LOW FREQUENCY SPATIALIZATION IN ELECTRO-ACOUSTIC MUSIC AND PERFORMANCE: COMPOSITION MEETS PERCEPTION

Roger T. Dean
austraLYSIS, Sydney, Australia
MARCS Institute, University of Western Sydney, Sydney, Australia
roger.dean@uws.edu.au

The article takes the perspectives of an electro-acoustic musician and an auditory psychologist to consider detection of localization and movement of low frequency sounds in reverberant performance environments. The considerable literature on low frequency localization perception in free field, non-reverberant environments is contrasted with the sparser work on reverberant spaces. A difference of opinion about reverberant environments has developed between on the one hand, audio engineers and many musicians (broadly believing that low frequency localization capacities are essentially negligible), and on the other, psychoacousticians (broadly believing those capacities are limited but significant). An exploratory auditory psychology experiment is presented which supports the view that detection of both localization and movement in low frequency sounds in ecological performance studio conditions is good. This supports the growing enthusiasm of electro-acoustic musicians for sound performance using several sub-woofers.

INTRODUCTION

Psychoacoustics now generally considers that there are three main mechanisms which allow humans to localize sounds, and detect movements of sonic sources (reviewed: [1-3]). The first two constitute the 'duplex' of inter-aural time differences (ITD) and inter-aural intensity differences (IID) distinguishing a single sound as heard at the two ears (see below for elaboration). The duplex theory was developed by Strutt (Lord Rayleigh) in the late 19th and early 20th century [4], following early observations on localizing the sound of a flute by Venturi in the 18th century. These two mechanisms are binaural: that is they require interaction between the auditory pathways flowing from both ears, and may exploit neural coincidence detection [5]. The third mechanism is monaural: the so-called spectral notch effect, whereby head-related transformations of the spectrum of an incoming sound, due mainly to the pinnae of the ears, provide location cues. All the mechanisms are influenced by head size and related anatomical features. The IID mechanism is largely restricted to frequencies above about 1000 Hz; the monaural mechanism operates in a fairly similar range. Only the ITD mechanism is usefully functional at lower frequency. Most psychoacoustic studies of sound localization have used free field conditions, that is conditions in which reflections and reverberation are virtually lacking, such as in an anechoic chamber or virtual environments in headphones. Those with primary involvement in auditory perception might note that this usage of the term free field is that of acoustics, and may not coincide with their normal usage.

None of these mechanisms of human sound localization are excellent when faced with low frequency sounds, and relatively few data specifically address frequencies below 250 Hz. Nevertheless, as I will describe, several fields of music exploit such very low frequencies substantially: for example music with or comprising of drones, noise music, and electroacoustic music at large. For this oeuvre, 250 Hz is hardly conceived as ‘low frequency’, rather more like mid-frequencies. Furthermore musical uses of frequencies below 250 Hz require presentation in reverberant environments, such as the performance studio, dance floor, or concert hall, and in moderately reverberant environments such as recording studios; conditions which contrast with those of the vast majority of psychoacoustic studies.

It is worth pointing out the distinctions between perceiving localization and lateralization. Biologically speaking, the important feature of a sound is the location of its source. However, in audio engineering and in electro-acoustic music in particular, more important is the degree of movement of the sounds and their components, expressed as their changing lateralization, that is their apparent positional spread in the listening space, whether virtual (headphones) or physical. One experimental approach to making this distinction is to have listeners represent the centre of a sound source on a one-dimensional scale from L to R [6], another to have them draw a graphical range of spread. Note the implication here that listeners may be perfectly aware of the disposition of the loudspeakers, yet perceive sounds as disposed almost anywhere in the listening space.

Thus the purpose of this article is to discuss low frequency perception in reverberant environments of ecological relevance to music, as an aspect of applied acoustics of particular importance to music composition, production and performance. I point to some of the gaps in the literature in relation to these applications, suggest and illustrate some useful analytical and experimental approaches, and also seek to provide pointers that may eventually be useful to electroacoustic composers and improvisers in constructing the spectral design of their music.

A SYNOPTIC REVIEW OF LOW FREQUENCY LATERALIZATION AND ITS IMPACTS

Sound spatialization in music performance and recording environments.

The sound of an orchestra is spatialized as a consequence of the disposition of the instruments in the performing space. For example, the low pitched instruments such as tuba (brass) and double bass (strings) are usually to the back, and at one or other side of the layout, while the high pitched flutes and oboes (woodwind) and violins (strings) are to the centre and further forwards, in relation to the position occupied by the conductor. Prior to the era of orchestral music, there were also notable compositional experiments with spatial dialogues between groups of instruments, as in the works of Gabrieli for brass chorale. Subsequent to the 19th/20th century dominance of the orchestra in western music, electroacoustic music since 1950 has enlarged this emphasis on spatialization to extremes, where hundreds of loudspeakers may be arrayed around a performing/listening space in a 3D organization, so that sounds can be projected from any point, or throughout, and can be 'moved' around [7, 8]. Composers have also sometimes had the opportunity to create grand architectural spaces (some temporary, some mobile) for such performances, for example Stockhausen and Xenakis [9], and in Australia, Worrall [10]. To electroacoustic music, even if presented in relatively humble stereo or quad, timbral flux around the listening space is a key feature.

