Feel It in My Bones: Composing Multimodal Experience through Tissue Conduction

Peter Lennox[1], Ian McKenzie[2] and Michael Brown[3]

[1,2,3] University of Derby, UK
p.lennox@derby.ac.uk; i.mckenzie1@derby.ac.uk; m.brown2@derby.ac.uk

Abstract. We outline here the feasibility of coherently utilising tissue conduction for spatial audio and tactile input. Tissue conduction display-specific compositional concerns are discussed; it is hypothesised that the qualia available through this medium substantively differ from those for conventional artificial means of appealing to auditory spatial perception. The implications include that spatial music experienced in this manner constitutes a new kind of experience, and that the ground rules of composition are yet to be established. We refer to results from listening experiences with one hundred listeners in an unstructured attribute elicitation exercise, where prominent themes such as "strange", "weird", "positive", "spatial" and "vibrations" emerged. We speculate on future directions aimed at taking maximal advantage of the principle of multimodal perception to broaden the informational bandwidth of the display system. Some implications for composition for hearing-impaired are elucidated.

Keywords: Bone conduction, Tissue conduction, multimodal perception, multimodal composition, pallesthesia

1. Introduction

During experiments to elicit descriptive terms of the experience of tissue-conducted (often referred to as "bone-conducted") sound, we observed that many respondents described the experience with extra-auditory terms, such as vibration, and feeling, and some remarked that the experience was markedly different from other modes of listening. We take the implicature that they may be experiencing the sensory input multimodally, leading us to investigate the possibility of deliberately composing multimodal experiences using this technique.

2. Tissue Conducted Sound

The source of a sound is something we might usually think of as external to one's body, our own voice, although originating from within might easily be perceived as exclusively peripheral. Listening to a play-back of our voice may sound strange to us as it is not in agreement with the sound heard while recording it; a percentage of the sounds heard whilst talking propagates internally and is not recorded, this part is tissue conducted sound. As a means of propagation it has been known and used for centuries and is historically known as bone-conduction (BC); 16^{th} century to assess ear pathology, early assisted hearing devices late 18^{th} century and tuning fork tests (TFT) evaluating hearing and ear pathology 20^{th} century to present day. As bone is only one of several transmission pathways the term "tissue conduction" is preferred as a more comprehensive description.

During vibro-tactile stimulation, skin, bone, fluids, soft tissue and the cranium contents are all transmission pathways to various extents. Pathways are frequency dependant contributing forces that act on parts of the hearing system; inertial forces on the ossicular chain and cochlear fluids, distortion of the temporal bone and cochlear shell, occluded ear canal resonance and fluid pressures [1-4]. While the importance of each pathways contribution remains in debate there is general agreement that their summed contribution result in passive travelling wave motion along the basilar membrane leading to auditory perception. Cancellation experiments suggest basilar wave motion to be similar whether elicited through air-conducted, bone-conducted or tissue-conducted sound [1, 5, 6]. In spite of the similarity in basilar motion, it was long believed that cues available for sound localisation in air, interaural time difference (ITD), interaural level difference (ILD) and spectral cues would not be available during tissue conduction. Recent TC research reports some success in lateralisation when appealing to interaural level and time differences [7-9]. The outer ear is not subject to stimulation during tissue conducted sound and in the absence of spectral cues provided by the pinnae, localisation of sound sources should remain 'in the head' on the horizontal plane. We report an interesting anomaly while using a multiple vibro-tactile transducer array; in the absence of spectral cues from their pinnae listeners continued to experience elevation and varying degrees of externalisation.

2.1. One Hundred Listeners

One hundred volunteers (24 female 16 -60 yrs. 76 male 16 - 62 yrs.) were invited to take part in a listening experiment; subjects received no prior instruction on any target attributes. Participants were invited to take part then offer any observations they may have on the experience in the form of written comments [10, 11], participants were also asked to record their age, sex, occupation and whether or not they were musicians (details are being used in ongoing correlative analysis). Audition elicited during the experiment was through a multi-channel vibro-tactile transducer array; five BCT-1 8Ω 90dB 1W/1 m tactile transducers held in a tensioned framework exerting contact force with skull were used to elicit auditory spatial impressions through tissue conduction.

2.1.1. Equipment

The prototype transducer array using five tactile transducers was used to display a range of spatial soundscapes and music. Discrete signal sets were constructed in Reaper DAW running on MAC and routed to five individual 1W amplifiers via fire-wire connection through a Focusrite PRO 26 i/o interface. Each transducer receives a unique set as a channel, channel 1 - left mastoid process, channel 2 - 25mm above left temple, channel 3 - midway between forehead and vertex, channel 4 - 25mm above right temple, channel 5 - right mastoid process. Banded style 3M Ear Plugs were available for listeners to experience any perceived differences between plugs in vs out.

