Variations in acoustical parameters in oral-binaural room impulse responses of a real and a computer-modelled room

Manuj Yadav, Luis Miranda, Densil Cabrera and William L. Martens

Faculty of Architecture, Design and Planning, The University of Sydney, NSW, 2006, Australia

ABSTRACT

Auralizing computer-modelled rooms can be accomplished with relative ease using room acoustical software such as Odeon, Ease etc., when the source(s) and the receiver(s) are located separately in space. However, the same task is not so straightforward when the source (human mouth) and receiver (two ears) are concentric, e.g., within the same head, which can enable the simulation of sound that one hears from one's own voice in different room environments. Previously, studies have employed humans, or more generally head and torso simulators for obtaining impulse responses from the mouth to the two ears of the same head in real rooms, referred to as an oral-binaural room impulse response (OBRIR). Measuring OBRIRs for more than one room and for different orientations with the same room can be very time-consuming and cumbersome, which can be a limitation in studies that require a large number of rooms with modular features. The present paper is addressing this issue with a preliminary study of the variations in the acoustical parameters derived from OBRIRs of a real room and a computer model of the same room, where important room modeling issues are highlighted for future studies.

INTRODUCTION

The characterization of the sound that reaches a performer, in the form of stage support (Gade 1989), has been well established as a measure that is used in the design and evaluation of large performance spaces (ISO 3382-1 2009), such as music auditoria. The study by Gade focussed on the performer and not the listeners in the audience area.

Recent research, however, has pointed out a few inadequacies in the utility of stage support as a parameter for smaller rooms, such as lecture theatres, everyday rooms, etc. (Brunskog *et al*. 2009). The inadequacies were related to the derivation of stage support from the room impulse response measurements. As stage support is measured as the fraction of energy in the impulse response later than 20 ms relative to the direct sound, this evaluation range is insufficient for small rooms where the direct sound and the first room reflections are usually more closely spaced. Also, as a distance of 1 m separates the source and the receiver in a stage support measurement, it limits the use of the parameter for cases where the source and the receiver are closer, e.g., when the source and the receiver are the voice and ears of the same (human or dummy) head.

An oral-binaural room impulse response (OBRIR) describes the room acoustical response from the mouth to the two ears of the same head. For a real room measurement, OBRIRs are calculated from the acoustic transfer function from the mouth loudspeaker to the two small microphones placed at the entrance of the ear canal of a head and torso simulator (HATS; Brüel & Kjær 4128C) (Cabrera *et al*. 2009). An OBRIR enables the characterization of parameters that relate to the sound that the listener hears from his/her own voice in a room, and is well suited for smaller volumes than stage support (Cabrera *et al.* 2011). Another possibility with OBRIRs is the simulation of autophony within different room acoustic conditions (measured with OBRIRs) in an anechoic room, which was implemented with real-time convolution solutions (Cabrera *et al.* 2009; Yadav *et al.* 2012a). However, the OBRIR measurement method for real rooms is a cumbersome exercise (Cabrera *et al.* 2009).

An alternative involves using computer-modelled rooms for deriving the OBRIRs for a range of head positions and orientations. The authors have described elsewhere how this could be accomplished within a room acoustics software (ODEON used in the paper), using higher-order ambisonics (HOA) to approximate the directivity of the HATS for a range of angles (Yadav *et al.* 2012b). This paper presents a preliminary study of the variation in the long-term characteristics of the OBRIRs of a real room and the respective computer-model of the same room based on the suggested impulse responses parameters presented in ISO 3382-1 (2009). The relevance of the parameters presented in this standard is that they relate subjective parameters to acoustic parameters based on impulse response measurements. The long-term characteristics explored in this paper are currently better understood in contrast to the fine structure characteristics of the impulse response. The fine structure characteristics, particularly of the early part of the impulse response, could have important implications, for example, in terms detecting gross variations in nearby objects. While the authors acknowledge this, the scope of this paper is limited to the long-term characteristics of the impulse response.

For the remainder of the paper, *measured OBRIRs* refers to the OBRIRs that were measured with a HATS in a real room, while *simulated OBRIRs* refer to the OBRIRs obtained from the computer model of the same room in ODEON using HOA for simulating the directivty of the HATS.

Within the long-term characteristics, two types of variations are studied- the varation within the measured and the simulated OBRIRs for yaw angles in the horizontal plane ranging from -60° to 60°. The variation gives an indication of whether the change in parameters between the various angles, from a reference position of 0° azimuth, would be noticible by humans. This is done by plotting the variation alongside the just-noticable differences (JNDs) for the parameters, as specified in ISO 3382-1 (2009).

