Spatial sound via cranial tissue conduction

PETER LENNOX[1] AND IAN McKENZIE[2]

[1] College of Arts, University of Derby, Derby, UK
e-mail: p.lennox@derby.ac.uk
[2] College of Engineering and Technology, University of Derby, Derby, UK
e-mail: I.McKenzie1@derby.ac.uk

September 23rd 2016

Abstract

This discussion paper and workshop describes progress and challenges in developing devices to satisfy human auditory spatial perception using tissue conduction technologies. The technique bypasses the ossicular chain, conducting signals through cranial tissue to the cochlea. The ear canals can remain unobstructed, for superposition of air conducted and tissue spatial sound. Alternatively, for users with conductive hearing loss, the technique provides a useful adjunct to, and potentially a replacement for, conventional hearing aid technology. The novel incorporation of spatial elements, via techniques such as ambisonics encoding can bring a more satisfying and informative experience.

1 Introduction

The potential for utilising bone conducted sound as a substitute for normal air conducted sound has been known for several hundred years [1], though the underlying biophysics has, until very recently, been poorly elucidated (for a discussion see: [2].

An important aspect of normal air conducted hearing is that useful spatial information is available. Schnupp and Carr [3] observe that strong evolutionary selective pressures appear to have given rise to spatial hearing in all phylogenetically-advanced species and that current generations of hearing assistive technologies do not adequately supply users with spatial information. There is therefore considerable scope for investigation and development of spatial competencies in bone conduction apparatus.

Here we discuss the current state of progress in terms of extant technologies, investigatory methods, applications and perceptual theory. We then briefly describe the necessary next steps in the investigation.

2 Extant technologies

These do not satisfy the requirements for coherent control of complex spatial sound signal sets for cranial tissue conduction of audio signals

Whilst the principle of bone conduction of audible signals is well established [4][5][6], we have adopted the slightly different terminology of "cranial tissue conduction". This term acknowledges the fact that the exact transmission pathways culminating in the cochlea are not unambiguously established; the transmission pathway includes soft tissue (between transducer and cranium), cerebrospinal fluid [7], and brain tissue. Additionally, we have found efficacious transmission via eyeballs, the neck, and jaw.

The transmission path issues are further complicated by the observation that certain transducer sitings can produce anomalous spatial impressions, so that a transducer applied to the right side of the jawbone can produce the perception of lateralisation to the left. Likewise, for some subjects, particular sitings can produce similar laterally inverted imagery. What is not clear at this stage is the degree of between-subject consistency in these inversions.

Although several research teams have reasonably established that lateralised imagery via tissue conduction (using transducer sites on the mastoids and condyle) is almost equivalent to that for binaurally-presented signals [4][5][8], the progression to full three-dimensional auditory perception is not straightforward.

Existing air-conduction spatial sound technologies are predicated on precise control of the signals arriving at the binaural hearing system – specifically, at the entrance to the ear canals. Signal sets can be spatially encoded utilising interaural amplitude differences, interaural temporal differences and pinnae-filtering. Additionally, interaural cross correlation (IACC) can be

managed in order to govern perceptual impressions of *spaciousness* [9] and *externalisation* [10].
The transmission path for airborne sound, from ear canal to cochlea is simple compared to the tissue conduction case which involves multiple paths (even for the case of a single transducer) and therefore different transmission speeds; once past the initial impedance of the outer skin sound waves propagate through several pathways in the human skull, the composition and therefore the impedance of these pathways is considerably different, propagation speeds can range from 250m/s to 400m/s. [11] resulting in a very complex signal set arriving at the basilar membrane. Precision control of inter-cochlear temporal differences (using multiple transducers) is confounded by this multiplicity. Similarly, given the observation of laterally inverted imagery (which we theorise may be partly due to the acoustic properties of the interior of the cranium), control of inter-cochlear amplitude differences is compromised.

Coherent control of the spatial impression of elevation (normally evoked by pinna-filtering) is hindered in that there is no extant cranial-tissue transfer function data, and hence no means of translating pinna-encoding to cranial conduction signal sets.
Since there is no established theoretical basis for adapting existing (air conduction) spatial sound encoding to the cranial-tissue case, investigation of the possible range of evoked spatial impression via the latter is through pragmatic and empirical means. An arbitrarily selected range of techniques have been utilised to manage inter-transducer signal relationships; these include pair-wise and triplet-wise amplitude panning, pairwise short (0-1ms) and long (1-15ms) delay panning, and first order ambisonics (for naturally recorded spatial material and for artificial spatial reverb). The transducer locations are as shown in figure 1 (and can be seen in the workshop presentation) and notably, a transducer is positioned on the top of the cranium for experimentation with impressions of elevation. Clearly this latter does not produce physical equivalence with pinnae encoding, but the initial question is whether any impression of 'up' is evinced at all.

