Proceedings of the International Symposium on Room Acoustics, ISRA 2010

29-31 August 2010, Melbourne, Australia

Characterising the variation in oral-binaural room impulse responses for horizontal rotations of a head and torso simulator

Densil Cabrera (1), Doheon Lee (1), Ralph Collins (1), Bastian Hartmann (1), William L. Martens (1) and Hayato Sato (2)

(1) Faculty of Architecture, Design and Planning, University of Sydney, NSW 2006, Australia
(2) Environmental Acoustics Laboratory, Kobe University, Rokkodai, Nada, Kobe, Japan

PACS: 43.55.Mc, 43.55.Gx

ABSTRACT

Oral-binaural room impulse responses (OBRIRs) describe the room acoustical response from the mouth to the ears of a head or dummy head. In this study, we measured OBRIRs in ten rooms, ranging from small to large. In each room, a head and torso simulator (HATS) was rotated at 2 degree increments to sample the room response at the selected measurement position. In rotating the HATS, the radiation pattern of the mouth rotates with the reception pattern of ears. This paper characterises the variation in room gain and interaural response of the tested rooms, and in doing so, we consider how OBRIRs can be usefully understood in terms of acoustical parameters.

INTRODUCTION

The sound of one's own voice during speech is influenced by room acoustical context, but the acoustical characterisation of this is still a relatively immature aspect of room acoustics. Stage support parameters [1, 2] are probably the best-known approach to characterising the room-reflected sound returned to a person, but they are difficult to apply in small rooms, and the bulk of stage support studies are concerned with music auditoria. A binaural approach to the problem was taken by Brunskog *et al.* [3], using a head and torso simulator to derive the acoustical parameter 'room gain', which quantifies the increased sound returned to the ear simulators due to the room (compared to anechoic conditions). The present authors, too, have used a binaural approach to measuring room-reflected sound with the aim of simulating environments for talking subjects [4], and the purpose of this paper is to examine the information in our measurements to provide a better understanding of how rooms affect oral-binaural sound transmission (i.e. from the mouth to ears of the same head).

Before dealing with the details of oral-binaural room measurements, we shall briefly consider what the purpose of such measurements might be. The study of Brunskog *et al.* [3] is concerned with the relationship between classroom acoustics and teacher voice strain, showing that room gain is one of the factors affecting the vocal power used by talkers. In stage acoustics (for example, in drama or opera theatres) sound returned to the actor or singer may provide 'stage support', and perhaps an oral-binaural measurement could quantify this as an alternative to conventional stage support measurements. Stage support measurements are essentially omnidirectional, but the spatial distribution of the early-reflected soundfield on a stage or in a classroom is unlikely to be even [5]. Something of this spatial distribution will be captured by binaural measurement due to the peripheral auditory system's spatial resolving properties. The speaking and singing voice are directional too (especially at high frequencies) [6, 7], and there may be some benefit in making measurements where sound is projected to the direction that a teacher, actor or singer might typically face.

Oral-binaural measurements also allow the implementation of an interactive simulation system (which was the primary purpose of our measurements [4]), and related work has also been done by Pörschmann [8] and Sato *et al.* [9] using synthesised (rather than measured) oral-binaural room characteristics. Such simulations are useful in the development of subjective attributes of room-reflected sound from one's voice.

Spatial hearing has mainly been studied in simple acoustical environments, usually anechoic, and the bulk of attention has been on directional perception. It is well-established that people learn to subconsciously interpret their own head-related transfer functions to discern the direction of sound sources [10], and it is likely that people also learn to interpret features of their environment by ear, for example, in near-field auditory echo-location [11], auditory distance perception [12], and indeed gross environmental features such as room size [13]. Hence, the study of oral-binaural characteristics should contribute to understanding such aspects of spatial hearing in the context of rooms.

Qualities of rooms derived from the sound of one's own voice are likely to extend beyond the simple loudness of the room-reflected sound. Other possible qualities may be somewhat analogous to those that are used in the more conventional room acoustics where sources are spatially separated: such as the prolongation of sound due to reverberance, binaural qualities, and spectral qualities.

This paper examines a set of oral-binaural room impulse responses (OBRIRs) measured in a variety of rooms, as described by Cabrera *et al.* [4]. While these OBRIRs were measured for the purpose of simulating the environments, the aim of the present paper is to examine acoustical features within these OBRIRs. The paper merely contains an objective analysis of the OBRIR features, and subjective data will be needed to provide a better understanding of the relevance and relative importance of the objective parameters to people's experience of their voice in rooms.

MEASUREMENT METHOD

Using a Brüel & Kjær 4128C head and torso simulator (HATS), we measured OBRIRs between a microphone at the mouth reference point (Brüel & Kjær type 4939) and microphones at the entrance of the ear canals (Brüel & Kjær type 4101). Although Brunskog *et al.* used the same model HATS, their approach was to measure from the loudspeaker to built-in ear microphones. The reason why we did not use the built-in microphones was that they include an emulation of the ear canal, which has a strong resonant peak – and we wished to avoid the need to invert this in the room simulation system. Our use of the mouth reference point (rather than the loudspeaker itself) as the origin of the transfer function was to remove the loudspeaker's response from the measurement, and because we could place a microphone in a similar location for a speaking subject in the simulation system.

Measurements were made using a swept sinusoid (logarithmic, 50 Hz – 15 kHz), which was recorded on both of the ear microphones and the mouth microphone. Calibration tones were also recorded on each microphone so that channel gain could be matched for transfer function calculation. The transfer function from the mouth microphone to each ear microphone was derived in the frequency domain, with spectral components below 100 Hz and above 10 kHz removed prior to returning to the time domain with the impulse response. In each room, measurements were made with the HATS mounted on a turntable, and OBRIRs were derived at 2º increments over a 120º rotation. The original purpose of this was so that head-tracking could be used to account for incidental head rotations of the subject using the simulation system – but this also allows us to examine how OBRIRs vary as a HATS rotates.

The turntable surface included some equipment such as microphone power supplies, and these were covered with a 50 mm layer of porous sound absorber (Tontine Acoustisorb 3) to reduce reflection from these and the turntable surface. The rooms in which measurements were made are described later in this paper. More details of the measurement method are given by Cabrera *et al.* [4].

POTENTIAL OBRIR PARAMETERS

There are two main reasons for deriving parameters from OBRIRs. The first is to reduce the large amount of information within OBRIRs to summary data, so that features are comprehensible. The second is to represent features of OBRIRs that are most relevant to the perception of corresponding room environments during speech – i.e. psycho acoustical parameters. In this paper we are mainly limited to the first reason because of the lack of direct experimental evidence to support the second.