The room that was used for generating the data used in this paper is a medium-sized lecture theatre (ALT1; volume = 610 m^3; $T20_{mid\text{-}frequency}$ = 0.6 s) in the Faculty of Architecture, Design and Planning, The University of Sydney. A model of the room, as used in ODEON, is presented in Figure 1. The location of the HATS, when the OBRIR measurement was done in the real room is presented in Figure 2.

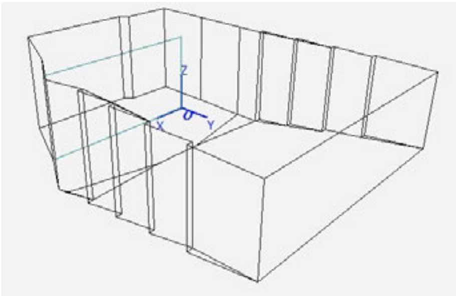

Figure 1. The model of the room used in ODEON

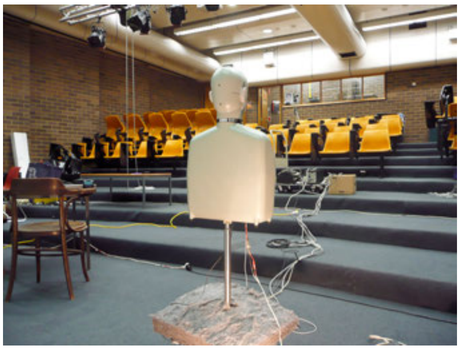

Figure 2. The location of the HATS in the real room used in the OBRIR measurement. The same location was used in the computer model in ODEON with the HATS mouth height of 1.5 m from the floor.

VARIATION OF ACOUSTICAL PARAMETERS IN REAL AND SIMULATED OBRIRS

Variation of acoustical parameters in real OBRIRs

In most normal rooms and with most acoustic sources (i.e. non-regular rooms and non-omnidirectional sources), we can expect variations in the decay of sound in a room based on the directional characteristics of the source exciting the room and the room configuration. These variations would occur even if the source were rotating along one axis, as is the case in the measured OBRIRs. Theoretically, this type of variations would be experienced by a talking-listener in a room and would be included in the overall acoustic perception of a room. This has also been shown to be valid in a subjective study using the measured OBRIRs in a head-tracked room acoustic simulation system (Yadav *et al.* 2011).

It should be noted that the calculation of the parameters in this paper do not take into account the direct sound (including the first floor reflection, which amounts to roughly 8 ms from the beginning of the impulse response) measured using the setups described previously (both measured and simulated). The direct sound does not vary between different head orientations and different rooms when the source and the receiver are concentric, as in OBRIR measurements. Morever, as we are only interested in variations from a reference measurement, we can assume that the direct sound will be constant in all measurements and removing it from the calculations will not yield any deviations from the measurement variations. The parameters whose variations were studied, were chosen on the basis of their relevance in characterising important subjective attributes of autophony in rooms, and are as follows:

1. EDT: To describe reverberation time, the parameter Early Decay Time (EDT) was used. EDT is defined as the time taken for the sound to decay to a level of -60dB, based on a 10dB interval and extrapolated to the 60dB decay point. This parameter was chosen as ISO 3382-1 (2009) regards this parameter as being strongly correlated to the subjective impression of reverberation. The single number values of EDT were obtained by averaging the EDT for the two ear for the 500 and 1000 Hz bands.

The JND value for EDT, where a human listener will experience a difference in perceived reverberation, is a 5% deviation from the reference value. The deviations of EDT (in percentage) from the reference, for the measured OBRIRs are depicted in Figure 3.

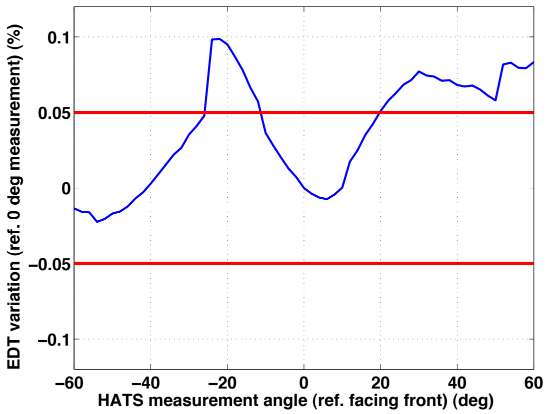

Figure 3. EDT variation for the measured OBRIRs in blue as a percentage of deviation from the 0 degree measurement. JNDs in two directions are shown in red.