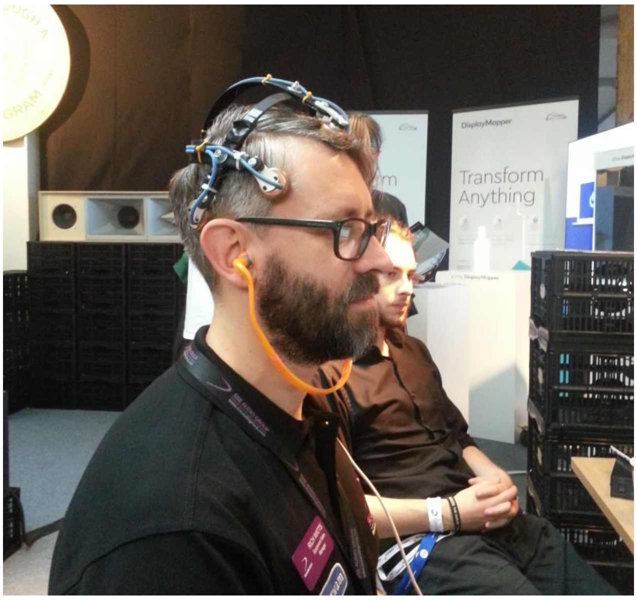

Fig. 1. Prototype five transducer headset array, PLASA London 2017.

2.1.2. Listening materials and conditions

Auditory stimuli was signal processed through a variety of FX plugins and routed simultaneously in different formats; stereo, modified stereo, ambisonics and direct feed. An ambient background was provided by a B format 1st order ambisonic recording of a country park captured using a Soundfield ™ microphone; stereo recordings of bird sounds, a steam train and music alongside mono FX clips were used to create the soundscape. Signal sets were spatially encoded using WigWare 1st order ambisonic panning and decoded through a Wig-Ware 1st order periphonic ambisonic decoder patched to the transducer array.

Listening tests took place across three days as part of the Exploratorium exhibit during PLASA London 2017 on the upper level of a large exhibition hall; a less than ideal high level noise floor was maintained throughout the day by sounds from other exhibitors and a large footfall. Attendees of the event who expressed any interest in our exhibit were immediately invited to take part in the listening test before any further discussion could take place. Participants were seated and the headset placed on their head, comfortable amplitude levels were arrived at during a short piece of music after which the five minute audition began. Upon completion participants were asked to record their comments on prepared forms (this method of recording data proved to be suboptimal as many of the participants described their experience in far greater detail during post-test discussions). Following the test period data was transcribed to a spreadsheet for manipulation and analysis.

2.1.3. Comments and Emergent Themes

The written comments were aggregated and analysed for emergent themes, primarily our interest lay in the arrays ability to convey auditory spatial information; 38% of the recorded comments were indicative of participants experiencing a 'sense of space' during the experiment. Other interesting themes include 'positive experience', 'interesting', 'spatial', 'vibration', 'weird' and 'clarity', figure 2 shows results correlated to 'positive'.

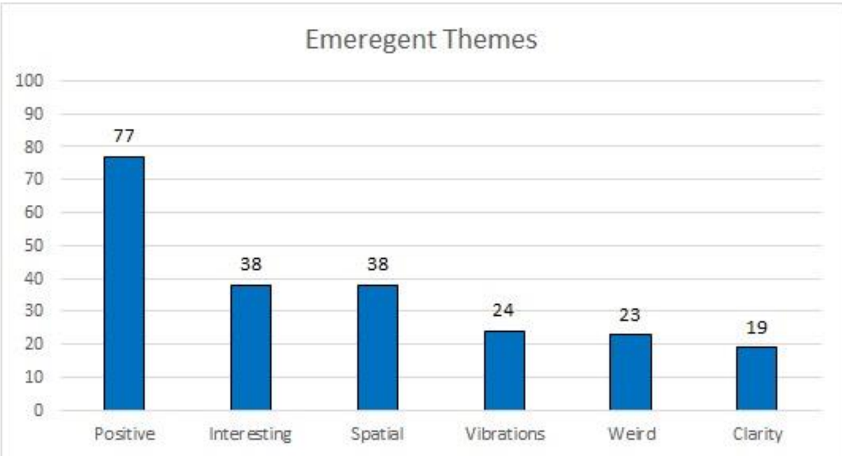

Fig. 2. Emergent themes from aggregated comments of 100 listeners correlated with comments of a 'positive' experience.