Thus it is a matter of concern that we lack comprehensive data and mechanistic understanding of sound localization in some parts of the frequency spectrum: notably low frequencies. This is so much the case that particularly amongst studio and sound projection audio engineers there is a repeatedly argued view that low frequencies are poorly localized. The argument consequently suggests there is little point in having multiple subwoofers (the specialized speakers which most fully represent frequencies below about 200 Hz) in a performance space, unless the purpose is solely enhancing the loudness of those low frequencies. Furthermore, it is relevant that electroacoustic projection does not provide such detailed visual cues to sound lateralization as an orchestra provides, so the difficulty with low frequencies is not reduced by such cues. In neatly presenting the somewhat opposed views of psychoacousticians and audio engineers, a spatialization team from the University of Derby [11] present a concise review discussion. They conclude that on balance it is to be expected that in most environments low frequency localization is much better than the audio engineer community admits, yet concede there are major gaps in the data and our understanding in relation to ecologically relevant musical environments, which are mostly reverberant.

Electroacoustic composers have continued apace to exploit sound spatialization, relying mainly on their compositional intuitions and perceptual impressions as to what is or is not audible. Empirical evidence suggests that perception of low frequency lateralization may be masked to some degree by the presence of broad band energy in higher frequencies [12]. But other work emphasizes the positive impact of low frequency sound on auditory image size, and the additional benefit of stereo (or multiple) subwoofer conditions for perception of overall image size [13, 14]. Stereo subwoofers were thought to be distinct from mono pairs, with a limited number of popular music tunes, but the subjective preference experiments which followed were not statistically conclusive [15, 16]. Nevertheless, multispeaker systems, vector based amplitude panning and ambisonics have been used in attempts to enlarge the 3D impact of sound projections, and to foster impressions of envelopment [17] and engulfment, the latter seemingly related to the degree to which 3D impressions do superimpose on 2D [18, 19]. The multi-speaker system at ZKM Karlsruhe, neatly called Klangdom, has been developed alongside corresponding sound diffusion software, called Zirkonium [8]. This allows a composer or sound designer to specify the location of a sound, as a combination of outputs from three adjacent speakers, and allows a compositional dissociation between audio track and speaker. In contrast virtually all other software assumes that an individual audio track is sent to 1 or 2 speakers (according to whether the track is mono or stereo); some software, such as Ardour, MAXMSP and ProTools does permit a single audio stream to be sent to any combination of speakers, but they do not allow specification in terms of localization.

The Zirkonium system, and a few others can also well represent one of the most notable developments in spatializing electroacoustic music: what Canadian composer Robert Normandeau calls timbral spatialization [8]. This is effectively the systematic presentation of different frequency bands from different parts of the performance space, and their movement around the space. Convenient and accessible software, Kenaxis, based on MAXMSP, allows easy approaches to this technique even in live performance.

Such developments in electroacoustic music further emphasize the importance of understanding our localization, lateralization and movement detection abilities in relation to every band of the audible frequency spectrum.

The importance of low frequencies in electroacoustic and other music: relations to environmental and speech sounds

I mention here a few aspects of the increasing but long-standing musical importance of low frequency sound. Certain ritualistic musics, such as some of Tibet, and of the Australian didgeridoo, use repeated low frequency timbres, in a manner akin to the drone in Indian music (which is actually a wide band frequency pattern with strong bass), and to the specialized forms of shamanic [20] or trance-music which use drones. A drone in this music is a long-sustained low band sound, often changing very slowly or not at all.

The evolution of popular music through jazz and rock, via amplification has resulted since the 1960s in much higher sound levels in the bass instruments: as for example, the acoustic contrabass has often been heavily amplified or replaced by the electric bass or by bass motives played on a synthesizer. The bass 'riff', or accompanying repetitive rhythmic-melodic form of much music in rock and jazz has consequently been able to take a much more foregrounded position in the overall sound mass, a position which has been enhanced by mixing technologies. A similar but less obvious phenomenon took place in certain forms of contemporary classical instrumental and electroacoustic music and in dance electronica, where more visceral sensations [21] were created than generally sought before (with notable exceptions such as the medieval composer Marais). Thus it is interesting to compare Xenakis' influential orchestral piece *Metastaseis* [22] with, for example, an Elgar symphony in respect of the contribution of low frequencies to the overall acoustic intensity: there are sections in *Metastaseis* in which even the score shows clearly the dominant sweep of instrumental glissandi (pitch slides) in the lower string instruments, something rare in previous music. Similarly, Xenakis's electroacoustic piece *Bohor*, a work of considerable power, acoustically and affectively, has sections in which transitions in the low frequency intensities are the dominant feature, requiring high quality reproduction for full appreciation.

Electroacoustic music points to the possible relation between low frequency timbres in music, speech and environmental sound. Speech (and voice) is relevant because it is a common component of electroacoustic music, often heard both raw and digitally transformed [23-25]. One interesting aspect of this is that across speakers and conditions, the median frequency of the speech FO (the fundamental pitch) is around 130 and 220 Hz for men and women respectively [26, 27], well below the range of most psychoacoustic experiments on other sounds, and well into the 'low range' in any conception. Environmental sound is relevant partly because of the importance of overt environmental sounds in soundscape and other aspects of electroacoustic composition, but also because some of the physical associations of low frequency sounds in the environment (with large mass, low position, slow movement) can provide important metaphorical cues in the music.