2. Clarity Index (C_{80}): Clarity index evaluates the amount of late energy compared to the amount of early energy, as seen in Equation 1.

$$C_{80} = 10\log\frac{\int_0^{80ms} h^2(t)\,dt}{\int_{80ms}^{\infty} h^2(t)\,dt}\,\mathrm{dB} \qquad (1)$$

Here, $h(t)$ is the instantaneous sound pressure of the OBRIR. C_{80} is closely related to the subjective impression of clarity, and is the ratio of the energy in the first 80ms compared to the energy from 80ms to the end of the impulse response, expressed in dB. As with EDT, in order to quantify the variation of late to early energy of the measured OBRIRs, the direct sound (8 ms including the first floor reflection) is assumed to remain constant through all measurements, and is removed from the evalution range. As a result, the time window of the early sound window is shortened in the C_{80} calculation. This allows keeping the temporal windowing for the C_{80} calculations and providing the means to depict the variations in early to late energy across measurements. The JND for C_{80} as it appears in ISO 3382-1 (2009) is 1dB. The single number C_{80} value was obtained by averaging the C_{80} values for the two ears for the 500 and 1000 Hz frequency bands. Variations of C_{80} for the measured OBRIRs are presented in Figure 4.

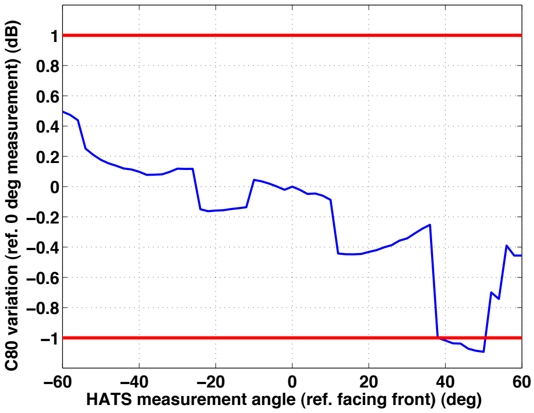

Figure 4. C_{80} variation for the measured OBRIRs in blue expressed as the difference in dB from the 0 degree measurement. JNDs in two directions are shown in red.

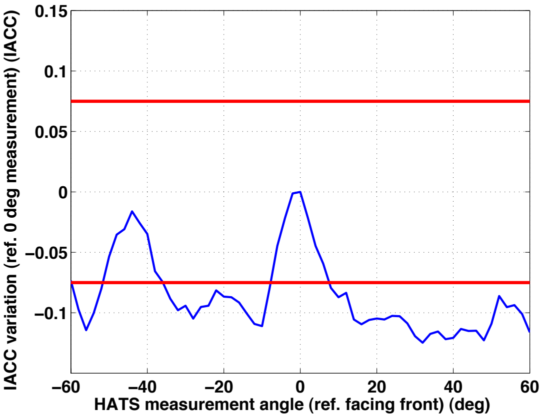

Figure 5. IACC variation for the measured OBRIRs in blue express as the difference in IACC from the 0 degree measurement. JNDs in two directions are shown in red.

4. Interaural cross correlation (*IACC*): The equation used for deriving the value of *IACC* in this paper is as described in ISO 3382-1 (2009), except the exclusion of the 8 ms of an OBRIR due to the reasons outlined above, and is expressed as:

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

Here, 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 the left and right ears, which correspond to the respective ear microphones that are placed at the entrance of the ear canal of a HATS. IACC describes the degree of similarity of two signals arriving at the two ears. IACC ranges from 1, where the signals are exactly similar, to 0, where the signals are completely uncorrelated. The JND for IACC suggested in ISO 3382-1 (2009) is a deviation of 0.075 from the reference IACC. Variations in IACC as the HATS rotates are shown in Figure 5 with the 0° measurement as reference.