2.2. Multiple Pathways

In the absence of pinnae related spectral cues, we hypothesise in respect to comments indicating spatial impressions that while presenting signal sets via multiple pathways we may achieve a degree of perceptual equivalence through more abstract relationships. Propagation through tissue conduction by means of multiple, frequency dependant transmission pathways may be responsible for differences in frequency component arrival times within grouped and segregated sensory data [1-4, 12]. There is also an intriguing notion that plausible interaural time difference fluctuations are created as complex sounds interact during transmission to the cochleae, eliciting a sense of spacious envelopment [13]. Due to the complex nature of propagation during multiple-location tissue conduction it is not yet possible to model the signal arriving at the cochleae; methods by which we may examine this remain in development and will feature in future papers.

2.3. Vibration

An unexpected result emerged from the aggregated comments, 24% referred to vibration in a neutral or positive context; 12% of these where positively associated with comments on spatial impression. The result was unexpected as we had assumed that any vibration from the array would generate negative terms, this was only the case in 3% of the responses. We note that in addition to audition, the disturbance of mechanoreceptors in the body can give rise to a range of perceptions. What is normally meant by 'sense of touch' can be either passive *receptive* touch or active, exploratory *haptic* touch. The former refers to stimuli detected by mechanoreceptors (usually near the skin surface) whereas the latter incorporates proprioceptive mechanisms; we manipulate, pick up, detect weight and so on. In sensory psychophysics the work of Weber and Fechner discovered that despite the diversity of sensations, all sensory systems convey four basic types of information, modality, location, intensity, and timing [14]. In our context, using vibrotactile transducers, tactition is the appropriate term, and this has dimensions of frequency, intensity, spatial location and direction.

There is another important aspect to touch in the form of *pallesthesia* (sensitivity to vibration). This can take place at the surface of the body, but can also be detected internally in skeletal structure, balance organs (in mammals: the otoliths) and soft tissues such as the abdomen and eyeballs. The contributions of pallesthesia to everyday perception are not well understood; it is normally studied within the context of high-energy low frequency vibration in conjunction with health and safety concerns and annoyance factors. As a potential information-channel, pallesthesia is often overlooked, but *might* play a significant role in multimodal perception; the current broad consensus is that most, if not all, higher, as well as lower level neural processes are in some form multisensory [15, 16].

3. Multimodal Perception

Multimodal perception refers to percepts formed of combined information from two or more sense-modes, involving multisensory integration. In the past four decades there has been a burgeoning of interest in the subject, often framed in the context of the binding problem, which is related to understanding how multitudinous sensory qualia (within particular sense modes, or across several sense modes) are combined into unitary perceptions of objects and events [17].

At a philosophical level, the difficulties in precisely defining, individuating and counting discrete senses are well known [18]. Nevertheless, study of perception has tended to be specialised in particular sense modes. Vision scientists have had only passing acquaintance with theories of auditory perception and vice versa [19]. Haptic, olfactory, gustatory and somatosensory sciences have traditionally had similarly little intercourse. This conceptual reification also tends to affect art forms, so that pictures and music are produced by different people and often enjoyed in separate places.

At the level of real world experience, unimodal sense-perception is a peculiarity, rarely occurring in real-environment perception; real events naturally produce information that is potentially available to two or more sense-modes.

Characterising multimodal perception in terms of performance, variance and neural mechanisms is therefore a central challenge in the study of perception.

Evidence from neuroscience indicates that multisensory neurons are ubiquitous in the central nervous system, at peripheral as well as central levels [20], suggesting that perception is intrinsically multimodal. Multimodality can be examined in terms of cross-modal effects, wherein information from one modality is modified by the influence of information in another. The McGurk effect typifies this; a unimodal auditory stimulus of repeated spoken syllables: " ba, ba, ba" is perceived accurately, but the addition of incongruent visual stimuli – a video of a speaker mouthing: "ga, ga, ga" produces reports of hearing " da, da, da" [21]. The interaction produces an intermediate perceptual conclusion that minimises incongruity, indicating that congruity is an important underlying principle in multimodal processing. For a discussion of cross-modal effects, [22].

Additive effects occur when the perceptual conclusions via a single modality might be weak, but the addition of other-mode information provides more robust perceptual conclusions – this might be typified in the problem of discerning the contents of speech in a noisy environment; if we can see the talker, we can *hear* the talker better [23]. Sometimes *supra-additive* effects can occur, whereby each unimodal stimulus is weak, but in combination add up to more than the simple sum of the two or more stimulus components [24].

In theory, at least, the addition of tactile stimulus elements to auditory stimulus material can be *sub-additive* (a multisensory neuron might show a given response to input from a single sense-mode input, and greater response for signals from two or more modes, but not a proportionate summing of unimodal response), additive, or even super additive. Another way to consider this is to say that, in conditions where the unimodal signals capture sufficient complexity of information, multimodal input provides less additional information than where unimodal signals are comparatively weak. So, in conditions of unimodal impairment (hearing deficits or high noise-floor, or combinations of both), multimodal additions can provide more robust perceptual conclusions.