Room gain

Room gain represents the energy of the room-reflected sound, in decibels (relative to the energy of the received sound without room reflections), and is defined by Brunskog *et al.* [3]. The difference between Brunskog's and our approaches is that our impulse responses are from the mouth reference point to the microphones outside the ear canal (rather than from the mouth simulator to the microphone within the ear simulator). However, this difference is relatively minor. Room gain (G_{RG}) is defined as the ratio of the squared impulse response ($h^2(t)$) energy to that of the direct sound, expressed in decibels (equation 1).

$$G_{RG} = 10\log_{10}\frac{\int h^2_{direct+room}(t)dt}{\int h^2_{direct}(t)dt} \qquad (1)$$

The direct sound of OBRIRs should include corporeal acoustic effects such as reflections from the shoulders, and so is not a simple band-passed impulse (but is nonetheless brief). Of course, the room gain may differ at the two ears, and so we take the power average of the value for each ear to derive an overall value for G_{RG}.

It should be noted that the direct sound energy does not represent the energy that a person hears when they speak in an anechoic room because it does not include corporeally transmitted sound (usually referred to as 'bone conduction'). Its significance is merely to represent the airborne component of the direct sound

Interaural parameters

We can also derive measures of the similarity of the signals between the two ears. The interaural level difference (*ILD*) and interaural cross correlation coefficient (*IACC*) are obvious possibilities for this. *ILD* is very simply calculated from the difference between the left and right ear room gains (equation 2) excluding the direct sound. Although another approach would be to include the direct sound (which has considerable strength, and a *ILD* of 0 dB), in this paper we focus on the *ILD* of room reflections.

$$ILD = 10\log_{10}\frac{\int h^2_{room,left}(t)}{\int h^2_{room,right}(t)} \qquad (2)$$

The purpose of measuring *ILD* is to indicate the left-right bias of the soundfield: i.e., is the reflection energy predominantly received by one ear or the other, or are they equally received by the two ears? This is to be distinguished from the more conventional use of *ILD* in auditory localization theory, where *ILD* of the direct sound from a source may be an indicator of lateralisation or source proximity (in concert with interaural time difference). In using *ILD* measured from OBRIRs, we are not concerned with the direct sound, but with a multitude of reflections from many directions distributed over a substantial time period, and so we are not making simple inferences about lateralisation.

IACC is more commonly used than *ILD* in room acoustics to characterise interaural dissimilarity or similarity, but there are many approaches to measuring and calculating *IACC*-related values. With regard to measurement, the ear microphone position and dummy head that we used is the same as that described in ISO3382-1 [2]: small microphones at the entrance of open ear canals. However, in its conventional application, *IACC* is applied to exocentric (rather than egocentric) sources, and the direct sound is included in such calculations. As discussed earlier, the direct sound of an OBRIR does not represent what a person hears, and so we cannot include it in the *IACC* calculation; and furthermore, the direct sound of an OBRIR is independent of the measurement environment, and so would not contribute meaningfully to an OBRIR-based

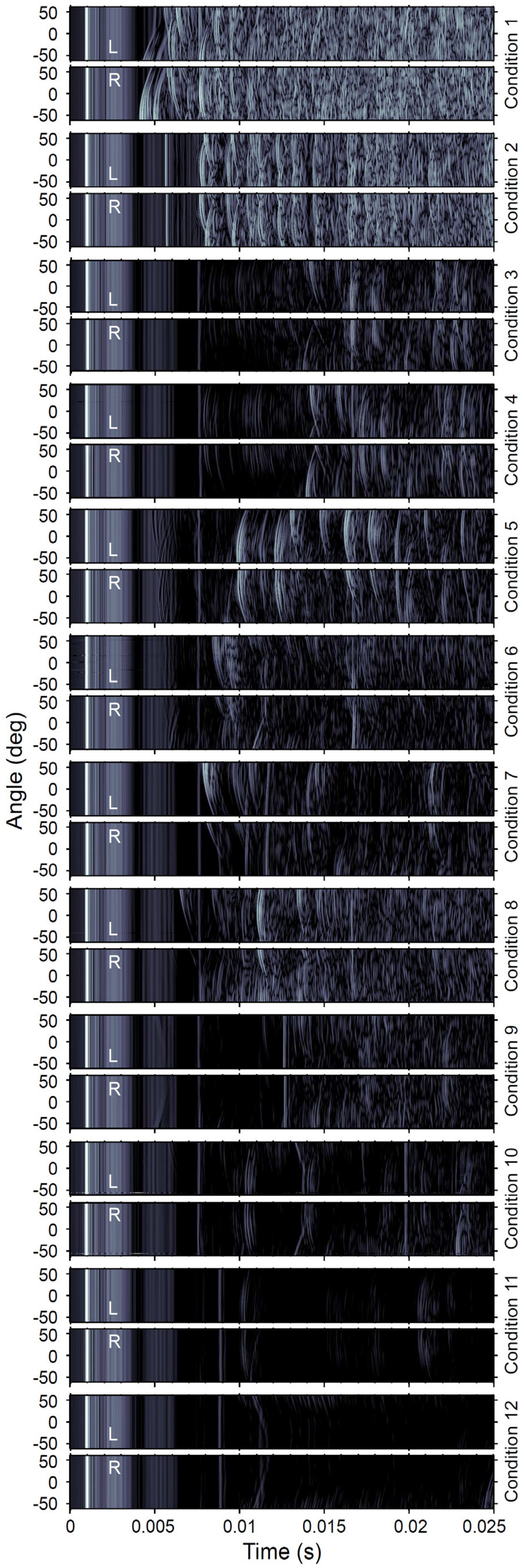

Figure 1. The envelope of the first 25 ms of the OBRIRs as a function of time and angle of rotation of the apparatus. For each condition, L refers to the left ear, and R to the right ear. Black represents an instantaneous sound level 100 dB less than white.

IACC measurement. Apart from omitting the direct sound, we calculate *IACC* conventionally following equation 3.

$$IACC = \max \left| \frac{\int h_l(t) h_r(t+\tau) dt}{\sqrt{\int h_l^2(t) dt \int h_r^2(t) dt}} \right| \tag{3}$$

In equation 3, τ is the lag offset used for cross correlation, which ranges between ±1 ms (which is about the natural range for interaural time difference). The subscripts l and r refer to left and right ears. The integration period is considered below.

Evaluation periods

The direct sound period in our measurements is well-defined, with essentially all of its energy within 3 ms at the start of the OBRIRs. Depending on the height of the HATS, this is followed by a floor reflection several milliseconds later (about 6.5 ms for a transducer height of 1.2 m, and 8.0 ms for a height of 1.5 m). In most of our measurements, the floor is the first room reflection, but in some measurements other nearby surfaces reflect earlier. Our approach is to exclude the direct sound (except for its use as a reference level), but to include the floor-reflected sound, in the evaluation period. It should be noted, though, that the partial absorption of the floor reflection (due to the sound absorptive material on the turntable) has some influence on the results reported in this paper. Considering that the thighs of a seated person are not modelled by a HATS, we can appreciate that there is room to improve the treatment of the first order reflections from below the head.