Musical tension, often created by controlling the degree to which expectation is fulfilled [28], can then exploit countermanding cues, such as rapid movement of some sonic objects at the same time as a lack of movement of low frequency objects. Conversely, electroacoustic music, and noise music in particular, often focus on movement of low frequency sounds, again raising fundamental questions for the psychoacoustics of reverberant environments. Noise music is a large genre, springing from the work of Xenakis but extending from classical acousmatic composition to underground rock (and the edges of techno and drum 'n' bass). In the core of noise music, high intensity sounds of at most slowly varying complex timbres (high spectral flatness, poorly defined spectral centroid) are used. The timbres often start close to white (or sometimes pink) noise, and sculpt them slowly. Whereas in most previous music, melodic or more generally motivic structure has most often been delivered largely in the high frequency bands, this is no longer true in noise music, nor in many other aspects of electroacoustic music [29, 30].

All these observations on the current usage of low frequencies in music emphasise that understanding the perception of spatialization and movement of timbres comprising frequencies below 250 Hz is needed for composers to most fully and powerfully exploit it. Hence this is a worthwhile topic in applied acoustics, but as yet has not attracted the attention it deserves. I turn next to a brief summary of what is known about this, with particular reference to environments which are ecologically apt for music: in other words, reverberant rather than free field (non-reflective) environments.

Some relevant psychology and psychoacoustics: perception of low frequencies in reverberant environments

It is worthwhile to ask what psychoacousticians treat as 'low frequencies' and why; and to contrast that with even conventional compositional perspectives, let alone electroacoustic ones. There is a fairly clear lower limit to the frequencies in which the IID aspect of the duplex theory provides significant information: around 1000 Hz. It seems this may have driven the psychoacoustician (and perhaps acousticians and audio engineers) to treat frequencies below 1000 Hz as 'low', and hence rarely to venture below 250 Hz (see for example [31, 32]). In contrast, the frequency at the centre of a piano keyboard is the note called 'middle C', and it has a fundamental frequency of only about 260 Hz. This note also appears right in the middle of the two staves which notate two part tonal music. So for a classical composer, using acoustic instruments and notation to make 'pitch-based' music, low pitches are those at least an octave below middle C, in other words, below about 130 Hz; around the pitch referenced as the common male speech fundamental. Electroacoustic composers are often influenced by perceptions of, and maybe experience in playing acoustic instruments, so they share the conception of low pitch being around 100 Hz, even when they make 'sound-based' music, in which pitch may not be apparent and is certainly of limited importance [33]. One of the clearest and fullest studies in support of the idea that localization is possible at such low pitches [6], though dealing with headphones rather than reverberant spaces, indicates sensitivity with narrow band sounds down to 31.5 Hz, but with quite lengthy static stimuli (800 ms).

I provide some pointers to frame our further comments on perception of sonic movement, and perception of both location and movement in reverberant environments. The above-mentioned early and influential experiments of Lord Rayleigh 'on the lawn' (and sometimes with the participation of 'Lady Rayleigh') involved a tuning fork of low frequency (128 Hz), and did also involve speech. They lead to the duplex theory, and it was interesting that in the earliest papers (e.g. 1876) Rayleigh considered the tuning fork a 'pure' sound, but by 1907, when he reviewed [4] his overall work in this area of sound localization, he emphasized that it is a more complex sound, and brings to bear mechanisms at many frequencies. This timbre has been studied in detail subsequently, usually consisting of at least three harmonic components and several side-components, over a wide frequency range [34]. The duplex theory has largely survived empirical testing, as summarized in the two recent general review articles I reference [2, 3]. In depth discussion is provided in some of the empirical studies such as those of Middlebrooks and collaborators. They summarise the situation as follows : 'the duplex theory does serve as a useful description (if not a principled explanation for) the relative potency of ITD and ILD [which is what I term IID] cues in low- and high-frequency regimes (p. 2233) [35].

The main factors which create ITD and IID are clear: the geometry of the head in relation to the sound source. What is perhaps less obvious is why frequency should impact on the potency of the ITD and IID cues: but it seems that the diffusion of energy around the head and ears is such as to annul most low frequency IID cues, whereas the ITD remains. A given time delay represents a smaller proportion of the long cycle times of low frequencies, and this may provide more discrimination than with high frequencies. Monaural localization depends primarily on the influence of the pinnae of the outer ear on the sound transmission, creating transformed power spectra often with notches in mid to high frequencies, which can provide location cues that can be learnt even with only one ear (see [3, 5]). Neural pathways and possible mechanisms for localization have been investigated [5, 36] and computationally modelled [37].

All three localization mechanisms can be influenced by head movement, which consequently is an advantageous feature of listener behaviour [38], perhaps particularly for low frequencies [39]. Many experiments have restricted head movement, so as to control this influence. Of course, audio engineers, and music creators and listeners are used to taking advantage of it and hence experiments of ecological relevance to them do not restrict it (as in the exploratory experiment below). In favourable free-field conditions, the minimum audible angle (MAA) for localization is about 1° for broad band noise and the minimum audible movement angle (MAMA) is much higher [2]. Data suggest that for a 500 Hz tone MAMA is about 8° at velocity 90 degrees per second, and >20° at 360 degrees per second [40]. There is also evidence that a minimum stimulus duration between 150-300 ms may be required for motion detection [41].