In the principle of psychophysical complementarity [25], we might assume that, since events in the world often or usually stimulate multiple sense-modes, integration of disparate information-streams should proceed according to evolved heuristics. That is, some form of "multimodal scene analysis" might be postulated, and this will take information from prior experience as well as on-going sensation; we are always contextually aware.

The key features of such multimodal scene analysis appear to be congruence and concomitance; synchronicity and spatial co-location would promote perception of concomitance, repetition of synchronicity (of stimuli) would reinforce it. Congruence might have higher-order implications; if an item in the environment visibly moves as though it has considerable inertial mass, and makes sounds with considerable (in amplitude) low frequency content, these would be congruent with the impression that the item is massive. Congruence-comprehension, the major component of *cognitive causal mapping* [26] must, in order to function in timely fashion, feature some efficient and rapid heuristics. Wilson and Sperber [27] in discussing Relevance Theory propose heuristics for comprehension:

a. Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.

b. Stop when your expectations of relevance are satisfied.

Although they are thinking specifically of language communication and comprehension, in theory a similar parsimonious approach might be appropriate for comprehension of other types of signal in the environment. Their use of "…interpretive hypotheses" and "…expectations" suggests that higher-order comprehension of context probably influences more peripheral neuronal responses, and by extension this could explain differences in multimodal processing. If unimodal information is incomplete, information from other modes can be recruited, whereas if the former is sufficient, the latter is redundant.

4. Compositional Use of Congruent Auditory and Tactile Spatial Signals

Music is inherently a multi-sensory experience; we *hear* in space, we may *see* or associate in our imagination a causal source, we may *feel* objective resonance and create internal narratives to support emotional constructs, guided by expressed extra-musical identities. Given this, how should music be defined and what would be the benefit of such a definition? Perhaps to offer illuminating insight and provide meaningful constraints for creative focus? A typical dictionary definition offers:

> *"the science or art of ordering tones or sounds in succession, in combination, and in temporal relationships to produce a composition having unity and continuity"*

(*Webster's Collegiate Dictionary*, online edition), which presents the obvious dominant characteristic and compositional preoccupation as naturally *sound*; regular patterns in which ideally result in listener coherence and sustained interest.

In a conventional sense, we might reasonably define music, in terms of this sonic attribute alone, as simply *'organised sound'*; the definition although seemingly superficial is satisfying since it is open and all-embracing of musical space, as was intended when Edgard Varése similarly expressed it [28] when discussing his own aesthetic sensibilities in relation to his recent excursion into multi-speaker tape composition: Poème électronique (1957-58). What is composition? A sonorous creative act, idea, performance or recording that might be considered new and valuable. To achieve value, this could involve be a transformation in an existing stylistic domain or the establishment of a new one that achieves cultural recognition. Alternative organisational designs in music we will call *style*, to mean the accepted *norms* of a musical period or individual. *Style* then in music refers to the common attributes and behaviours within a musical form; in any given *style*, certain features are considered normal and others anomalous.

All sound may be considered *musical*, which is the virtue of the above definition, but in each culture musicians tend to admit only a subset of acceptable sounds, frequency arrangements, combinations and temporal patterns, into sonic expression. Given the self-imposed limitations, music remains a system of sufficient complexity to allow for combinatory and sequential variation accommodating novelty, identity and meaning. Innovation in music however, requires more than mere *novelty*; the newness must have a context for it to be validated by the domain gatekeepers, as Frank Zappa said [29]: "Without deviation (from the norm), 'progress' is not possible…In order for one to *deviate successfully*, one has to have at least a passing acquaintance with whatever *norm* one expects to deviate from".

There may be no single intercultural definition of music and the boundary between musical sounds and noise may be culturally blurred. Varése speculated [28] upon the future of music *"the score of the future would need to be seismographic in order to illustrate their full potential"*, citing the definition of music given by Józef Maria Hoene-Wroński: "the corporealization of the intelligence that is in sound", as being particularly influential in shaping his musical imagination. Music could dispassionately be regarded as an abstract sonic temporal construction, constrained by pre-formed elements organised in predefined relationships; a product perhaps more of *discovery* than invention, that might conceivably be determined or computed; the permutations and combinations of acceptable outcomes calculated and selected according to stochastic design. From this perspective, mechanised musical culturally verified artefacts might be fabricated according to audience requirements for expectation, consistency, coherence and originality.

Educationally it is not uncommon to study the *craft* of composition by learning the characteristics of archetypal work, through systematic analysis codifying behaviours tested through re-creation; creative motivation and method is somewhat less often addressed and there are other important peripheral attributes of musical expression and experience *missing* from the analysis, that may offer new perspectives and valuable insights such as:

- The creative process: which is very likely non-linear; does music have to be experienced along a fixed timeline?