There are various end points to evaluation periods commonly used in room acoustics. *IACC* and some other parameters of ISO3382-1 use an 80 ms evaluation period for the early sound [2]. However, clarity index as applied for speech uses a 50 ms evaluation period [14], and it could be argued that this is more relevant to OBRIRs, because they are often concerned with the sound of speech (rather than music). On the other hand, the period from 20 ms to 100 ms is used for early stage support, which is also similar in concept to OBRIRs because it is an egocentric room acoustical measurement. Other parameters use the entire impulse response, which is the approach taken for room gain (as per Brunskog *et al.*).

In view of these options, in this paper we present data for two evaluation periods: *full* (from just after the direct sound to the end of the OBRIR); and *early* (from just after the direct sound to 80 ms after the direct sound).

Spectral weighting

There are many ways in which OBRIR parameters could be weighted, and one consideration might be the spectral characteristics of the voice. However, in this paper we take a simple approach, which is to follow the weighting schemes used previously for room impulse response parameters. The spectral weighting of room gain, as deployed by Brunskog *et al.*, is the arithmetic mean of the octave band room gains from 125 Hz to 4 kHz. For consistency, we use this approach here. We use most commonly applied spectral weighting method for *IACC* ($IACC_3$), which is the average value of the octave bands centred on 500 Hz, 1 kHz and 2 kHz. We have used the same octave band range mean for *ILD*.

SELECTED ROOMS

Twelve measurements were made within 10 rooms (i.e., in two of the rooms, two measurements were made in different positions or with changed room acoustic treatment). The rooms were all in the University of Sydney, mostly within the Wilkinson Building (where the Faculty of Architecture, Design and Planning is housed). The one room in another building was the Verbrugghen Hall, a recital hall at the University's Conservatorium of Music. Rooms were selected for variety. In this section we outline the physical and acoustic characteristics of these measurement conditions, in order of room volume.

The 12 measurements considered in this paper, referred to as conditions 1 to 12, are described over the following pages. Figure 1 provides a summary of the early reflection patterns (0-25 ms) for the twelve conditions (in the figure, the direct sound is at 1 ms). Data in Figure 1 are the absolute value of the Hilbert-transformed OBRIR, on a logarithmic greyscale, with a 100 dB range (black is 100 dB less than white). The figure shows how each OBRIR varies as the HATS rotates over a 120º range (±60º). The floor reflection is visible in many of the conditions at 7 ms (and 9 ms for conditions 11 and 12, where the HATS ear height was 1.5 m rather than 1.2 m), and the timing and strength of the floor reflection is unaffected by rotation. In some conditions a ceiling reflection is also apparent (very early in condition 2), and this is also unaffected by rotation. Lateral reflections are maximally affected by rotation, and these display a change arrival time and level as a function of angle, which differs between the two ears.

Figure 1 also clearly shows a large contrast between the reflection densities of the conditions (at least for the very early period). Furthermore, in some conditions (e.g. 2 and 5) the sequence of early reflections is fairly regular, whereas in others (e.g. 1 and 8) the sequence is more random - presumably due to more complex room surfaces.

Conditions 1 & 2: Very small reverberant rooms (toilet 9.1 m^3; mirror chamber sky 15.3 m^3)

The first two rooms that we examine are a toilet room (Condition 1) and a mirror chamber sky room (Condition 2). Both are small and reverberant. Although the acoustics of toilets may not be a major stream of room acoustics research, such rooms are interesting from a phenomenological standpoint – they are small volume rooms with hard surfaces and are experienced by people every day. As such they should make a significant impression on our learnt association between physical features of the environment and the sound that we hear. The particular room measured for Condition 1 is relatively large for disabled access, and without any internal partitions. It has a mid-frequency reverberation time of 1.2 s. The apparatus was set up in the least cluttered part of the room.

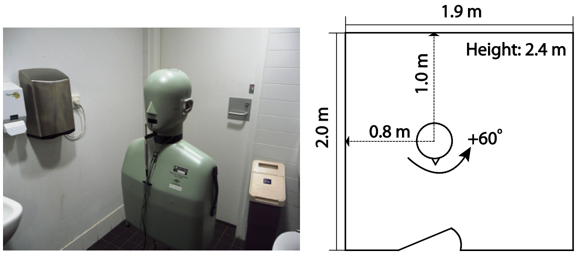

Figure 2. Photograph of the measurement apparatus in the toilet room (left) and a plan showing the position and orientation of the apparatus (right).

Figure 1 shows that condition 1 has the earliest reflections of all the conditions, with the first reflection from the nearest wall reaching its maximum strength in the right ear when the mouth and right ear are both turned towards the wall (about -50º on the chart). The initial time delay gap (ITDG), which we could define as the delay between the direct sound and the first major non-floor reflection (because the timing of the floor reflection is constant), is only 3 ms at its shortest (HATS rotated to -50º), lengthening to 3.8 ms with the HATS rotated to +60º).

A mirror chamber sky is designed for the examination of the light distribution cast by an overcast sky upon and within architectural models. The diffuse sky is created from a diffuse artificial light source covering the entire ceiling, which is optically extended by mirrors on each of the four walls (Figure 3). In Condition 2, the ceiling luminaire is the major sound absorber in the room. The volume of the room is greater than Condition 1, and the reverberation time is less (1.0 s), but this is still a highly reverberant very small room. The ceiling is low (2.1 m), making it the source of the first reflection (ITDG of 4.7 ms).

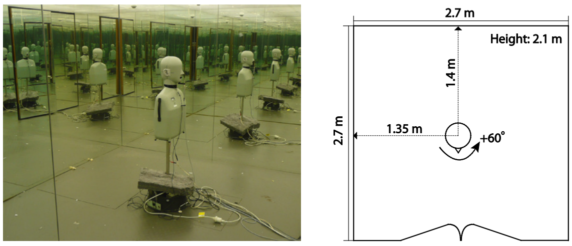

Figure 3. Photograph of the measurement apparatus reflected in the mirror chamber sky room (left) and a plan showing the position and orientation of the apparatus (right).

The OBRIR parameters for Conditions 1 and 2 are shown in Figure 4, as a function of the angle of rotation of the apparatus. Notable features are the very high G_{RG} values, the relatively small range of *ILD* (especially Condition 1), the very low *IACC* values in Condition 1 (which contrasts with the prominent IACC peak in condition 2), and the >5 dB (i.e. large) separation between early and full G_{RG}.

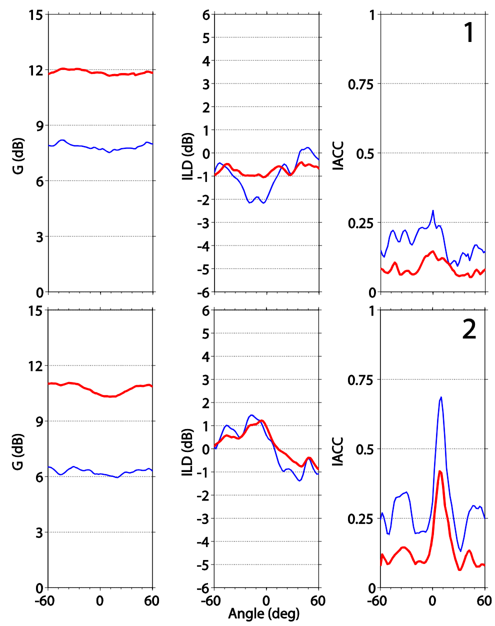

Figure 4. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the toilet room (top) and mirror chamber sky (bottom). The blue lines are for the early evaluation period, and the red (heavier) lines are for the full period.