There are also possible non-auditory components of low frequency localization, involving the vestibular system, or vibrotactile information, perhaps registered on the face, nose or other body parts [21, 42]. Of ecological relevance here is that many noise artists, used to very high sound intensities, have learnt to conserve their hearing by the use of ear plugs: for them, such non-auditory components of input sounds may be of even greater importance. The author, like many in electroacoustic music or electronic dance music, has experienced disconcertingly extreme intensity sound in some underground sound clubs, and used earplugs both there and when performing noise music. At some venues, earplugs are always given out at the entry. On the lower scale of acoustic intensity, orchestral wood wind players, usually seated just in front of the brass section, commonly have protective screens behind them in rehearsal and performance: unlike the screens used in recording studios, their purpose is simply reducing the sound impact on the woodwind players. The issue of earplugs and sonic localization is worthy of in depth experimental study.

Finally, I briefly summarize some of the known impacts of reverberation (i.e. rooms, or some partially sequestered spaces such as valleys). Many of the relevant studies use simulated reverberation, and have not been fully corroborated in experiments in real environments. The importance of reverberation diminishes when listeners are less than about 1 metre from the sound source, or when they are substantially eccentric to the speakers and/or the space. While speech perception is more difficult in reverberant environments, in some respects musical listening may be enhanced, and there is an important industry concerned with the design of environments and architectures for events, concerts, domestic and studio listening and living (e.g. [43]). Reverberation facilitates distance judgements concerning sound sources [44], but it generally diminishes other aspects of localization discrimination (reviewed [45, 46]). An important ongoing series of studies on localization of sound in rooms is being undertaken by Hartmann (e.g. [47, 48]). From this series, one salient observation is that the utility of the ITD declines below 800 Hz, though few data exist concerning frequencies below 200 Hz: sensitivity (in terms of ITD threshold time) is about twice as good at 800 as 200 Hz [49]. Some ITD-related theories of localization fail in this low frequency range [50]. Low frequency IIDs are even further reduced in reverberant environments, nevertheless, low frequencies may still be highly weighted in resultant localization judgements [46, 51].

Two illusions are particularly relevant to low frequencies and to reverberant environments. Precedence effects occur when two sound sources are separated in time, and determine whether the sounds undergo fusion: the opening sound may dominate the localization, preventing the realistic fission between the two. The duration of the sound is important in this, and in free field (particularly, anechoic) environments, the crucial transitions generally occur between 1 and 20 ms. However, this is much less clear for low frequency sound and for reverberant environments (see review [52]). The Franssen illusion is related: in this even a gradual transition of acoustic energy from one loudspeaker to another can be missed, and localization be determined by the opening onset location [2, 3, 53].

EXPERIMENT

An exploratory experiment on low frequency localization and movement detection in reverberant environments

Bearing the preceding discussion in mind, I conducted an exploratory experiment on this topic. It investigated whether and how fast a low frequency sound can be perceived to move in a performance studio environment (with quite short reverberation time). It complemented this with measures of location of sounds, using a four-alternative forced choice approach. I hypothesized that at frequencies below 200 Hz, localization to L or R of a listener would remain feasible in a reverberant environment, even for filtered noise. It was expected that accuracy would be higher as duration of the sound increased 0.2 s<2 s<6 s, the latter two durations chosen to correspond to feasible compositional durations for spatial movement. Similarly, I hypothesized successful detection of sonic motion at the latter two times, but to be poor or non-existent at 0.2 s. Note that the forced choice design does not aim to distinguish between localization and lateralization, as described above.

Our experimental procedure and studio environment is described more fully in the Appendix, and the legend to Figure 1, but the essence was as follows. I used two Genelec

7060B subwoofers as sound sources, each at 45° to, equidistant from, and at the same horizontal level as our seated listeners, who were musically trained people. There were five participants. Sounds were of three kinds, each low-pass filtered to reduce the presence of higher frequencies, and synthesized within MAXMSP software: sine tones of 30 and 60 Hz, and white noise. These were presented either for 0.2, 2 or 6 s, with 10ms on/off ramps, with the SPL (dBA) at the listeners' head set unequally on the basis of readily acceptable loudness for the three sounds. During presentation the sound originated

Table 1. Mixed effects analysis of accuracy, across all five participants.
Optimised Model: accuracy ~ duration * location + duration * sound + trial + (1 | participant). This means duration, location, sound and the interactions duration*location and duration*sound are the fixed effect predictors, and random effects for the intercepts by participant are required.

Predictor			
	Sequential ANOVA of the model		
	Degrees of Freedom (DF)	Sum of Squares	Mean Sum of Squares (by DF), which is also the F-value
Duration (sonds)	1	218.51	218.51
Location/Movement	3	221.55	73.85
Sound	2	25.16	12.58
Trial number (centred, rescaled)	1	3.11	3.11
Duration x Location	3	54.61	18.20
Duration x sound	2	36.37	18.18
	Random Effects parameter and Confidence Intervals for the Fixed effect coefficients		
Random effects	S.D.		
Intercept by participant	0.57		
Fixed Effects	2.5% Confidence Interval	97.5% Confidence Interval	
Intercept	2.41	4.36	
COEFFICIENTS:			
Duration	-0.10	+0.34	
Location: moving LR	-6.16	-4.41	
Location: moving RL	-5.59	-3.85	
Location: R	-1.70	+0.39	
Sound: 30Hz sine	0.87	1.73	
Sound: 60Hz sine	1.40	2.27	
Trial number	0.0004	0.15	
Duration x Location (moving LR)	0.51	0.95	
Duration x Location (moving RL)	0.42	0.85	
Duration x Location R	-0.10	+0.44	
Duration x Sound (30Hz)	-0.47	-0.15	
Duration x Sound (60Hz)	-0.64	-0.32	