- The communication and expression of emotive design through dynamic physiological gestures in performance; music has at times been considered a language with linguistic syntactical structure [30]. The imprecision within the symbolic representation (notation) is also profitable for performers, allowing for expressive individual interpretation.

- The tactile sensations of performing/composing upon instrument, sensing and responding to the resonant vibrations within a space. Each performer has individual muscle memories and patterns of behaviour that may be meaningfully codified outside of sound.

If we could transform and translate our perspective, music might be qualified in another way; Varése experienced such an epiphany [31] when listening to a Beethoven symphony: *"I became conscious of an entirely new effect produced by this familiar music. I seemed to feel the music detaching itself and projecting itself in space. I became conscious of a third dimension in the music".* Could music then not be considered and appreciated, as Clifton [32] suggests, in terms of its inherent expressive qualities alone, conveyed as ordered events within which meaning and significance might be inferred? The materials of musical meaning traditionally involving sonic patterns in time becomes no less of a musical expression, for the indoctrinated, when conveyed as mute symbolic notation indicating the potential for sonic experience or opportunity for realisation in the musical imagination. For a performer of music, a significant tactile dimension is also present; the vibration of the instrument against the body as the sound is activated and muscular tensions are important extensions of expressive intent that are ordinarily absent in musical reception. Within this paper, we are considering the extension and translation of compositional patterns into a tactile domain, to explore the experience of composition and challenge the tradition of *sound* being the only medium within which musical intelligence might be conveyed and appreciated. In application, musical designs realised through tissue conduction alone offer constrained frequency and dynamic space previously inaccessible to the composer; a new enhanced complementary dimension, in addition to airborne vibrations, within which to convey a hitherto missing dynamism in musical design becomes a possibility. There are obvious implications for the hearing impaired, in the development of an application to provide a unique spatial insight into what was considered a formerly sonic experience and the potential for the composer to explore a *new musical domain* composing explicitly for a tissue conduction system.

4.1. Spatial Composition

Whilst it might not be aesthetically imperative to adhere to strict Euclidian representations, *spatial* parameters can be useful adjuncts to *musical* parameters, since they introduce dimensions of blending/separability, ensemble depth [33] and motion. For a discussion see [34].

In the case of blending/separability, it may be that, at times the desired percept is of a harmonious whole, even if formed by distinct elements; much classical orchestration is of this nature. On other occasions, the composer may be desirous of the ability for individual elements to perceptually break away from the whole, becoming individually distinct. Utilising this dimension can produce effects of swarming, dispersing and coalescing musical elements. Manipulation of perceived spatial location, including position-within-ensemble-depth (where some sources are apparently nearer to the perceiver than are others) can be effective in controlling blended/separate percepts. Perceived motion can include literal trajectories through three dimensions, with concomitant parameters of speed, change of loudness, Doppler effects, and changes in simulated early reflections. A particularly perceptually impressive parameter is that of auditory looming [35], humans are intrinsically inclined to pay increased attention to items in the environment that are approaching, and this is exacerbated if the items are fast, unfamiliar or even unpleasant-sounding [36]. Musically, this could be used to surprise, generate anticipation or divert attention.

The more interesting compositional domain lies perhaps in the metaphorical mappings between space and musical parameters such as loudness, frequency, tempo, movement and acceleration. Here, it is relatively unexplored what might be perceived as congruent; temporal congruence (items that start together and modulate together are probably part of the same causal stream, or ‘event’) and spatial congruence (items that spatially coincide are probably related) seem reasonable first principles. More complex principles might stem from observation of relationships between items in environments; items in environments affect each other. Antiphony, or “Call-and-response” exemplifies the notion of musical elements interacting, giving the impression of communication; this can be very effective when combined with spatial separation of instruments.

Imposing perceptual constraints to essentially explore a tactile experience with limited frequency and dynamic responsiveness would offer a composer unique insight into the propagation of musical and spatial information for the hearing impaired, potentially revealing a comparatively novel experience by conveying intimate resonances within performance expression. What a tissue-conduction compositional system might look like and to what extent such a system could be artistically significant is under consideration.

The system as a whole would not necessarily need to be confined to tissue conduction alone; a supplementary low frequency component would be a meaningful addition (for low frequency body vibration equivalent to video gaming chairs see the Marvel Avengers Vybe Haptic system). For those able to perceive airborne vibration, the system could potentially offer a sensory duality within which an experiential mirror or dialogue could be established between an ‘inner’ conductive system and an ‘outer’ sonic array. The sub-frequency content might be reasonably mirrored in both the sonic and vibrotactile domain. A system that offered the potential for the composer to explore, and audience to experience, such an enhanced expressive domain, including outer surround airborne content with subsonic material along with inner tactile detail with spatial information, would indeed be a novel one with multiple applications.