While the parameters of Figure 4 are indicative of small reverberant rooms, the contrast in *IACC* between Conditions 1 and 2 makes a telling comparison. The mirror chamber sky is square and completely unfurnished, and so the Condition 2 OBRIR envelope of Figure 1 shows a more regular sequence of reflections. The IACC peak in Condition 2 occurs when the HATS is pointing directly at one of the walls. This effect is emphasised by the position of the apparatus in the centre of the room plan, which should lead to very similar impulse responses at each ear at this orientation. In Figure 1, it can be observed that the left and right ears are approximate mirror images of each other over the 120º range of rotation. The built-in furniture of the toilet room evidently makes a more diffuse soundfield than that of the unfurnished mirror chamber sky.

Conditions 3 & 4: Listening room (125 m^3)

This is a rectangular room with a disproportionately high ceiling (4 m). (As part of a subsequent renovation to make it more suitable as a listening room, a suspended ceiling was installed.) Three of the four walls have drapes that can be drawn across most of their surface, allowing for simple adjustable acoustics. The wall with the door, which the HATS faced at 0º orientation, did not have drapes, and so is an important reflection source when the drapes are drawn. The room's reverberation time is 0.6 s with the drapes gathered, and 0.4 s with the drapes drawn (note that the room also had some other sound absorbing material in it when the measurements were made). The room also contained some furniture, although as the measurements were made on different days, the furniture positions may not have been the same in the two conditions.

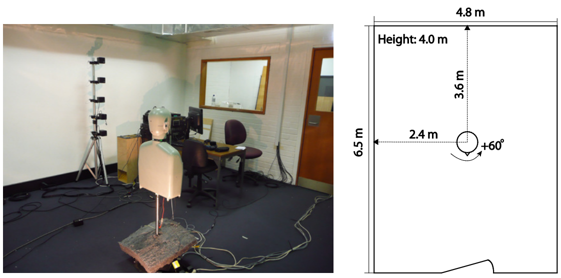

Figure 5. Photograph of the measurement apparatus in the listening room without curtains on the walls (left) and a plan showing the position and orientation of the apparatus (right).

Condition 3 has the drapes drawn across the walls, and Condition 4 has them gathered. Although drawing the drapes appreciably changes the sound, it does not cause a large change in the visualisation of the first 25 ms of the OBRIR in Figure 1. Early reflections from the furniture are weak, and the first wall reflection, which we could call the ITDG, is 12.8 ms after the direct sound (at its earliest). This side wall reflection is noticeably attenuated by the drapes. The ceiling reflection is seen almost 16 ms after the direct sound.

The parameters (Figure 6) represent the room condition in interesting ways. Gathering the drapes results in added G_{RG}, especially when the full integration period is used. The G_{RG} values are much lower than the small reverberant rooms (Conditions 1 and 2). The single bare wall results in a wide range of *ILD* values as the HATS turns, and gathering the drapes reduces the *ILD* range. *IACC* is not so responsive to the asymmetric reflection of the single bare wall, and merely decreases due to added reverberation when the drapes are gathered. Like many other rooms, there is an *IACC* peak corresponding to the HATS being aligned with the room plan, which is partly due to the first reflection from the wall facing the mouth arriving at the ears simultaneously.

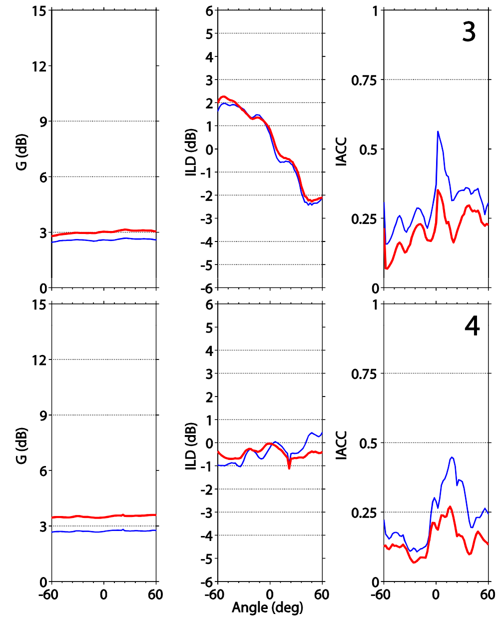

Figure 6. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the listening room, with curtains drawn (top) and gathered (bottom). The blue lines are for the early evaluation period, and the red (heavier) lines are for the full period.

Condition 5: Reverberant room (130 m^3)

This is a rectangular reverberant room which is part of the University of Sydney's acoustics laboratory. Although the room did not have diffusing panels in it at the time of measurement, it did contain various pieces of equipment unrelated to this project (most notably, two dodecahedral loudspeakers and two large public address loudspeakers). One of the dodecahedral loudspeakers was not far from the measurement position, as can be seen in Figure 7, producing weak early reflections. The door was slightly open for the measurement, which together with other absorptive elements in the room, reduced the mid-frequency reverberation time to 4.7 s.

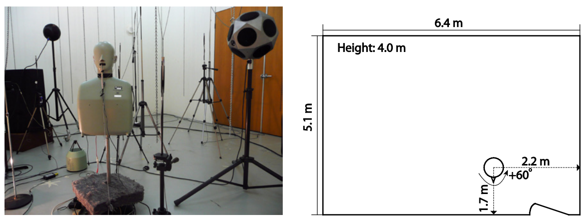

Figure 7. Photograph of the measurement apparatus in the reverberant room (left) and a plan showing the position and orientation of the apparatus (right).

The apparatus was positioned relatively close to two of the room walls. Being a simple room, larger than the previous ones described, the early reflection patterns are well separated in the first 25 ms (Figure 1).

The G_{RG} values in the reverberant room are high, with a large difference between early and full evaluation period values. Proximity to the two walls results in large ranges of early *ILD* and *IACC* (whereas the full range values vary little due to the strength and spatial diffusivity of the reverberant field). When the HATS's mouth is pointing directly towards the nearby wall, the first reflection at the two ears is highly correlated, contributing to a sharp increase in early *IACC*.

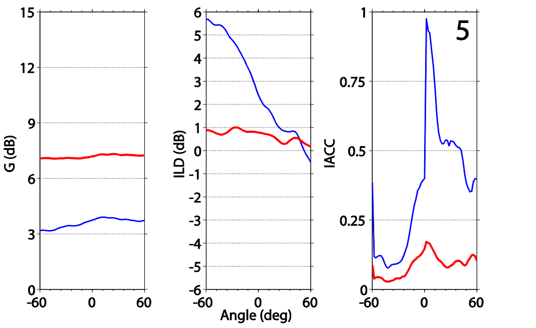

Figure 8. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the reverberant room. The blue lines are for the early evaluation period, and the red lines are for the full period.