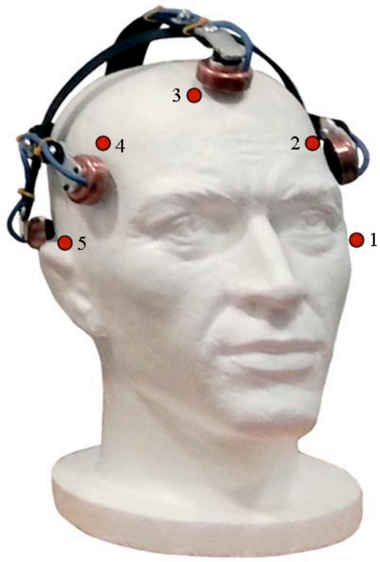

Fig 1.
Sound presented at:

- 1–Left Mastoid
  2–1" above Left Temple
  3 – Point between Forehead and Vertex
  4 – 1" above Right Temple
- 5- Right Mastoid

3 Investigative Methodologies

Although some methods exist for comparing tissue conduction (TC) with air conduction (AC), they are not predicated on examining and comparing spatial performance

Whilst Stanley & Walker [4]and MacDonald & Letowski [5] found equivalent performance between cranial tissue conduction and binaural air conduction in lateralisation experiments (where lateralisation is used as a proxy for direction perception), there are difficulties adapting classical psychophysical methods for investigations into externalisation, range-perception, spaciousness and elevation perception. In binaural experiments, individualised or generic head-related transfer functions (HRTFs) can be utilised whereas, for tissue-conduction experiments generic transfer function datasets do not exist and individualised transfer functions would necessarily measure between each specific transducer and the signal at the auditory nerve (i.e. after the cochlea), which would entail invasive procedures. Therefore, intra-subject comparisons may not reliably be produced.
The matter is further complicated in our observation that there is a significant learning period, during which performance on tasks such as direction discrimination, externalisation and spaciousness perception improve with continued and repeated exposure to the

experimental apparatus. The inter-subject variation in this learning period is very large, ranging from several hours to several hundred hours; whilst an optimal standardised learning period might be desirable for psychophysical investigations, it is too early to conjecture what that might be.

New methods for investigating 'active auditory perception'.

The search behaviour of subjects during the learning phase, whilst using the TC technology, indicates that conventional psychophysical methodologies can fail to capture important data.

We have often observed, especially in the early stages of exposure to the cranial tissue conduction apparatus, distinctive head movements in the tilt, tumble and rotate axes as well as head repositioning facilitated by torso movements. Naturally, without head tracking and corresponding adjustments to transducer signals, such 'search behaviour' elicits no useful information. Nevertheless, such perceptually-motivated behaviour indicates that, especially on receipt of incongruous or unfamiliar circumstances, listeners do not rely solely on analysing passively-received signals. Theories of embodied cognition [12][13] emphasise the intrinsic role of behaviour in refining perceptual conclusions. We hypothesise that, with appropriate head tracking experiments, this search behaviour would interact with the training period, significantly shortening exposure time required to reach performance asymptote.

4 Potential applications

Standard audiological tests do not finely characterise the contributory proportions of sensorineural and conductive deficits in age-related or other auditory deficits in the general population, especially in respects of spatial hearing.

Audiological testing does utilise comparisons between bone conduction and air conduction to elicit information on proportions of conductive and sensorineural components in hearing deficits. Many initial inexpensive methods involve a tuning fork applied to bony areas such as the mastoid. More extensive (and therefore expensive) methods can be deployed to obtain more detailed information, but these are generally reserved for situations where potentially serious medical conditions are suspected.

In discussions with volunteers with diagnosed hearing deficits, we find that many have only vague (and often inaccurate) knowledge of their condition. Certainly, in respect of knowledge of deficits in *spatial* performance, we have found paucity. We conjecture that, with an apparatus comprising binaural signals plus spatial tissue conduction headset, it is feasible to produce a detailed characterisation, for wide frequency and amplitude ranges, of individuals' spatial hearing performance and the underlying mechanisms.