4.2. Spatial Qualia for Different Displays.

Whilst potentially exciting, there are some limitations that are intrinsic to the display methods for music; the range of perceptual significance [37] is constrained. For instance, loudspeaker presentations cannot simulate distances from horizon to the peripersonal (within-reach) space of the listener; on the other hand, headphone presentations can produce ‘in the head’ soundstage at the expense of externalisation and precision of ensemble depth.

Another, sometimes-overlooked deficit refers to vibration. In real environments, many events are accompanied by very low frequency sound and ground or structure-borne vibrations and these can be used to detect the approach of some items before their audible output reaches supraliminal proportions. Indeed, snakes rely on this type of information exclusively, as they lack air-conduction hearing apparatus [38]. There is evidence that the saccule in the vestibular system is sensitive to low frequencies [39] and can be recruited to a compensatory role in deaf… [40]. It may even contribute to enjoyment in traditionally-loud music genres [41, 42]. As the saccule is involved in the detection of self-movement in the vertical plane, it could be that "being moved by the music" is more than a figure of speech.

It could be that finer control of stimuli that engage the saccule can grant us finer control of the 'in-head/externalised' axis of experience, since the saccule is normally stimulated by a person's own speech, as Trivelli *et al* note [40]. It might also be used to govern the perception of size of objects independently of range; very large events (such as distant thunder) may be only marginally louder than the local noise-floor, yet can rattle the windows.

So whilst it is a convention to refer to "sound" as those airborne vibrations with frequencies of between 20Hz and 20 kHz, there is potential information available to perception below the lower frequency. In artificial presentations using loudspeakers, for practical purposes, the frequency range of 20Hz to 40 Hz is problematic as it is expensive to convey, similarly expensive to arrest, and tends to comprise noise to neighbours. Extending signal content down to 1 Hz would exacerbate such problems. However, using direct injection techniques such as vibrotactile transducers might allow us to explore the information-terrain that is normally inaccessible to composers and listeners.

An important caveat is that (at this stage of our research) we are largely considering passive tactile sensation, as distinct from active tactile sensation. For a discussion of the distinction, see [43]. That is, we are limited to stimulation of the somatosensory system in respect of tactile sensation through the activity of vibrotactile transducers. Tactile discrimination of frequency can be learned [44] and spatial location of stimulus can be finely identified (we know where on our body a stimulus is occurring). However, interactions between temporal and spatial characteristics of tactile perception require further study.

Interoception of vibrotactile stimuli is even less well understood; anecdotal evidence suggests that people can feel high amplitude low frequency sound in various parts of the body, varying with frequency. This could be caused by different internal structures of the body having different resonant frequencies, in which case internal spatial perception would be governed by frequency (rather than the actual position of actuators).

Given that hearing impairment clearly does not entail insensitivity to vibration, and may indeed be associated with increased discrimination of vibration [45-47], multimodal composition for hearing-impaired is theoretically feasible; experiences thereof may qualitatively differ from the normal experiencing of music.

5. Conclusions

Originally, we had embarked on an exploration to facilitate auditory spatial perception through tissue-conducted sound. Incidental input to the saccula notwithstanding, we thought of this as essentially an equivalent means to stimulate auditory perception. The reactions of one hundred listeners in our unstructured elicitation listening tests lead us to suspect some qualitative differences; we now think that the vibrotactile input amounts to something distinct from, and possibly additive to, audition.

The rules of congruence of concomitant multimodal stimuli await explication. Largely, this is because the relative contributions of low frequency vibrations (apprehended through the saccula and through the resonances in the body's internal structure) to ostensible auditory perceptions are not thoroughly understood. We suppose that some model of "multimodal scene analysis" (equivalent to Bregman's auditory scene analysis) could be developed. This would include heuristics that describe how inputs are bound together in perception on the basis of synchronous onsets, offsets and modulations, or separated due to asynchronicity along certain dimensions. It should also account for cross-modal effects, where the conclusions in one modality are modified by input in another and especially where particular signal-features in one modality are additive (in perception) to concomitant features in another (such as the exaggeration of transients)

Composing for amodal spatial perception, and utilising cross-modal effects via multimodal inputs is at a comparatively immature stage; parameters of the target qualia await elucidation.