either solely in the Left (L) or Right (R) speaker, or 'moved' with a linear constant power transit L->R or R->L over the whole sound duration. Listeners were required after the end of the sound (and not before) to indicate as quickly as possible which of the four categories of location/movement event they had just heard. There were in total 3(duration) x 3(sound source) x 4 (location/movement) = 36 different stimuli. These were presented in randomized order, each stimulus 8 times in each of two blocks, with the listener requested to take a break between the two. Thus each participant responded to 576 stimuli. This achieved the commonly used 'roving' of sound sources in terms of intensity and frequency, intended to minimize listeners recognition of specific colouring attached to individual speakers or positions in the space.

A mixed effects analysis of the data (using glmer from the lme4 package in R since the data are binomial) to model the accuracy of responses is summarized in Table 1. There were random effects for the intercept by participant; accounting for this reduces the likelihood of Type 1 errors in the analysis (e.g. [54]). The sequential ANOVA of the glmer model suggests that the main explanatory power is provided by duration, sound location/movement and their interaction; though it is important to note that the exact values in such a sequential ANOVA depend somewhat on the order in which the parameters are entered. The confidence intervals (which also reveal the mean coefficient, as the centre of the range) show that the effect of Duration is largely carried in the interaction with location (Duration itself is not effective). The moving sounds are both much worse identified than the static ones, which are not different from each other (as shown by the fact that the CI on the coefficient for R, which is referenced against L as base, breaches zero). The two sine tones are both better located than the noise sound (which is the base level, not shown in the table), though this modest effect is reduced as the sound duration increases: but note that they were not matched in dB at the listening position as I did not seek to understand the influence of timbre. There is a small improvement as the experiment proceeds (reflected in the positive coefficient on Trial number, which was centred and rescaled before analysis).

The modelling was guided by Bayesian Information Criteria values, coefficient significance and parsimony, and optimized by likelihood ratio tests. The optimized model had a BIC of 1814.74. Confidence intervals for the fixed effects parameters were determined by lme4's 'confint' function, using the 'profile' technique. The more approximate 'Wald' technique gave very similar values. The random effects are modelled as a distribution whose standard deviation is measured, but are not of primary interest here. The fixed effects are expressed as coefficients and confidence intervals (which are here symmetrical), and where a predictor has several categories (sound, location) or distinct continuous values (duration), the coefficients are the difference from the 'base' level, which is the level which does not appear in the table. The model was worsened by treating the location as comprising fixed vs moving, hence this approach is not shown.

Figure 1 summarizes the salient comparisons, as judged by the mixed effects analysis, and shows that for the static sounds, accuracy was extremely high, but it was much worse for the moving sounds, increasing with their duration. It also shows that participants had difficulty in locating the very short sounds, and this was almost entirely due to the short moving sounds (the interaction shown in the figure as Moves/0.2 seconds). The difference in performance for the different sounds (which were in any case intended as roving stimuli rather than maximally controlled, and so not shown) was very small.

The 0.2 s movement sounds had movement rates of 450 degrees/sond, beyond the rate at which MAMAs are optimal in free field conditions (discussed above). The inability to judge movement in these stimuli was thus expected. Correspondingly, Table 2 shows an aggregated contingency table for the responses in relation to the stimulus location/movement. It shows that the moving sounds created confusion, where generally the starting point of the movement was taken to be its static location when the participant failed to recognize the movement. This effect is similar to the Franssen illusion already described.

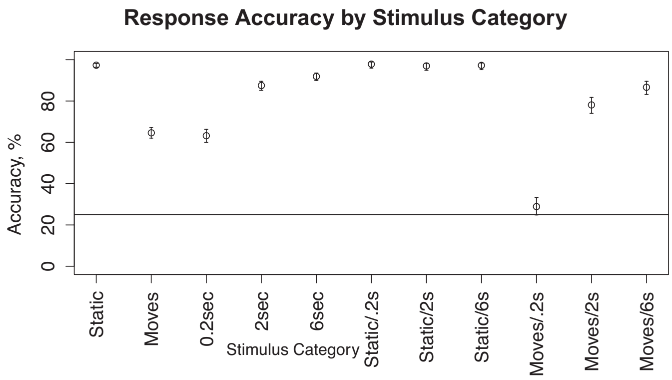

Figure 1. Summarized accuracies in detection of localization and movement in the various conditions. The graph shows percent hit rate with 95% confidence intervals, based as conventionally on pooling all five participants' data. Thus the categories overlap in what is shown: the combinations Static/Moves; the three durations; and the 6 duration/location interactions; each include all the data (2804 responses). All individual results (except Moves/.2s) are highly significant at $p < 0.00001$ in comparison with the chance rate (25%, shown as a horizontal line). A full analysis by mixed effects modelling is provided above. The text and Appendix describe the conditions in more detail. Confidence intervals were also determined by a more correct statistical meta-analysis of the separate confidence intervals determined on each (independent) participant for the stimulus categories shown, and these were a little larger, but confirmed all the conclusions.