5 Potential revisions to current auditory spatial perceptual theory

Especially, elevation and externalisation perception, which are putatively products of pinna encoding (and so should not be amenable to manipulation for a tissue-conduction system which bypasses the pinnae) should be re-examined.

Some subjects report, after repeated listening, perceptual impressions of overhead sources. For natural recordings, this could be due to expectations in respect of the subject material (birds and thunderstorms might be expected to be 'up'). In the case of musical tones, this argument is less compelling, though there may be an interaction between frequency content and spatial conclusions (so high tones sound 'high'). To eliminate the possibility of accidental air conduction from transducer to outer ear, we occluded the ear canals with ear buds providing 25dB attenuation. A possible candidate contributing to the elevation perception is a multimodal cue: we know that, at high signal levels, the vibration of the transducer is felt; one subject reported the sensation as unpleasant. It may be that, at lower levels, the haptic component is still significant (though unreported). Alternatively, it could be that the audio quality for a transducer on top of the cranium differs from that for other transducer locations, and this can be learned as 'up'. Finally, it could be that all these factors contribute, to varying degrees, to a unitary perception of 'up'. We have used the term 'up' instead of the more usual 'elevation' because we have not yet attempted to evoke the perception of 'down', nor have we more finely investigated gradations in elevation perception.

Whilst cranial-tissue conduction of pure tones evokes 'in head' perception, naturalistic recordings or artificial spatial reverberation readily evokes reported impressions of externalisation and some degree of range perception. Interestingly, externalisation and range perception are poorest for the frontal segment of the field and currently we do not use a transducer sited on the nose or forehead (both excellent conductive sites).

We have not yet systematically investigated perceptions of sources' movements. Initial impressions are that, unsurprisingly, lateral movements are clearly perceptible, and in conjunction with externalisation measures, sources apparently pass overhead. Front-back and down-up movements are less clearly defined, though diagonal movements (that intrinsically feature lateralisation changes) are perceptible.

Overall, if we assume that we are not producing signal qualities that are directly equivalent to those utilised by the binaural air conduction system to arrive at spatial conclusions, then we need to develop an alternative theoretical framework for spatial hearing in cranial-tissue conduction.

6 Development of future research areas

The research areas for the future can broadly be categorised as technological and perceptual, and these interact.

Technological:

Transducer design and deployment.

Transducer design has not enjoyed a long evolutionary developmental process similar to that for acoustic transducer design. Transducers are simple piezoelectric motors covering the audible frequency range, though their frequency response is uneven and they become increasingly inefficient with decreasing frequency and have poor response below approximately 200Hz. Since multi-transducer spatial arrays have not been used, component matching has not been prioritised and performance varies between units.

As there is a physical impedance mismatch, for efficient energy transfer the transducers must be pressed firmly against the tissue, resulting in possible discomfort. Variations in contact force produces frequency dependent variations in effective energy transfer, which is problematic for multiple-transducer arrays.

It may be that an alternative approach entails different treatments for different frequencies, similar to two-or three-way loudspeakers, so that transmission can be optimised for narrower frequency bands

Signal processing:

Existing spatial sound encoding regimes are clearly not intended for direct-to-cochlea tissue conduction. The theoretical basis for control of spatial attributes such as imagery, movement, externalisation and spaciousness requires formulation. It may even be that some attributes are simply not amenable to coherent control. It may also be that processing for lateralisation and for front-back discrimination differs substantively; similarly for matters of elevation. This would imply that different processing strategies would be required for different spatial attributes.

Development of a dedicated signal-processing regime relies on perceptual testing, which in turn requires control of experimental variables that are inherent in current technologies.

Perceptual:

Classical behavioural psychoacoustic methods have a valuable role, as do explorations of cognitive and multimodal dimensions. The target is an understanding of the potential 'informational bandwidth' – how much information *could* be accessed by users of cranial-tissue conduction apparatus, and how could this be optimised? Of particular interest here is the question of individual differences in duration of learning periods. Clearly, the technique is not immediately intuitively accessible to all users. The prolonged training period implies that some processes akin to neural plasticity are at work; the signals produced at the cochlea are not physically equivalent to those for air conduction, but, inasmuch as they are coherent, they may be reliably utilisable. One promising line of research lies in investigating the strategies subjects use to maximise 'information pickup' [14]. Proactive 'search behaviour' exemplifies this and a direct investigation of such behaviour would rely on monitoring and categorising head and body movement strategies (see:[15]).