6. References

1. Stenfelt, S., Goode, R.L.: Transmission properties of bone conducted sound: Measurements in cadaver heads. The Journal of the Acoustical Society of America. 118: pp. 2373--2391. DOI: 10.1121/1.2005847 (2005)
2. Sohmer, H.: Reflections on the role of a traveling wave along the basilar membrane in view of clinical and experimental findings. European Archives of Oto-Rhino-Laryngology. 272: 531. DOI: 10.1007/s00405-014-3045-z (2015)
3. Jun, K.C.: Objective measurements of skull vibration during bone conduction audiometry [thesis]. Medical Faculty: Zürich University; (2009)
4. Lenhardt, M.L., Shulman, A., Goldstein, B.A.: Bone-Conduction Propagation in the Human Body: Implications for High-Frequency Therapy. International Tinnitus Journal. 13, 2: 81–86 (2007)
5. Dietz, A.J., May, B.S., Knaus, D.A., Greeley, H.P.: Hearing Protection for Bone-Conducted Sound. In: New Directions for Improving Audio Effectiveness; Meeting Proceedings RTO-MP-HFM-123, Paper 14. Neuilly-sur-Seine, France: RTO. 2005; pp. 1--18 (2005)
6. Chordekar, S., Kriksunov, L., Kishon-Rabin, L., Adelman, C., Sohmer, H.: Mutual cancellation between tones presented by air conduction, by bone conduction and by non-osseous (soft tissue) bone conduction. In: Hearing Research, 283, 1–2; pp. 180--184. DOI: 10.1016/j.heares.2011.10.004. (2012)
7. Stanley, R., Walker, B.N.: Lateralization of sounds using bone-conduction headsets. In: Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society, 16-20 October 2006; San Francisco, CA. p. 1571-1575 (2006)
8. MacDonald, J.A., Henry, P.P., Letowski, T.R.: Spatial audio through a bone conduction interface. International Journal of Audiology. 45: pp. 595--599. (2006)
9. Stenfelt, S., Zeitooni, M.: Binaural hearing ability with mastoid applied bilateral bone conduction stimulation in normal hearing subjects. The Journal of Acoustical Society of America.134; 1: pp. 481--493. DOI: 10.1121/1.4807637 (2013)
10. Berg, J., Rumsey, F.: Spatial Attribute Identification and Scaling by Repertory Grid Technique and Other Methods. In: Proceedings of the Audio Engineering Society International Conference on Spatial Sound Reproduction (AES16); 10-12 April 1999; Rovaniemi, Finland. (1999)
11. Zacharov, N., Koivuniemi, K.: Audio descriptive analysis and mapping of spatial sound displays. In: Proceedings of the 7th International Conference on Auditory Display. 29 July–1st August 2001; Helsinki, Finland. pp. 95--104. (2001)
12. Bregman, AS. S.: Auditory scene analysis. MIT Press: Cambridge, MA; (1990)
13. Rumsey, F., Mason, R.: Interaural time difference fluctuations: their measurement, subjective perceptual effect, and application in sound reproduction. In: Proceedings of the Audio Engineering Society Conference: 19th International Conference: Surround Sound-Techniques, Technology, and Perception (AES19); 21-24 Jun. 2001; Schloss Elmau, Germany (2001)
14. Gardner, E., Martin, J.: Coding of Sensory Information. Principles of Neural Science 2000; chapter 21, 411--429. McGraw-Hill, fourth edition. (2000)
15. Tozzi, A., Peters, J.F.: A Symmetric Approach Elucidates Multisensory Information Integration. Information. 8(1):4. doi:10.3390/info8010004 (2017)
16. Klemen, J., Chambers, C.D.: Current perspectives and methods in studying neural mechanisms of multisensory interactions. Neuroscience Biobehaviour Review 36, pp. 111--133. doi.org/10.1016/j.neubiorev.2011.04.015 (2012)
17. Schnupp, J.W., Dawe, K.L., Pollack G.L.: The detection of multisensory stimuli in an orthogonal sensory space. Exp Brain Res. 2005 Apr;162(2):181-90. DOI:10.1007/s00221-004-2136-2 (2005)
18. Macpherson, F.: Taxonomising the senses. Philos Stud 153:123–142 DOI 10.1007/s11098-010-9643-8 (2011)
19. Fodor J.A.: Modularity of mind: an essay on faculty psychology. MIT Press, Cambridge, Mass (1983)
20. Cappe, C., Thut, G., Romei, V., Murray, M.M.: Auditory–Visual Multisensory Interactions in Humans: Timing, Topography, Directionality, and Sources. Journal of Neuroscience 22 September 2010, 30 (38) 12572-12580; DOI: https://doi.org/10.1523/JNEUROSCI.1099-10.2010 (2010)
21. McGurk, H., MacDonald, J.: Hearing lips and seeing voices. Nature. 264(5588):746-8 (1976)
22. Landry, S.P., Guillemot, J.P., Champoux, F.: Audiotactile interaction can change over time in cochlear implant users. Front. Hum. Neurosci. https://doi.org/10.3389/fnhum.2014.00316 (2014)
23. Sumby, W.H., Pollack, I.: Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. 26:212–215. (1954)
24. Werner, S., Noppeney. U.: Superadditive Responses in Superior Temporal Sulcus Predict Audiovisual Benefits in Object Categorization. Cerebral Cortex. doi:10.1093/cercor/bhp248 (2009)