Conditions 6 & 7: Sound recording studio control room (152 m^3) and recording room (170 m^3)

This sound recording studio has a control room and recording room adjacent to each other, and both rooms have relatively high sound absorption for their volume. The control room (Condition 6) has acoustic treatment on rear and side walls which is intended to have broadband sound absorption and to yield diffuse reflections. This appears to be borne out in the visualisation of the OBRIR in Figure 1, with a relatively high early reflection density. Although there are some minor reflections from the console furniture to the right of the HATS, the first major non-floor reflection is from the glass doors behind the HATS. These doors and the remainder of the room's three-section front wall are not treated for absorption and diffusion (since the studio loudspeakers are mounted flush with the front wall). The mid-frequency reverberation time was 0.35 s.

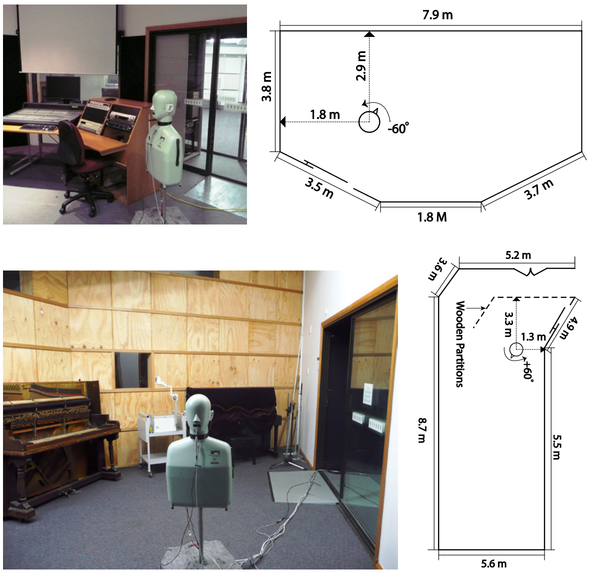

Figure 9. Photograph of the measurement apparatus in the studio control room (above left) and the studio recording room (below left), with respective plans (different scales) to the right of each photograph.

The recording room of the sound studio is larger than the control room, with a mid-frequency reverberation time of 0.4 s. It has a sloping ceiling, with absorptive and diffusive treatment on many of the walls. However, similar to Condition 6, the apparatus was positioned quite close to a simple wall, and this has a marked effect on the recorded OBRIRs. This wall reflection is seen in Figure 1, most prominently in the left ear at high angles of rotation.

In both rooms, the G_{RG} values are lower than the previously considered reverberant rooms, with a relatively small difference between early and full values. The presence of a strongly reflecting nearby surface, horizontally displaced from the measurement point, leads to a large range of *ILD* values as the HATS rotates. Greatest *ILD* occurs when the mouth and one ear are both directed towards the reflecting surface. *IACC* also varies considerably with angle of rotation, but in these cases the HATS does not strictly face the nearby reflective surface, and this is probably why a peak is not seen (like the peak in Condition 5).

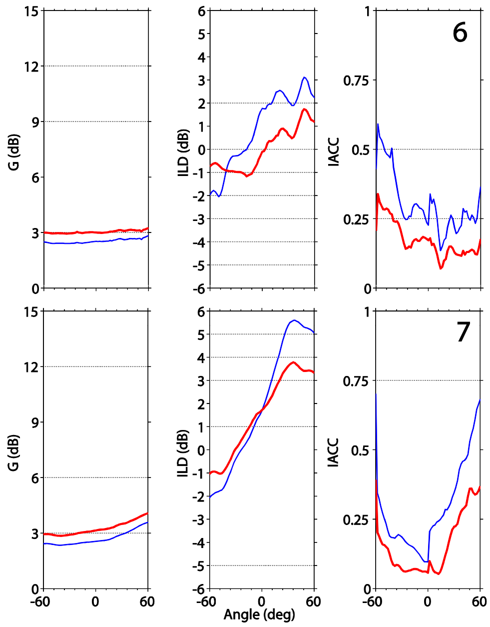

Figure 10. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the sound recording studio, with the control room in the upper chart, and the recording room in the lower chart. The blue lines are for the early evaluation period, and the red lines are for the full period.

Condition 8: Photometric laboratory (188 m^3)

This is a laboratory room containing various devices for photometric measurements, such as a photometric bench and an integrating sphere (on the left and right of the photograph respectively). The room is mainly hard-surfaced, with clutter around the walls and below the bench's shelf. The mid-frequency reverberation time is 0.9 s. Measurements were made in the least cluttered part of the room, as shown in Figure 12. The OBRIR envelope in Figure 1 exhibits a relatively high reflection density.

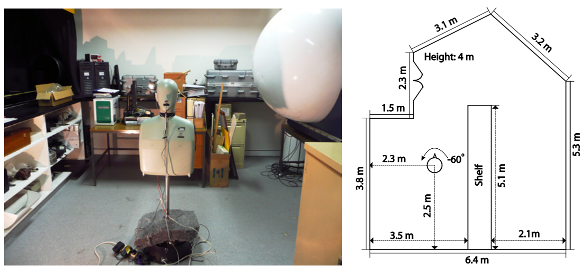

Figure 11. Photograph of the measurement apparatus in the photometric laboratory (left) and a plan showing the position and orientation of the apparatus (right).

Figure 12 shows the OBRIR parameters. As the HATS turns towards the corner near the door (which is occupied by filing cabinets, and so is effectively closer than shown in the plan) the early *ILD* increases due to a prominent early reflection to the left ear (about 10 ms after the direct sound). *IACC*, however, fluctuates in a way that is not so simple to interpret. Room gain is greater than that of the sound studio rooms, due to the longer reverberation time and the multitude of nearby reflective surfaces.

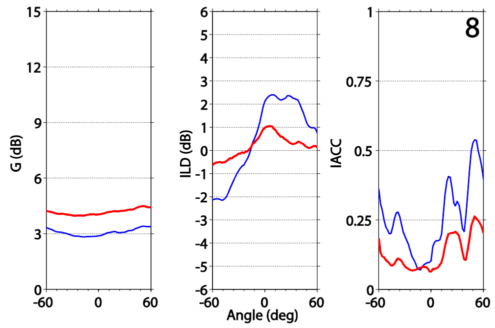

Figure 12. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the photometric laboratory. The blue lines are for the early evaluation period, and the red lines are for the full period.

Conditions 9 & 10: Lecture theatres (310 m^3 & 610 m^3)

These two lecture theatres seat 70 and 110 people (Conditions 9 and 10 respectively). Both are raked, and the smaller one has its rake steps in concentric circles (Figure 13). Reverberation times are 0.5 s and 0.6 s respectively. In both lecture theatres, the HATS was positioned in the middle of the front of the auditorium, perhaps as a lecturer would stand addressing the class.