Table 2: Contingency table of percentages of responses to each location/movement category of stimulus. Numbers on the diagonals are the correct response percentages.

Location/	Response:			
Movement	L	R	L->R	R->L
L	97.5	0.7	1.6	0.1
R	0.4	97.2	1.3	1.1
L->R	26.9	9.6	61.3	2.1
R->L	4.3	25.0	2.9	67.9

It is important to bear in mind (as detailed in the Appendix), that while the digital signal leaving the MAXMSP synthesis in each case contained very little energy at frequencies above 200 Hz, that generated by the subwoofers contained some energy detectable above the acoustic background of the studio, in some cases at frequencies up to about 600 Hz. Correspondingly, sending high dB sine tones to the speakers produced pitch-discernible audible sound at least up to 1000 Hz. Thus even 'clean' low frequency sounds as presented by these excellent speakers will always contain higher frequencies. The same observations held for larger (much more expensive!) Meyer subwoofers.

CONCLUSIONS

Can a composer use low frequency lateralization and movement, and hope for perceptibility?

I conclude that musically-experienced listeners have good capability in relation to location/lateralization and movement perception of low frequency sounds in our reverberant studio environment, though movement accuracy is much lower than location accuracy, as normally observed. There was worse performance with very short sounds. Our listeners may have learnt much about low frequency listening from their musical experience; but nevertheless, their performance improved slightly with trial as the experiment proceeded. On the other hand, I argue that most people learn from environmental sound around them, and that there is a biological advantage in gaining ability to localize even low frequency sounds, especially if they are moving. This remains to be more fully tested.

The results support the view that sound projection systems with multiple sub-woofers can add timbral flux and spatial control to composers' armories. It will be interesting to assess the influence of sub-woofers at different elevations [55], in 3D space, in addition to different azimuths in 2D space, as I have done here. This is especially so given the readily available Max patches for sound diffusion and movement, the panning software VBAP, and specialized facilities like those at ZKM introduced above, and the 22:4 system in our studio. In 3D spatialization, issues of front-back discrimination also come into play (not discussed here). These are generally far more problematic for listeners than lateral location or movement detection [3], and hence will require considerable attention.

Interesting in the longer run, both psychoacoustically and musically, are questions concerning the possible competition between musical low frequency and high frequency timbre spatialization. While already in practical use, the impact of these perceptually is little understood. Experimental acoustics and psychoacoustics will clearly have more to contribute to composers and performance space design in this area.

REFERENCES

[1] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. MIT Press, Cambridge, MA (1997)

[2] C. Stecker and F. Gallun, "Binaural hearing, sound localization, and spatial hearing", *Translational perspectives in auditory neuroscience: normal aspects of hearing*. San Diego (CA): Plural Publishing, Inc, (2012)

[3] T.R. Letowski and S.T. Letowski, *Auditory spatial perception: auditory localization*. 2012, DTIC Document.

[4] L. Rayleigh, "XII. On our perception of sound direction", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **13**, 214-232 (1907)

[5] A.R. Palmer, "Reassessing mechanisms of low-frequency sound localisation", *Current opinion in neurobiology* **14**, 457-460 (2004)

[6] M.I.J. Mohamed and D. Cabrera, "Human sensitivity to interaural phase difference for very low frequency sound", in *Acoustics 2008 - Australian Acoustical Society Conference*. p. p47.pdf.

[7] P. Lennox, Spatialization and Computer Music, in *The Oxford Handbook of Computer Music*, ed., R.T. Dean, Oxford University Press, New York, USA, 2009) p. 258.

[8] R. Normandeau, "Timbre Spatialisation: The medium is the space", *Organised Sound* **14**, 277 (2009)

[9] S. Kanach, *Music and Architecture by Iannis Xenakis*. 2008, Hillsdale, Pendragon Press.

[10] D. Worrall, "Space in sound: sound of space", *Organised Sound* **3**, 93-99 (1998)

[11] A.J. Hill, S.P. Lewis, and M.O. Hawksford, "Towards a generalized theory of low-frequency sound source localization", *Proceedings of the Institute of Acoustics* **34**, 138-149 (2012)

[12] P. Laine, H.-M. Lehtonen, V. Pulkki, and T. Raitio. "Detection and lateralization of sinusoidal signals in presence of dichotic pink noise", in *Audio Engineering Society Convention 122*. 2007. Audio Engineering Society.

[13] D. Cabrera, S. Ferguson, and A. Subkey. "Localization and image size effects for low frequency sound", in *Audio Engineering Society Convention 118*. 2005. Audio Engineering Society.

[14] W.L. Martens. "The impact of decorrelated low-frequency reproduction on auditory spatial imagery: Are two subwoofers better than one?" in *Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction*. 1999. Audio Engineering Society.

[15] A.J. Hill and M.O. Hawksford. "On the perceptual advantage of stereo subwoofer systems in live sound reinforcement", in *Audio Engineering Society Convention 135*. 2013. Audio Engineering Society.

[16] T.S. Welti. "Subjective comparison of single channel versus two channel subwoofer reproduction", in *Audio Engineering Society Convention 117*. 2004. Audio Engineering Society.