It may be that other sensory-modal information (e.g. haptic cues via vibrotactile stimuli) is contributing to the auditory perception in subjects during trials with TC. This multimodal perception should be the subject of an important line of inquiry. The intrinsically multimodal nature of everyday perception can easily be overlooked in the laboratory or in our deployment of technology-facilitated perception. But as our goal is to optimise efficiency of information transactions, proactive and multimodal theories of perception cannot be ignored.

7 Conclusions

As average life expectancy increases, so does the probability that many of us will spend a significant proportion of our lives with depleted hearing [16]. The simple pleasures of listening to music or watching the television are lost, and there is evidence that hearing impairment may contribute to accelerated cognitive decline [17]. A significant component of age-related hearing deficit is conductive hearing loss-which may also be a causal factor in sensorineural and auditory-cognitive impairment. Cranial tissue conduction of auditory signals has been known of since the 16^{th} century; the application of modern technologies and signal-processing techniques could provide significant benefits to society.

References

[1] A. Mudry & A. Tjellström, "Historical background of bone conduction hearing devices and bone conduction hearing aids" *Advances in Oto-rhino-laryngology*. 2011;71:1-9

[2] T. Tchumatchenko & T. Reichenbach, "A

cochlear-bone wave can yield a hearing sensation as well as otoacoustic emission" *Nature Communications*. 5:4160 (2014)

[3] J.W.H. Schnupp & C.E. Carr, "On hearing with more than one ear: lessons from evolution" *Nature Neuroscience* 12, 692 – 697 (2009)

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

[5] 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.

[6] A.J. Dietz, B.S. May, D.A. Knaus, H.P. Greeley, "Hearing Protection for Bone-Conducted Sound" In *New Directions for Improving Audio Effectiveness* (pp. 14-1 – 14-18). Meeting Proceedings RTO-MP-HFM-123. Neuilly-sur-Seine, France: RTO. (2005)

[7] H. Sohmer, S. Freeman, M. Geal-Dor, C. Adelman, I. Savion "Bone conduction experiments in humans – a fluid pathway from bone to ear" *Hearing Research*; 146:81-8 (2000)

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

[9] M. Conceição, & D. Furlong, "Influence of Different Test Room Environments on IACC as an Objective Measure of Spatial Impression or Spaciousness" In *Proceedings of the 131st Audio Engineering Society Convention* October 2011 New York

[10] R. Mason, & F. Rumsey, "An assessment of spatial performance of virtual home theatre algorithms by subjective and objective methods" In: *Proceedings of the 108th Audio Engineering Society Convention,* 2000-02-19 - 2000-02-22, Paris, France

[11] S. Stenfelt, "Acoustic and physiologic aspects of bone conduction hearing" *Advances in Oto-Rhino-Laryngology*, (71) (2011)

[12] A. Clark, *Being There: Putting Brain, Body and World Together Again*. Cambridge MA: The MIT Press (1997)

[13] A. Clark, "Embodied Prediction" In T. Metzinger & J. M. Windt Eds. *Open MIND:* 7(T). Frankfurt am Main: MIND Group (2015)

[14] J.J. Gibson, "The theory of affordances" *Perceiving, Acting, and Knowing: Toward an Ecological Psychology* 67-82. Hillsdale, NJ:Lawrence Erlbaum (1977)

[15] E. J. Gibson, "Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge" *Annual Review of Psychology* 39: 1-41 (1988)

[16] K.J. Cruickshanks, T.L. Wiley, T.S.Tweed, B.E. Klein, R. Klein , J.A. Mares-Perlman, D.M. Nondahl, "Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study" *American Journal of Epidemiology* 1998 Nov 1;148(9):879-86

[17] F.R. Lin, K. Yaffe, J. Xia, Q.L. Xue, T.B. Harris, E. Purchase-Helzner, S. Satterfield, H.N. Ayonayon, L. Ferrucci , E.M. Simonsick; Health ABC Study Group. "Hearing Loss and Cognitive Decline in Older Adults"*Journal of the American Medical Aassociation Internal Medicine*. 2013;173(4):293-299