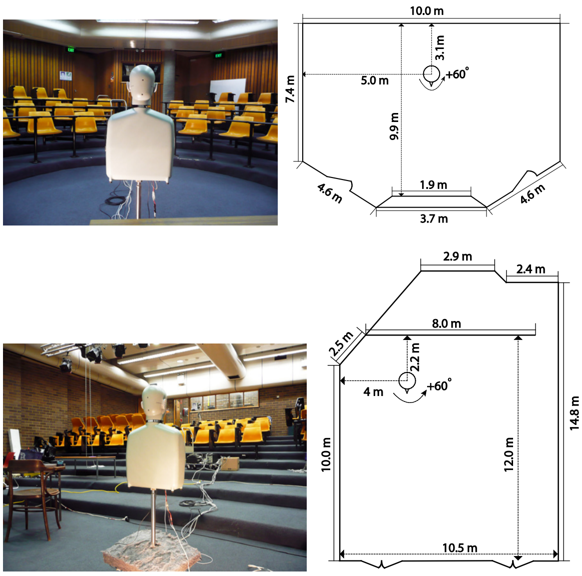

Figure 13. Photographs of the measurement apparatus in the small lecture theatre (above left) and the large lecture theatre (below left), with respective plans (different scale) to their right.

Room gain values are a little lower in the larger lecture theatre (Figure 14). In both, there is a relatively large range in *ILD*, consistent with the effect of the first reflection from the nearby walls (mainly the wall behind the HATS, but probably also the short angled wall to the right of the HATS in Condition 10). In the larger lecture theatre, there is a peak in *IACC* when the HATS aligns with the room geometry. In Figure 1 it can be observed that the left and right OBRIRs vary with angle roughly as mirror images of each other, which is consistent with the 0º *IACC* peak.

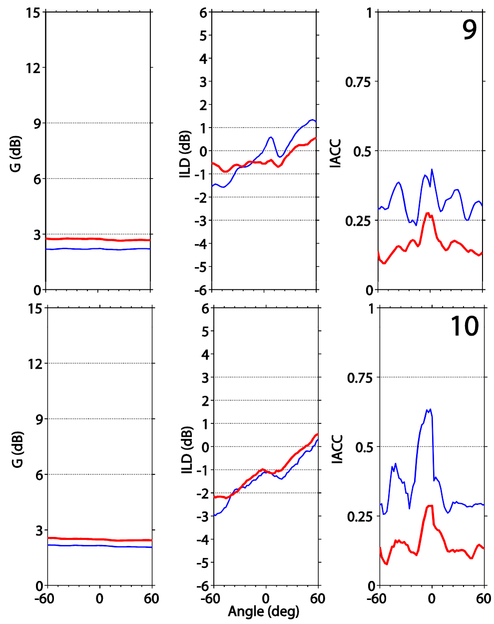

Figure 14. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the large lecture theatre. The blue lines are for the early evaluation period, and the red lines are for the full period.

Conditions 11 & 12: Recital hall (7650 m^3)

This is a music auditorium with a large stage and a mid-frequency reverberation time of 1.7 s on the day that we measured (it has adjustable reverberation time). We measured two positions on the stage of this auditorium: (i) downstage halfway across the stage width; and (ii) upstage, near the rear of the stage (Figure 15). Because a singer would almost never be seated if performing on stage, the HATS transducer height was adjusted to 1.5 m (while the same argument can be made for the lecture theatres, they were measured with a 1.2 m transducer height).

Being the room with the most absorption, these measurements yielded the lowest room gain (Figure 16), albeit still higher than might be found in a large auditorium (on a previous occasion, we measured stage support *ST1* value of -11.7 dB in the vicinity of the downstage position). It might be expected that placing the transducer near the stage walls would increase the $G_{RG,Early}$ values, but there was scarcely any increase found in our measurements. The increase from some directions is balanced by a reduction in sound energy from other directions, and a consideration with using a head and torso simulator is that the head shadow reduces the level at the far ear in the mid-high frequency range. The large range of *ILD* values upstage reflects this. The mid-stage position probably received some early support from a large piano about 2.5 m from the apparatus. *IACC* exhibits a prominent peak when the HATS is aligned with the auditorium in the downstage position.

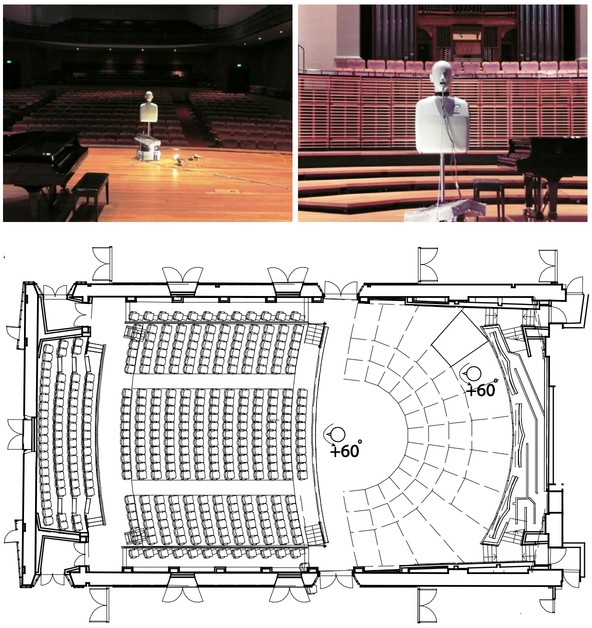

Figure 15. Photographs of the measurement apparatus in the recital hall downstage position, and a plan of the auditorium, indicating the two positions on stage.

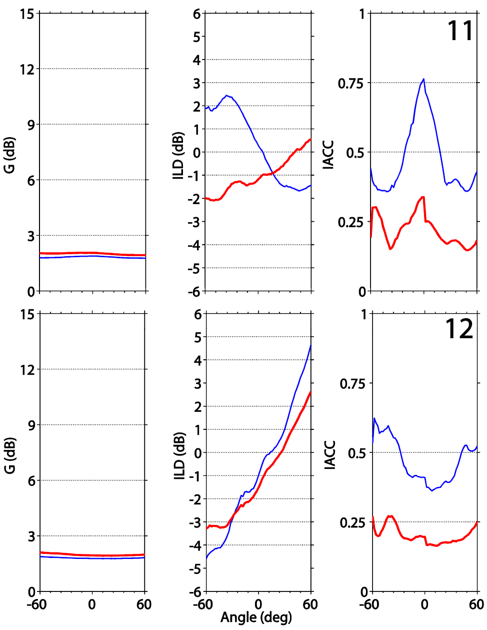

Figure 16. Values of G_{RG}, *ILD* and *IACC* as a function of angle in the recital hall, for the downstage position (top) and upstage position (bottom). The blue lines are for the early evaluation period, and the red lines are for the full period.

