial") music might provide enhanced experience in terms of involvement and immersivity. In information-transmission terms, incorporating spatial parameters facilitates greater information throughput, allowing finer detail to be depicted and discerned.

It is acknowledged that perceptual tasks in music listening differ from those in environment listening. In the latter, requirements for timely detection of threat and reward are presumed to have exerted evolutionary influence on phylogenetic development. Notwithstanding, exaptation [14], whereby evolved mechanisms or capabilities can become co-opted for other uses, provides that our spatial abilities are available for the experiencing of music.

Stereo [15] provides for a loudspeaker-feed signal set that generates interaural differences (in amplitude and phase) at the ideal listening position that can produce the powerful illusion of a left-right discriminable stage with multiple spatially-separate musical sources, either static or moving. Additionally, spatial reverberant fields (captured or synthesized) can give some sense of ensemble depth (some sources closer than others) and spaciousness. The effect is of a proscenium arch presentation. The stereo signal set can be listened to over headphones; however the effect is generally of a soundstage distributed left-right between the ears, giving a particular "in-the-head" experience. A binaural signal set can be used (either binaurally recorded or synthesized) to promote "externalization" (for a discussion see: [16]) and in the optimal case, where the head-related transfer function (HRTF) used in the production of the signal set closely matches the HRTF of the listener, strong impressions of an externalized, three-dimensional environment can ensue. However, such close matching is rarely feasible and the usual experience falls short of the theoretical optimum.

Surround sound, where a complex signal set is fed to multiple loudspeakers surrounding the listener(s) can depict many source-locations, movements and a sense of being immersed in a whole spatial environment. However, perceptions of depth-of-field (variations in perceiver-source distance) remains limited. Systems range from fairly simple (e.g. Dolby 5.1 surround) to complex (e.g high-order ambisonics or wave field synthesis).

The spatial qualia engendered by the various approaches differ; a large and complex system may well give experiences of large environments but may be less competent in producing "intimate" ones with sources close to the listener. The converse is generally the case with small, intimate systems. Composers of spatial music are thus constrained in what qualia they can attempt to offer.

For discussion of spatial music compositional concerns, see for example [17].

In all the above cases, listeners with bilateral or unilateral hearing deficits will experience degraded spatial musical qualia, reducing immersion and impairing enjoyment of the material.

4. SPATIAL TISSUE CONDUCTION

Auditory perception elicited by means of mechanical transduction, i.e. a tuning fork pressed against the cranium, has long been known. Single vibro-tactile transducers have been in use in audiology and the hearing aid industry for decades. Until fairly recently spatial audio was not thought possible through tissue conduction, theorised interaural level differences due to interaural attenuation were not considered sufficient; studies have shown this not to be the case [18][19][20]. In all three experiments to assess lateralisation, stimuli were presented bilaterally with transducers placed in contact with either the mastoid process behind the ear or the condyle just in front of the ear; all produced similar results to that of headphones. These experiments indicate that when sound is presented through tissue conduction we still make use of the same binaural cues as for air conducted (AC) sound.

Auditory localization is dependent on the physiological and anatomical properties of the auditory system as well as behavioral factors. The textbook primary cues for auditory localization are interaural differences and spectral cues [21][22][23]. The ridges and folds in the outer ear reflect and absorb certain frequency components of a sound wave, the spectral characteristics of a sound wave will differ if approaching the ear from different directions. Due to the shape of the pinnae providing this filtering effect the elevation and position of sound sources is encoded in direction-dependent spectral cues allowing us to localize sound sources. Many literary sources agree that vertical information derives exclusively from position-depending differences in the frequency filtering properties of the external ear.

Whilst interaural differences akin to air conduction may result when sound is presented through tissue conduction, no sound is presented to the outer-ear specifically the pinnae and vertical information should be absent; some comments suggest this is not the case. This anomaly may arise out of fine differences in arrival times caused by propagation along multiple signal pathways from transducer to the basilar membrane. There is also an intriguing possibility of multimodal cueing; binaural auditory cues merging with additional information provided through the somatosensory system via haptic cues [10][24][25][26]. When using a multiple transducer array vertical information is available to the listener as well as externalisation of the perceived sound; how this is the case continues to be the subject of further investigation.