[17] J. Braasch, W.L. Martens, and W. Woszczyk. "Identification and discrimination of listener envelopment percepts associated with multiple low-frequency signals in multichannel sound reproduction", in *Audio Engineering Society Convention 117*. 2004. Audio Engineering Society.

[18] T. Nakayama, T. Miura, O. Kosaka, M. Okamoto, and T. Shiga, "Subjective assessment of multichannel reproduction", *Journal of the Audio Engineering Society* **19**, 744-751 (1971)

[19] G. Paine, R. Sazdov, and K. Stevens. "Perceptual Investigation into Envelopment, Spatial Clarity, and Engulfment in Reproduced Multi-Channel Audio", in *Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio*. 2007. Audio Engineering Society.

[20] M. Tucker, Dreaming with open eyes. The shamanic spirit in twentieth century arts and culture. Aquarian, London (1992)

[21] N.P.M. Todd and F.W. Cody, "Vestibular responses to loud dance music: A physiological basis of the "rock and roll threshold"?", *The Journal of the Acoustical Society of America* **107**, 496-500 (2000)

[22] R.T. Dean and F. Bailes, Event and process in the fabric and perception of electroacoustic music in *Proceedings of the international Symposium: Xenakis. The electroacoustic music* (http://www.cdmc.asso.fr/sites/default/files/texte/pdf/rencontres/intervention11_xenakis_electroacoustique.pdf), ed., (Centre de Documentation Musique Contemporaine, 2013) pp. Intervention11, pp.12.
[23] H. Smith, The voice in computer music and its relationship to place, identity and community, in *The Oxford Handbook of Computer Music*, ed., R.T. Dean, (Oxford University Press, New York, 2009) pp. 274-293.
[24] R.T. Dean and F. Bailes, "NoiseSpeech", *Performance Research* **11**, 134-135 (2007)
[25] H.A. Smith and R.T. Dean, "Voicescapes and Sonic Structures in the Creation of Sound Technodrama", *Performance Research* **8**, 112-123 (2003)
[26] M. Södersten, S. Ternström, and M. Bohman, "Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers", *Journal of Voice* **19**, 29-46 (2005)
[27] R.C. Smith and S.R. Price, "Modelling of human low frequency sound localization acuity demonstrates dominance of spatial variation of interaural time difference and suggests uniform just-noticeable differences in interaural time difference", *PLoS One* **9**, e89033 (2014)
[28] D. Huron, *Sweet anticipation.* MIT Press, Cambridge, MA (2006)
[29] R.T. Dean, *Hyperimprovisation: Computer interactive sound improvisation; with CD-Rom.* A-R Editions, Madison, WI (2003)
[30] P. Hegarty, "Just what is it that makes today's noise music so different, so appealing?", *Organised Sound* **13**, 13 (2008)
[31] D.R. Perrott and J. Tucker, "Minimum audible movement angle as a function of signal frequency and the velocity of the source", *The Journal of the Acoustical Society of America* **83**, 1522-1527 (1988)
[32] K. Saberi and D.R. Perrott, "Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate", *The Journal of the Acoustical Society of America* **87**, 1732-1737 (1990)
[33] L. Landy, Sound-based music 4 all, in *The Oxford Handbook of Computer Music*, ed., R.T. Dean, (Oxford University Press, New York, USA, 2009) pp. 518–535.
[34] T.D. Rossing, D.A. Russell, and D.E. Brown, "On the acoustics of tuning forks", *Am. J. Phys* **60**, 620-626 (1992)
[35] E.A. Macpherson and J.C. Middlebrooks, "Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited", *The Journal of the Acoustical Society of America* **111**, 2219-2236 (2002)
[36] J. Ahveninen, N. Kopčo, and I.P. Jääskeläinen, "Psychophysics and neuronal bases of sound localization in humans", *Hearing Research* **307**, 86-97 (2014)
[37] J. Liu, D. Perez-Gonzalez, A. Rees, H. Erwin, and S. Wermter, "A biologically inspired spiking neural network model of the auditory midbrain for sound source localisation", *Neurocomputing* **74**, 129-139 (2010)
[38] W.O. Brimijoin, A.W. Boyd, and M.A. Akeroyd, "The contribution of head movement to the externalization and internalization of sounds", *PLoS One* **8**, e83068 (2013)
[39] W. Martens, S. Sakamoto, L. Miranda, and D. Cabrera. "Dominance of head-motion-coupled directional cues over other cues during walking depends upon source spectrum" in *Proceedings of Meetings on Acoustics*. 2013. Acoustical Society of America.
[40] D.R. Perrott and A. Musicant, "Minimum auditory movement angle: Binaural localization of moving sound sources", *The Journal of the Acoustical Society of America* **62**, 1463-1466 (1977)
[41] D.W. Grantham, "Detection and discrimination of simulated motion of auditory targets in the horizontal plane", *The Journal of the Acoustical Society of America* **79**, 1939-1949 (1986)
[42] N.P. Todd, A.C. Paillard, K. Kluk, E. Whittle, and J.G. Colebatch, "Vestibular receptors contribute to cortical auditory evoked potentials", *Hearing Research* **309**, 63-74 (2014)
[43] B. Blesser and L.-R. Salter, *Spaces speak, are you listening?* : MIT press, (2007)
[44] E. Larsen, N. Iyer, C.R. Lansing, and A.S. Feng, "On the minimum audible difference in direct-to-reverberant energy ratio", *The Journal of the Acoustical Society of America* **124**, 450-461 (2008)
[45] H. Al Saleh, *Effects of reverberation and amplification on sound localisation.* 2011, University of Southampton.
[46] A. Ihlefeld and B.G. Shinn-Cunningham, "Effect of source spectrum on sound localization in an everyday reverberant room", *The Journal of the Acoustical Society of America* **130**, 324-333 (2011)
[47] W.M. Hartmann, "Localization of sound in rooms", *The Journal of the Acoustical Society of America* **74**, 1380-1391 (1983)
[48] W.M. Hartmann and E.J. Macaulay, "Anatomical limits on interaural time differences: an ecological perspective", *Frontiers in neuroscience* **8**, (2014)
[49] A. Brughera, L. Dunai, and W.M. Hartmann, "Human interaural time difference thresholds for sine tones: the high-frequency limit", *The Journal of the Acoustical Society of America* **133**, 2839-2855 (2013)
[50] W.M. Hartmann and A. Brughera, "Threshold interaural time differences and the centroid model of sound localization", *The Journal of the Acoustical Society of America* **133**, 3512-3512 (2013)
[51] F.L. Wightman and D.J. Kistler, "The dominant role of low□ frequency interaural time differences in sound localization", *The Journal of the Acoustical Society of America* **91**, 1648-1661 (1992)
[52] R.Y. Litovsky, H.S. Colburn, W.A. Yost, and S.J. Guzman, "The precedence effect", *The Journal of the Acoustical Society of America* **106**, 1633-1654 (1999)
[53] W.M. Hartmann and B. Rakerd, "Localization of sound in rooms IV: The Franssen effect", *The Journal of the Acoustical Society of America* **86**, 1366-1373 (1989)
[54] H. Quené and H. Van den Bergh, "Examples of mixed-effects modeling with crossed random effects and with binomial data", *Journal of Memory and Language* **59**, 413-425 (2008)
[55] V.R. Algazi, C. Avendano, and R.O. Duda, "Elevation localization and head-related transfer function analysis at low frequencies", *The Journal of the Acoustical Society of America* **109**, 1110-1122 (2001)
[56] C. Hak, R. Wenmaekers, and L. Van Luxemburg, "Measuring Room Impulse Responses: Impact of the Decay Range on Derived Room Acoustic Parameters", *Acta Acustica United with Acustica* **98**, 907-915 (2012)