GENERAL OBSERVATIONS

Room gain is inversely related to room volume, which is evident in the summary given by Figure 17 (showing the median G_{RG} for each room condition, ordered from smallest to largest volume). There is a correlation of -0.57 between G_{RG} and the cube root of volume (which improves to -0.85 if the largest room volume is omitted, considering that it is much larger than the other rooms). Whether or not the largest room is included, the correlation between absorption and G_{RG} is -0.84 (with absorption estimated from the mid-frequency reverberation time and room volume). Typically the early G_{RG} is close to the full G_{RG}, but in highly reverberant rooms the difference is up to 4.5 dB. Hence, early G_{RG} also correlates well with the cube root of room volume, in fact somewhat better than the full version (r = -0.57 for all rooms, or -0.87 when the largest room is omitted). As the HATS rotates there is generally little variation in G_{RG} (the median range is 0.18 dB for early, and 0.21 dB for full G_{RG}). The notable exception is Condition 7 (1.2 dB for both early and late G_{RG}), where the HATS's speech simulator's directivity is apparent as it turns towards a nearby reflective wall in an absorptive room.

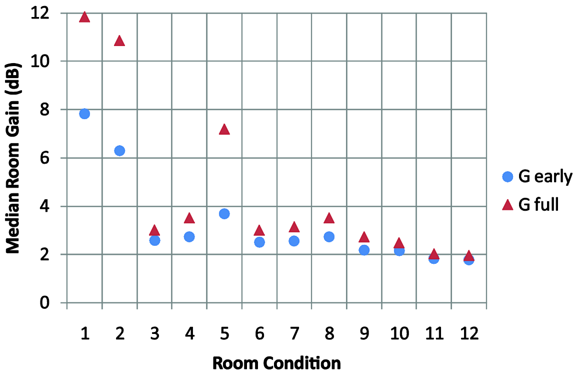

Figure 17. Median G_{RG} for each of the room conditions.

ILD, in itself, may be of interest in considering a particular orientation of the HATS in a particular room position, but if the HATS were to rotate a full circle, we would expect the median *ILD* to be near 0 dB regardless of the room condition. Therefore, *ILD* is considerably less interesting to us than the *range* of *ILD* values encountered as the HATS rotates. The range of *ILD* values over a rotation of the HATS is more sensitive to the horizontal spatial distribution of strongly reflective surfaces than the range of G_{RG} values, because the individual directivities of the ears (facing opposite directions) are accounted for, along with the directivity of the mouth. Figure 18 summarises the *ILD* range values of the twelve measured conditions. That a larger range is evident for early reflections (compared to the full evaluation period) is unsurprising, because the late reverberation is likely to be more diffuse than the early reflections (however, in Condition 3 there was very little late reverberant energy, and so the early and full *ILD* ranges are the same). Being relatively close to a wall (compared to the distance from other walls) is associated with large *ILD* ranges in Conditions 5, 6, 7 and 12. Considering that the *ILD* range is primarily determined by early reflections, the early evaluation range is probably more relevant than the full range.

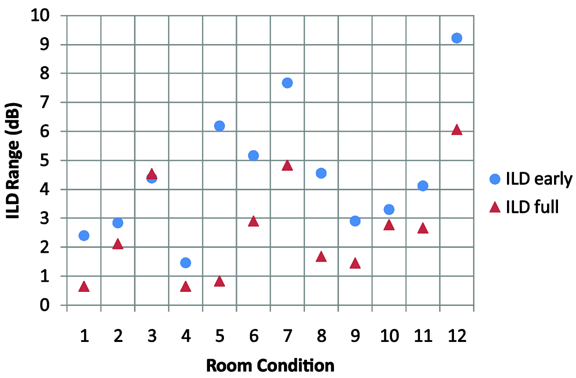

Figure 18. The range between maximum and minimum *ILD* for each of the room conditions.

Median *IACC* values are shown in Figure 19. The full period *IACC* values are consistently low relative to conventional measurements, which is unsurprising considering that the direct sound was omitted. The early *IACC* values vary more between the rooms, and so might be more useful in evaluating OBRIRs. The main factors that appear to influence the median early *IACC* of each room are the room size and surface diffusivity (high values are seen for large rooms with simple surfaces). There is little correlation between median *IACC* and the *ILD* range (refer to Table 1), with *IACC* better-correlated to room gain. It is probably to be expected that greater room gains would be associated with lower *IACC* values because the room gain represents the strength of the reverberant field.

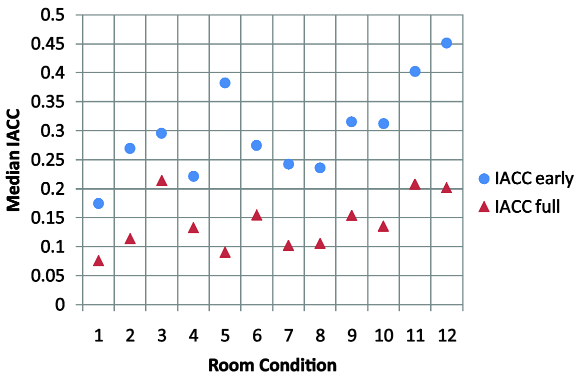

Figure 19. Median *IACC* for each of the room conditions.

In the individual room analyses we saw that early *IACC* tends to be very sensitive to HATS orientation in conditions with a strong early reflection, or where the HATS can align with room geometry. The highest early *IACC ranges*, seen in Figure 20, are examples of such conditions. It should be noted that in its conventional use (with an exocentric source, incorporating the direct sound), *IACC* has also been shown to be sensitive to head orientation [15]. Of course, in using range as an indicator of room acoustics (for *ILD* or *IACC*), a full 360º rotation of the HATS would be preferable to the rather arbitrary 120º range used here.

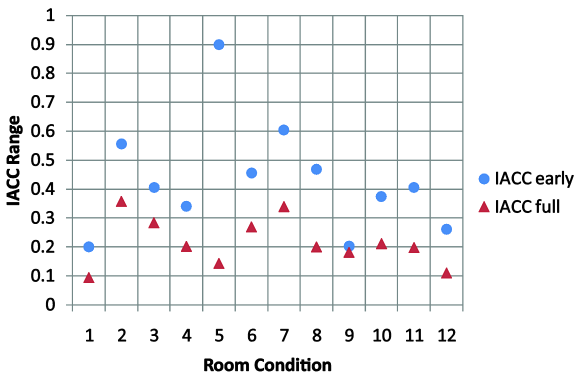

Figure 20. The range between maximum and minimum *IACC* for each of the room conditions.

Correlations between parameters are shown in Table 1. Correlations are generally weak, with the strongest being between *IACC* and G_{GR}, especially for the full evaluation period. The correlations between early and full measurements are all positive: G_{RG} 0.99; *ILD* range 0.76; *IACC* 0.63; and *IACC* range 0.37.

Table 1. Correlations between parameters.