5. APPARATUS, MUSICAL MATERIALS AND LISTENING CONDITIONS

5.1 Apparatus

Psychophysical investigation of the dimensions of experience of spatial tissue-conduction listening will prove useful, but for many first-time listeners, bases for comparison may be lacking; a training period targeting specific attributes may be required. Determining what those attributes might be is the aim of the present study.

Figure 1. Multiple transducer array

Sounds presented at:

- 1-left mastoid
- 2 – 25mm above left temple
- 3: between forehead and vertex
- 4: 25mm above right temple
- 5: right mastoid

Figure 2. BCT-2 10W transducer

A prototype headset transducer array using five BCT-1 8Ω 90dB 1W/1 m tactile transducers has been used to display a range of spatial soundscapes and music. Each transducer receives a discrete signal set through an individual amplifier; a Focusrite PRO 26 i/o interface provides fire-wire connection to a mac mini running Reaper DAW. A single BCT-2 10W transducer was also available for listeners to position on the jaw, zygomatic arch or back their head/neck. A set of banded style 3M Ear Plugs were available for listeners to use and compare the experience with the plugs in vs out.

5.2 Listening Materials

Environmental and musical stimuli was processed using a variety of effects and routed in different formats; stereo, modified stereo, ambisonics and direct feed. A 1st order ambisonic recording of a country park captured using a Soundfield ™ microphone provides the ambient background; stereo recordings of bird sounds, a steam train and music alongside mono FX clips were used to create the soundscape. Signals were processed using Reaper ® DAW; signals were spatially encoded using WigWare 1st order ambisonic panning and decoded through a WigWare 1st order periphonic ambisonic decoder patched to the transducer array.

5.3 Listening Conditions

The proto-type has been on demonstration at IOA Birmingham, ICMEM Sheffield and PLASA London. At PLASA we recruited one hundred untutored listeners, with a mixture of expertise; none reported experience of tissue conduction. Auditions were of five minutes duration, no prior instructions were given and volunteers were invited to record initial reactions and observations on paper immediately after auditioning. The listening tests took place in non-ideal conditions, as part of the Exploratorium exhibit we shared the space with four other exhibitors. The Exploratorium was located on the upper level of the large exhibition hall, a large footfall and other exhibitors using amplified sound produced a considerable noise floor.

5.4 Limitations

Equipment and calibration: The transducers in use have the following known limitations:

- Frequency response: 200Hz to 16 KHz, low frequencies are not well served, resulting in a 'thin' sound for some musical material.
- Component matching: the manufacturers do not publish information on performance matching.

With a cohort of 100 and a wide variety of head sizes, precision in determining matched contact force for all transducers was infeasible, possibly resulting in different spatial experiences for different listeners. Additionally, as audiological testing was impossible, variations in hearing acuity could not be taken into account

The demonstrations took place in an environment with high levels of ambient sound, especially in vocal ranges, entailing concomitant constraints on dynamic range and hence subtlety of detail.

The method of recording responses proved to be suboptimal, as many volunteers described the experience in greater detail verbally than subsequently on paper.

6. RESPONSES AND ANALYSIS

The responses were tabulated for analysis to identify key themes.

Of interest were the variations in descriptive language across such a mixture of untrained listeners varying in age, gender, expertise and listening ability. A broad synonymic approach was taken, whereby terms were loosely grouped to form themes. So, for instance, the category "weird" included terms such as "eerie", "strange" and "unusual".

Theme	Descriptors in Class
Positive	Nice, Incredible, Amazing, Awesome, Excellent, Loved, Good, Enjoyed, Cool, Wonderful, Extraordinary, Impressive, Effective
Negative	Muddy, Muffled, Lacking, Limited, Quiet, Dull, Distortion
Hearing Loss	Hearing Loss
Spatial	Spatial, Surround, 3D, Virtual Reality, Image location, Movement, Image positioning, 360 soundfield, External
Clarity	Clarity, Clear, Crisp, Pure
Interesting	Interesting, Fascinating
Weird	Weird, Unusual, Surreal, Strange, Uncanny, Ethereal, Eerie, Bizarre
Vibrations	Vibrations, Tickling, Tickling
Feel	Feel, Felt, Feeling, Natural, Sensorial
External	Distant, Immersive, Overhead, Above, Around, Spacious, Outside
Headphones	Headphones