Appendix A - The studio and experimental set up

The performance research space at MARCS Institute is roughly rectangular, 8 x 6.1 metres. It has 22 speakers in the roof area, up to 4 subwoofers on or near the floor, and two large retractable projection screens. These screens are on adjacent corrugated acoustic walls. The speakers for the experiment were 0.7 metres from these walls, and 3.3metres from the listening position, facing inwards. The other two walls of the studio contain glass windows allowing adjacent rooms to function as control rooms. The windows have retractable curtains, and these make a large difference to the coloration of sound in the studio, so the experiments were done with them completely open. The speaker and listener positions were chosen such that there was no consistent audible coloration distinction between the speakers with a wide range of input tones.

The Harlequin brand floor has a vinyl cover on top of wooden squares, providing suitable spring for dance use, and there were some normal studio items inside the room. The reverberation time (T30) measured from impulse responses in accordance with ISO 3382 [56] was 400 ms, and at 30 Hz this was extended to about 700 ms. The sounds were intended to be roving stimuli, and so they were not exactly matched for intensity: at the position of the listeners' heads they were measured to be 40-47 dB(A), using a Bruel and Kjaer 2250 sound level monitor. The noise sound was set at the lower intensity for listener comfort. Background noise levels in the studio were c. 30 dB.

The sounds were generated in MAXMSP as white noise, and as 30 Hz and 60 Hz sine tones. Each sound was digitally filtered through a MAXMSP low pass resonator (to reduce frequencies in the tones above 100 Hz, with 24 dB per octave roll off). The differences between the three spectra were obvious, and as expected. A MOTU 896 mk. III digital interface was used (sampling rate 44.1 kHz; output level -3dB), and the Genelec speakers were at default settings, with their nominal cut-off (equivalent to a cross-over) frequency being 120 Hz. They were visible, and the room was illuminated. It was noted that the loudness of all the sounds ramped rapidly to a maximum, but if sustained, then after 7-8 seconds it dropped to a new steady state which continued unchanged for a prolonged time. This was also observed with large Meyer subwoofers, and hence durations longer than 6 seconds were not suitable, and not used. The power spectrum of the sounds at the listeners' position was measured, and it showed the vast majority of energy to be below 200 Hz, but at higher frequencies there was slight energy above background levels, declining strongly and progressively with frequency. The 30 Hz stimulus was above background up to 400 Hz, the 60 Hz to 600 Hz (particularly during the transient ramp on), and the noise tone to 500 Hz.

The 5 participants (mean age 39.0 years, s.d. 17.2) all had musical experience; 4 also had experience of recording technology and practice; and there was one female. During an experiment, the stimuli were presented in randomized order, and in two blocks, each of 288 stimuli; participants could take a break between the blocks. They fixated on the computer screen while listening, and were asked to make their judgement of location/movement as quickly as possible after each sound had ended. Number keys were pressed to indicate the location (1 for L; 2 for R; 3 for LR; 4 for RL). Data was also recorded in the MAXMSP patch.