25. Shepard, R.N. Psychophysical complementarity. In: Kubovy, M., Pomerantz, J.R. (eds.), Perceptual Organisation. Erlbaum, Hillsdale, N.J. pp.279-341 (1981)
26. Lennox, P.P.: Cognitive Maps in Spatial Sound. In: Proc. Audio Engineering Society's 52nd International Conference: Sound Field Control- Engineering and Perception. AES (2013)
27. Wilson, D., Sperber, D.: Relevance Theory. In: Horn, L., Ward, G. (eds.) The Handbook of Pragmatics. Wiley-Blackwell. Oxford (2005)
28. Goldman, R.F.: "Varèse: Ionisation; Density 21.5; Intégrales; Octandre; Hyperprism; Poème Electronique. Instrumentalists, cond. Robert Craft. Columbia MS 6146 (stereo)" (in Reviews of Records). Musical Quarterly 47, no. 1. (January): pp. 133--34. (1961)
29. Zappa, F.: Occhiogrosso P. The Real Frank Zappa Book. London: Picador. (1989)
30. Bernstein, L.: The Unanswered Question: Six Talks at Harvard (Charles Eliot Norton Lectures). Cambridge, USA. Harvard University Press. (1990)
31. Music and The Drama, Varese, Ultra-Modernist Composer, Prophesies Symphonies in "Space", The Lewiston Daily Sun, December 8, (1936)
32. Clifton, T.: Music as Heard: A Study in Applied Phenomenology. New Haven and London: Yale University Press. (1983)
33. Rumsey, F.: Spatial quality evaluation for reproduced sound: Terminology, meaning and a scene-based paradigm. Journal of the Audio Engineering Society. 50(9):651-666 (2002)
34. Lennox, P.P.: Music as artificial environment: spatial, embodied multimodal experience. In: Wöllner, C. (ed.) Body, Sound and Space in Music and Beyond: Multimodal Explorations. Routledge, Oxford. (2016)
35. Seifritz, E., Neuhoff, J.G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., Elefante, R., Di Salle, F.: Neural processing of auditory looming in the human brain. Current Biology, 23, 2147–2151. (2002)
36. Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., Västfjäll, D.: Embodied auditory perception: the emotional impact of approaching and receding sound sources. Emotion. Apr;10(2):216-29. doi: 10.1037/a0018422 (2010)
37. Lennox, P.P., Myatt, A., Vaughan, J.M.: From surround to true 3-d. In: Proc. Audio Engineering Society's 16th International Conference: Spatial Sound Reproduction. AES. (1999)
38. Knight, K.: Snakes hear through skull vibration. Journal of Experimental Biology. 215:ii doi: 10.1242/jeb.069104 (2012)
39. Emami, S.F., Pourbakht, A., Daneshi, A., Sheykholeslami, K., Emamjome, H., and Kamali, M.: Sound Sensitivity of the Saccule for Low Frequencies in Healthy Adults. ISRN Otolaryngology Volume 2013, Article ID 429680, http://dx.doi.org/10.1155/2013/429680 (2013)
40. Trivelli, M., Potena, M., Frari, V., Petitti, T., Deidda, V., Salvinelli, F.: Compensatory role of saccule in deaf children and adults: novel hypotheses. Medical Hypotheses, vol. 80, no. 1, pp. 43--46 (2013)
41. Todd, N.: Evidence for a behavioral significance of saccular acoustic sensitivity in humans. Journal of the Acoustical Society of America, vol. 110, no. 1, pp. 380–390 (2001)
42. Todd, N.P.M., Cody, F.W.: Vestibular responses to loud dance music: a physiological basis of the rock and roll threshold? Journal of the Acoustical Society of America, vol. 107, no. 1, pp. 496--500 (2000)
43. Gibson, J.J.: Observations on active touch. Psychological review 69(6):477-491. http://dx.doi.org/10.1037/h0046962 (1962)
44. Imai, T., Kamping, S., Breitenstein, C., Pantev, C., Lutkenhoner, B., Knecht, S.: Learning of Tactile Frequency Discrimination in Humans. Human Brain Mapping 18(4): pp. 260--71 (2003)
45. Shibata D.K, Zhong J.: Tactile vibrations are heard in auditory cortex in the deaf: study with FMRI. Annual Meeting of the Radiological Society of North America; Chicago, Ill;2001; p.259
46. Levanen, S., Hamdorf, D.: Feeling vibrations: enhanced tactile sensitivity in congenitally deaf adults. Neuroscience Letters 301, 75 - 77 (2001)
47. Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., Hari, R.: Touch activates human auditory cortex. NeuroImage Volume 30, Issue 4, 1 May 2006, Pages 1325–1331 (2006) http://doi.org/10.1016/j.neuroimage.2005.11.020 (2006)