Figure 3: Main themes and descriptive terms

The emergent key themes in aggregated comments were: "positive" [77%] (expressed as having enjoyed the experience), "spatial" [38%] (including surround/surrounded, spacious, distant, immersive, above etc.), "interesting" [38% (including "fascinating", "amazing" "incredible" etc.,), "weird" [23%] (including "eerie", "strange" "unusual"), "vibrations" [24%] (expressed directly as vibrating, vibrations) , "clarity" [22%] (clear, pure,), "feelings" [28%] (distinct from vibrations, such as "felt very pleasant" "felt dreamlike" "felt like I was in the soundscape"), and "negative" [19%] (expressed as "not clear enough" or "couldn't hear the bass"). A complication arose in the overlap of the positive and negative categories, 10**%** of respondents gave comments that included both. 14% of comments were classed as "neutral".

6.1 Participant comment samples

1) Male age 30, Sound Engineer, non-musician.

"Very surreal distant sounding. Passing sounds such as the train and plane felt closer and move forward. The higher sounds such as water felt harder to make out. Fidelity sometimes felt lost when many sounds were overlapped. As strange as it sounds it was like a memory or dream of a sound."

Recorded classes for comment 1:

Positive; Negative; Spatial; Surround; Feel; External; Weird.

2) Female age 36, Stage Manager, non-musician.

"Although the sound was still 'one sided ' to a certain degree I felt for the first time that I was immersed in a soundscape and that my hearing loss was not making me lose out on part of the effect. The train in particular really felt 360, especially with the chin transducer on my right cheek bone."

Recorded classes for comment 2:

Positive; Hearing Loss; Spatial; External; Feel.

3) Male age 62, Concert Producer, Musician

"Sounded slightly "muffled" some spatial "separation" but not dramatic"

Recorded classes for comment 3:

Negative; Spatial.

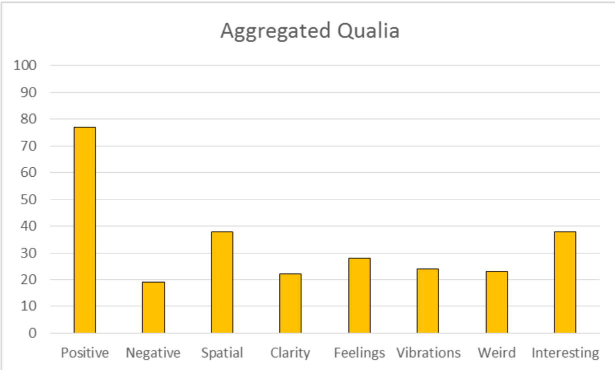

Figure 3. Aggregated Qualia

Notably, the degree of emphasis placed on each attribute (for instance, "quite spacious" or "very spacious") was not distinguished here.

6.2 Co-occurring themes

We then mapped each attribute class against "positive" to find what it was about the experience that people found rewarding.

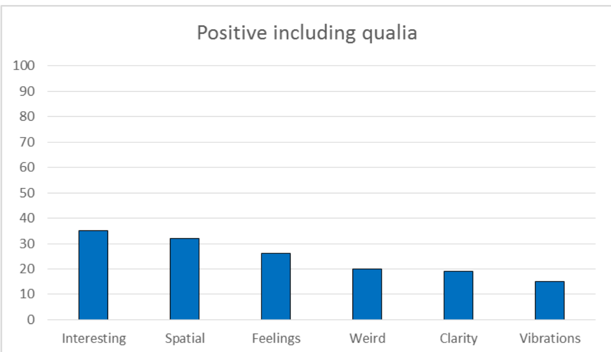

Figure 4: Qualia included with positive comments.

We found that the attribute class that mapped most strongly to positive expressions was the "interesting" category; 35% of comments included interesting and positive descriptors. This was followed by the spatial class – 38% had used spatial terms; 32% had used spatial *and* positive terms. Those that referred to the way they felt about the experience also correlated highly with positive – 26% featured positive *and* feelings. Clarity was referred to in conjunction with positive comments in 19%, "vibrations" were mentioned along with positive comments in 15% of cases.