Early	G_{RG}	***ILD* range**	***IACC***
***ILD* range**	-0.33		
IACC	-0.52	0.56	
***IACC* range**	0.02	0.32	0.15
Full	G_{RG}	***ILD* range**	***IACC***
***ILD* range**	-0.50		
IACC	-0.69	0.22	
***IACC* range**	0.04	-0.03	0.00

DISCUSSION AND CONCLUSIONS

In this study we have examined three potential parameters that could be used to characterise rooms based on oral-binaural room impulse responses (OBRIRs): room gain (G_{RG}), interaural level difference (*ILD*), and interaural cross correlation coefficient (*IACC*). The first of these was defined previously by Brunskog *et al*., and our survey shows how it varies between positions across a variety of rooms, from very small to large. G_{RG} represents the amount of room-reflected sound from the mouth to the ears, and so is perhaps self-evidently useful as a parameter. Brunskog *et al.* also show that it is one of the factors contributing to the vocal effort of teachers. However, the rooms in our survey include ones with very high and quite low G_{RG} values, and speaking is difficult at both extremes. Our measurements show how the difference between early and full evaluation periods for G_{RG} becomes larger in highly reverberant rooms. G_{RG} is related to gross physical and acoustical aspects of the rooms such as room volume and the rooms' sound absorption.

As OBRIRs are binaural, they are eminently suitable for examining interaural features. The interaural parameters (*ILD* and *IACC*) are somewhat speculative at present, as we still need support for their usefulness from subjective experimental data. However we have seen that they are sensitive, in different ways, to the room condition, and so they provide insight into the spatial distribution of reflections, which may contribute to the acoustic quality of the room for a talker. Clearly, interaural parameters are most relevant when reverberation is most audible – namely in rooms with high room gain or long reverberation time: and therefore interaural parameters need to be considered alongside the question of their salience. Early *ILD* varies greatly when there is a prominent strong early reflection. Early *IACC* peaks when the HATS is aligned with a simple room geometry, or else when it is directly facing close wall. These variations in interaural parameters with HATS rotation provide support to the notion that head tracking can helpfully contribute to OBRIR simulation systems.

It is easy to informally check the apparent significance of interaural variation in OBRIRs for oneself, by listening to one's own speech in real environments whilst deliberately repositioning (e.g., rotating) one's head (and possibly the torso too). In some situations, changes in the sound are clearly evident with body rotation, but generally the changes are subtle. Positions with prominent reflections from a particular direction – like those Conditions where *ILD* was observed to change greatly – are among the most striking. However, it is likely that these changes in mouth to ear transfer function are generally suppressed as one listens to one's own speech, and instead a holistic impression of the room acoustical environment is constructed at a subconscious level (the phenomenon of reverberation suppression certainly occurs in listening to exocentric speech [16]). It should also be borne in mind that a talker's head is constantly moving, and so there may be little opportunity to fully hear the effect of a single time-invariant OBRIR in normal speech. Hence, in situations where OBRIR parameters are highly sensitive to angle (such as the *IACC* peak observed with apparatus alignment with the geometry of some rooms), it would be reckless to draw the inference that the auditory experience of the room fluctuates greatly with small head movements in natural speech at the measurement position.

The acoustic quality of rooms for talkers may also depend on gross spectral features – i.e., the balance of high to low spectral components in the transfer function. In auditorium acoustics, bass ratio and treble ratio were designed to account for such aspects of acoustic quality [16]. We investigated this in the present set of measurements, using the power spectral centroid of the octave band G_{RG} values, but found a high positive correlation with G_{RG} itself (0.96 for early, and 0.93 for late periods), and so we have not included these values in the paper. That is, the larger rooms had lower G_{RG} and a duller sound.

The survey of rooms in this paper is limited, with only ten rooms tested, and one position tested within most rooms. Furthermore, the HATS was rotated over 120º, rather than 360º in each measurement situation. Nevertheless, the quantity of data is considerable, and the results are sufficient to gain insight into some aspects of the parameters examined.

REFERENCES

1 A.C. Gade, "Investigations of musicians' room acoustics conditions in concert halls, Part1: Methods and laboratory experiments," *Acustica* **68**, 193-203 (1989)
2 ISO 3382-1 Acoustics – measurement of room acoustics parameters – Part 1: Performance spaces (2009)
3 J. Brunskog, A.C. Gade, G. Payá Bellester, L. Reig Calbo, "Increase in voice level and speaker comfort in lecture rooms," *J. Acoust. Soc. Am.* **125**, 2072-2082 (2009)
4 D. Cabrera, H. Sato, W.L. Martens and D. Lee, "Binaural measurement and simulation of the room acoustical response from a person's mouth to their ears," *Acoustics Australia* **37**, 98-103 (2009)
5 D. Cabrera, R. McCarthy, J. Bassett and D. Gilfillan, "The measurement of directional stage support in a drama theatre: Quantifying the effect of a set," *Building Acoustics* (in press)
6 W.T. Chu and A.C.C. Warnock, *Detailed Directivity of Sound Fields around Human Talkers* (Inst. for Res. in Construction, National Res. Council Canada, Ottawa, ON, Canada, Tech. Rep. IRC-RR-104, 2002)
7 D. Cabrera, P.J. Davis, A. Connolly, "Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment," *Journal of Voice* (in press)
8 C. Pörschmann, "One's own voice in auditory virtual environments" *Acustica* **87** pp. 378-388 (2001)
9 H. Sato, M. Morimoto and K. Fukunaga, "Effects of reverberation sounds on conversing difficulty in living rooms" *Tech. Rep. Architectural acoustics Acoust. Soc. Jpn.* No. AA2008-52 (2008) pp. 1-8 (in Japanese)
10 J. Blauert, *Spatial Hearing* (MIT Press, Cambridge MA, 1997)
11 T.A. Stroffregen and J.B. Pittenger, "Human echolocation as a basic form of perception and action," *Ecological Psychology* **7(3)**, 181-216 (1995)
12 P. Zahorik, D.S. Brungart and A.W. Brokhorst, "Auditory distance perception in humans: A summary of past and present research," *Acustica* **91**, 409–420 (2005)
13 R. McGrath, T. Waldmann and M. Fernström, "Listening to rooms and objects" *Proc.16th Audio Eng. Soc. Int. Conf.*, Rovaniemi, Finland (1996)
14 J. Bradley, "Predictors of speech intelligibility in rooms," *J. Acoust. Soc. Am.* **88**, 837-854 (1986)
15 I. Witew, A. Lindau, J. van Dorp Schuitman, M. Vorländer, S. Weinzierl and D. De Vries, "Uncertainties of IACC related to dummy head orientation, *Fortschritte der Akustik: Tagungsband* d. **36**. DAGA. Berlin (2010)
16 P. Zahorik, E. Brandewie and V.P. Sivonen, "Spatial hearing in reverberant rooms and effects of prior listening exposure," *International Workshop on the Principles and Applications of Spatial Hearing*, Zao, Japan (2009)
17 L.L. Beranek, *Music, Acoustics and Architecture* (John Wiley & Sons, New York, London, 1962)