Variations of acoustical parameters in simulated OBRIRs

Approaching similar values of acoustical parameters between a simulated room and measurements obtained from a real room could simply be matched via an iteration process. This iteration process would require changing room surface finishes in the simulated room until the reverb times closely match the measured values. While this process can be time consuming, it is relatively straightforward. Instead of aiming for similarity in the acoustical parameters, this paper is focussing on exploring the variations that occur in the simulated impulse responses as the source and receiver are rotated on the horizontal plane in the simulated space. These variations might be caused by several factors including the room configuration, the HRIR interpolation and the choice of filters used to broaden the frequency content of discrete reflections that are provided by ODEON and further processed with HOA in Matlab (Yadav *et al.* 2012b). For this paper, the simulated OBRIRs are created using the method described in Yadav *et al.* (2012) for room presented in Figure 1 and the position presented in Figure 2, to correspond to the measured OBRIRs.

The observed variations in EDT, C_{80} and IACC are presented in Figures 6, 7 and 8. The single number values of EDT and C_{80} are obtained by averaging the measurement of the two ears for the 500 and 1000 Hz octave bands, and averaging the 500 Hz and 1000 Hz octave bands at the two ears for the IACC.

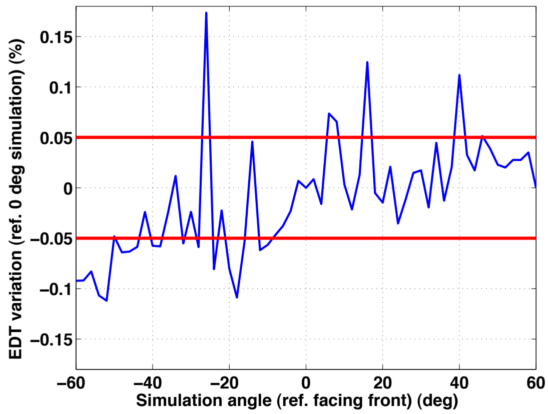

Figure 6. EDT variation for the simulated OBRIRs in blue as a percentage of deviation from the 0 degree measurement. JNDs in two directions are shown in red.

DISCUSSION

From inspecting Figures 3-8, deviations up to and beyond the JND is noticed for all the parameters, as the HATS is rotated. While this is to be expected in the IACC as it is very sensitive to the spatial situation of the receiver (de Vries *et al.*

2001), it is interesting to see these variations in the parameters describing the energy decay in the room. A variation in the energy parameters suggests that talking-listeners are likely to notice these variations.

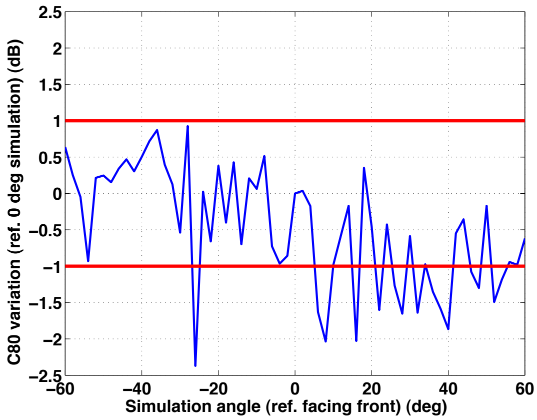

Figure 7. C_{80} variation for the simulated OBRIRs in blue expressed as the difference in dB from the 0 degree measurement. JNDs in two directions are shown in red.

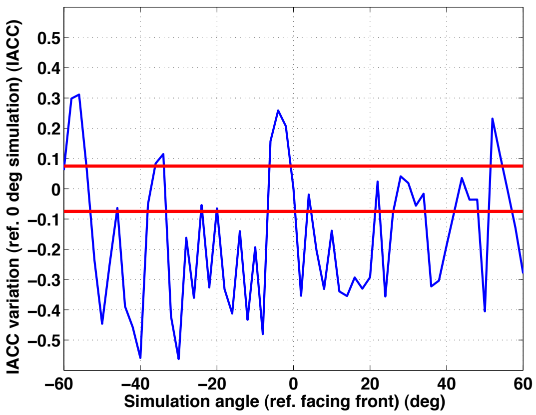

Figure 8. IACC variation for the simulated OBRIRs in blue express as the difference in IACC from the 0 degree measurement. JNDs in two directions are shown in red.

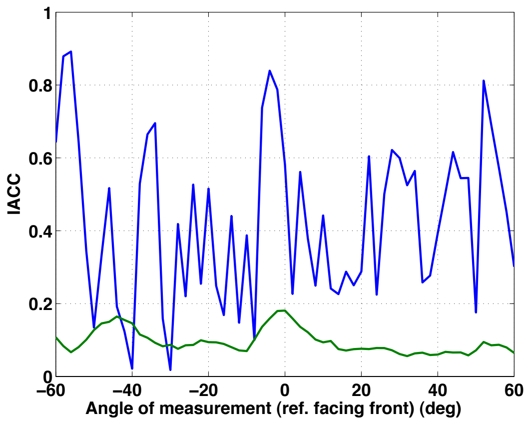

Figure 9. IACC of the measured room in green and simulated room in blue.