7. DISCUSSION

Results may indicate that tissue conduction is of more utility to some than others; variations in comments might also indicate variances in biomechanical and/or neurological auditory processing.

Some volunteers (3%) reported some degree of bilateral or unilateral hearing deficit, but nevertheless reported in spatial and positive terms. Some others (2%) reported spatial anomalies that might indicate a degree of unilateral deficit ("the sound field sounded shifted to the left") but it was not within our experimental purview to comment or diagnose. Likewise, some that had used very positive and clarity terms may actually have been observing differences between their normal air-conducted hearing and this experience.

The surprising concomitance of reports of vibrations (which we might have thought was an undesirable percept) and positive comments prompts us to speculate that program-material modulated haptic input can contribute to the experience.

Notably, in the case of the "weird category (23%), weird and positive comments appeared in conjunction in 11% of all comments; there may be an overlap in the "weird" and the "interesting" categories, depending on individuals' use of language. It does appear that the novelty of the experience may be conflated with positive reports, and this in itself does not imply improvement in informational throughput.

8. CONCLUSIONS AND FURTHER WORK

This work is the early stage of investigation as to what might provide valuable experiences in tissue conduction, for whom.

Early indications are that the qualia associated with this kind of spatial experience may be similar but not identical to those for binaural presentations. Hence more structured methodologies should not precisely mimic those for air-conducted hearing.

At this stage of prototypical development, display of spatial parameters cannot be deemed accurate. Precision localization (of sources) in terms of azimuth, elevation and source-perceiver range is currently infeasible. Nevertheless, the fact that some degree of externalization and sense of spaciousness were alluded to in listeners' observations, is of interest. Research into refined processing of the signal set dedicated to tissue conduction is indicated. Ambisonic encoding has been used as a methodological convenience; its advantages for some attributes (such as ambient spaciousness) might not be matched for others (such precision localization). Different spatial audio attributes may be favored in different applications, of which personal music listening is only one. Similarly, it may be that a single spatial music encoding regime will not be appropriate for all listeners.

This work has enabled us to identify the following development areas for future research:

Technological: improved signal processing, improved transduction, improved apparatus comfort, developments in multimodal stimuli.

Methodological: Precise characterization of listener hearing capabilities, investigation of training periods and of individual preferences for encoding. Parameterization of qualia for spatial music listening.

Possible benefits of competent spatial tissue-conduction apparatus include:

- Enhanced quality of life for those with conductive hearing loss, through access to personal music listening.
- Augmented private perception where unimpeded air-conducted hearing is required.
- Diagnostic procedures to identify and isolate conductive hearing loss components.
- Improved methodologies for the investigation of mutimodally-augmented perception.

9. REFERENCES

[1] G. Martin, W. Woszczyk, J. Corey and R. Quesnel "Controlling Phantom Image Focus in a Multichannel Reproduction System" in Proc. 107th Conference of the Audio Engineering Society (AES 107), New York, 1999

[2] J. Blauert, W. Lindemann, "Auditory spaciousness: Some further psychoacoustic analyses," in J. Acoustical Society of America (JASA), 1986, 80(2):533-42

[3] T. Neher, T. Brookes, F. Rumsey, "Unidimensional simulation of the spatial attribute 'ensemble depth' for training purposes. Part 1: Pilot study into early reflection pattern characteristics," in: Proc. Audio Engineering Society 24th International Conference on multichannel Audio (AES 24), 2003, pp. 123-137

[4] R. Einhorn "Observations From a Musician With Hearing Loss" in Trends Amplif. 2012 Sep; 16(3): 179–182. doi: 10.1177/1084713812468513

[5] R. Guski, "Psychological methods for evaluating sound quality and assessing acoustic information" *Acustica* 83, 1997 pp 765-774

[6] N. Ford, F. Rumsey, B. de Bruyn, "Graphical Elicitation Techniques for Subjective Assessment of the Spatial Attributes of Loudspeaker Reproduction – A Pilot Investigation" in proc. Audio Engineering Society (AES) 110th Convention, Amsterdam, 2001

[7] World Health Organisation "WHO global estimates on prevalence of hearing loss", Report on WHO website, 2012,

http://www.who.int/pbd/deafness/WHO_GE_HL.pdf?ua=1 last accessed Feb 2017

[8] World Health Organisation, “Hearing loss in persons 65 years and older based on WHO global estimates on prevalence of hearing loss; Mortality and Burden of Diseases and Prevention of Blindness and Deafness” report on WHO website, 2012, http://www.who.int/pbd/deafness/news/GE_65years.pdf?ua=1 Last accessed Feb2017

[9] R.R Ruby “Conductive hearing loss in the elderly” in J. Otolaryngology, 1986, 15(4): pp. 245-247

[10] A. Tajadura-Jiménez, A. Väljamäe, N. Kitagawa, and H. Ho, “Whole-Body Vibration Influences Sound Localisation in the Median Plane”, Proceedings of the 10th Annual International Workshop on Presence, Barcelona, Spain, Oct. 2007.

[11] M. Gerzon, “What's wrong with quadraphonics?” Studio Sound, 16(5), 1974pp. 50–56.

[12] R. Zvonar, “A history of spatial music: Historical antecedents from renaissance antiphony to strings in the wings,” in eContact!: The online journal of the Canadian Electroacoustic Community (CEC), 7(4), 2005, [Last accessed October 2015]

[13] P.P. Lennox, P. P. “The philosophy of perception in artificial auditory environments: Spatial sound and music,” PhD Dissertation, Dept. Music, Univ. York, UK. 2005

[14] S. Gould and S. Vrba, “Exaptation—a missing term in the science of form” Paleobiology 1982 8: 4-15.

[15] W. Snow, (1953). “Basic principles of stereophonic sound,” J. Society of Motion Picture & Television Engineers (SMPTE), 1953, 61, pp. 567–587.

[16] D. Griesinger, “General overview of spatial impression, envelopment, localization and Externalization,” In Proc. of the 15th International Conference of the AES on small room acoustics (AES 15), Denmark, 1998, pp. 136–149

[17] E. Bates “The Composition and Performance of Spatial Music” – PhD dissertation. Dept. Music, and Dept. Elect. Eng., Trinity College Dublin, Republic of Ireland, 2009

[18] R. Stanley, & B.N. Walker, “Lateralization of sounds using bone-conduction headsets”, Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES) San Francisco, CA. 2006 pp. 1571-1575

[19] J.A. MacDonald, P.P. Henry, & T.R. Letowski, “Spatial audio through a bone conduction interface”, International Journal of Audiology. 2006 45, pp. 595-599.

[20] S. Stenfelt, & M. Zeitooni, “Binaural hearing ability with mastoid applied bilateral bone conduction stimulation in normal hearing subjects”, Journal of Acoustical Society of America, 2013 (134), 1, 481-493.

[21] W. A. Yost, “Lateral position of sinusoids presented with intensitive and temporal differences”, Journal of the Acoustical Society of America, 1981 70:397–409, 1981.

[22] R. O. Duda, “Elevation dependence of the interaural transfer function,” in Binaural and spatial hearing in real and virtual environments (R. H. Gilkey and T. R. Anderson, eds.), Mahwah, New Jersey: Lawrence Erlbaum Associates, 1997 ch. 3, pp. 49– 75.

[23] B. C. J. Moore, “An introduction to the psychology of hearing”, Academic Press, London, fifth edition, 2003.

[24] M.A. Meredith, & B.E. Stein, B. E, “Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration”, Journal of Neurophysiology Published 1 September 1986 Vol. 56 no. 3, 640-662 DOI:

[25] C. Pantev, A. Wollbrink, L.E. Roberts, A. Engelien, & B. Lu¨tkenho¨ner, “Short-term plasticity of the human auditory cortex”, Brain Research 842 1999 192–199.

[26] A. Good, M.J. Reed, & F.A. Russo, “Compensatory Plasticity in the Deaf Brain: Effects on Perception of Music”, Brain Sci. 2014, 4, 560-574; doi:10.3390/brainsci4040560