This implies that a study of these parameters could be useful when subjective responses are analysed for correlations with objective parameters.

The variations in EDT and C_{80} of the simulated OBRIRs are larger than the variations in real rooms in all measured parameters. In the case of EDT the variations are as large as 17% compared to variations of up to 10%. In the case of C_{80}, variations of over 2dB are observed compared to variations of just over 1dB. This could be attributed to the factors mentioned above, namely, the HRIR interpolation and the choice of filters used to broaden the frequency content of discrete reflections of the simulated OBRIRs. Currently, the authors are exploring the possibilities of aligning the variations in both the real and the simulated rooms.

Furthermore, while there are noticeable variations in C_{80} and EDT, they are not as pronounced as the variations observed in IACC. When the IACC data of the original and simulated rooms are compared (Figure 9), it is noticeable that the IACC in the original room is very low. These extremely low values (see e.g. de Vries *et al.* (2001) for values recorded in a concert hall) at all locations point at possible problems in specifying the surface scattering coefficients. If we look at the obtained values for the measured and simulated room, presented in Figure 9, we will see that the measured room overall has lower IACC values. In a room where a diffuse field dominates the impulse response, we can expect lower IACC values as the energy arrival behaves randomly.

CONCLUSIONS

The variations in the acoustical parameters in the present paper are promising in terms of displaying the possibility to incorporate simulated OBRIRs in subjective experiments. If these variations were noticible by human participants, the process of acquiring OBRIRs for computer-modelled rooms in which autophony is simulated would be simplified to a great extent. Creating rooms that closely match the variations experienced in real rooms is an important step towards this goal.

The analysis presented in this paper also points out at the importance of considering surface scattering in the room simulation process. To extend the validity of this analysis method, a further study is being conducted by increasing the number of rooms under investigation.

ACKNOWLEDGEMENTS

This research is supported by an Australian Postgraduate Award for Manuj Yadav and Luis Miranda and a CONACYT (a Mexican government initiative) scholarship from the Becas al Extranjero program for Luis Miranda. The travel award from the NSW division of AAS covered the expenses related to attending the conference for Manuj Yadav.

REFERENCES

Brunskog, J, Gade, AC, Bellester, GP & Calbo, LR 2009, 'Increase in voice level and speaker comfort in lecture rooms', *The Journal of the Acoustical Society of America*, vol. 125, no. 4, pp.2072–2082.

Cabrera, D, Sato, H, Martens, WL & Lee, D 2009, 'Binaural measurement and simulation of the room acoustical response from a person's mouth to their ears', *Acoustics Australia*, vol. 37, no. 3, pp.98–103.

Cabrera, D, Lee, D, Collins, R, Hartmann, B, Martens WL & Sato, H 2011, 'Variation in Oral-Binaural Room Impulse

Responses for Horizontal Rotations of a Head and Torso Simulator', *Building Acoustics,* vol. 18, no. 1, p.277.

De Vries, D, Hulsebos, EM & Baan, J 2001, 'Spatial fluctuation in measures for spaciousness', *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp – 947-954.

Gade, AC, 1989, 'Investigations of Musicians' Room Acoustic Conditions in Concert Halls. Part I: Methods and Laboratory Experiments', *Acustica*, vol. 69, pp.193–203.

International Organization for Standardization 2009, *Acoustics –Measurement of room acoustic parameter- Part1: Performace spaces*, ISO 3382-1:2009, International Organization for Standardization, Geneva.

Yadav, M, Cabrera, D & Martens, WL 2012a, 'A system for simulating room acoustical environments for one's own voice', *Applied Acoustics*, vol. 73, no. 4, pp.409–414.

Yadav, M, Miranda, L, Cabrera, D & Martens, WL 2012b, 'Simulating autophony with auralized oral-binaural room impulse responses', *133rd Convention of the Audio Engineering Society*. San Francisco, CA, USA.

Yadav, M, Cabrera D, Collins, R & Martens WL 2011, 'Detection of headtracking in room acoustic simulations for one's own voice', *Proceedings of the Australian Acoustical Society*. Gold Coast, Australia.