Proceedings of the International Symposium on Room Acoustics, ISRA 2010

29-31 August 2010, Melbourne, Australia

Equal reverberance contours for synthetic room impulse responses listened to directly: Evaluation of reverberance in terms of loudness decay parameters

Doheon Lee, Densil Cabrera and William L. Martens

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

PACS: 43.66.Lj, 43.55.Br

ABSTRACT

This paper examines effects of listening level and reverberation time on the perceived decay rate of synthetic room impulse responses (RIRs). A listening test was conducted with synthetic RIRs having a range of listening levels and reverberation times: in the test, subjects adjusted a physical decay rate of the RIRs to match the perceived decay rate of reference stimuli. In this way, we constructed equal reverberance contours as a function of sound pressure level and reverberation time. The experiment results confirm that listening level and reverberation time both significantly affect reverberance. The study also supports our previous findings: that the loudness decay function in can be used to predict reverberance better than the conventional reverberance predictors.

INTRODUCTION

Reverberance, which in general terms refers to the subjective impression of reverberation, is usually assessed using early decay time (EDT) [1]. Atal *et al*. [2] proposed EDT following the study of Haas *et al*. [3], which found that early reflections are important in the human perception of sound. According to Soulodre and Bradley [4], EDT correlates with the reverberance of a tested music stimulus better than reverberation time (T). Despite the importance of reverberance in room acoustics, since the study of Atal *et al*. [2], there has not been much work on developing a new or improved reverberance predictor. Recent work by the authors [5, 6, 7] has found that listening level significantly affects the reverberance of music stimuli and impulsive stimuli: greater level is associated with greater reverberance. This effect of stimulus level on reverberance is also supported by Hase *et al*. [8]. Their study shows that sound pressure level and reverberation time independently affect the reverberance of music and speech stimuli. Moreover, in their experiment, sound pressure level had a stronger effect on reverberance than reverberation time.

As described by Zwicker and Fastl [9], the human auditory system is not simple and the physical sound pressure is not sufficient to explain the subjective perception of sound strength (namely, loudness perception). For example, two sounds with the same weighted sound pressure level may differ in loudness due to their spectral content (e.g., white noise tends to be perceived louder than a pure tone, and pure tones at mid-frequency tend to be perceived louder than pure tones at low-frequency). In the case of two tones of the one frequency and power, but different durations, the one having a longer duration is perceived louder than the one having a shorter duration, up to certain duration. Hence, in order to simulate the human perception of sound, these factors (which are related to the effects of auditory filtering and temporal integration) need to be carefully considered, in addition to other important factors such as spectral masking and the functions relating auditory excitation to specific loudness and so forth. When sound fluctuates over time, it is more complex to predict its loudness because the level and frequency content of sound at a particular time may strongly affect the loudness at a subsequent time.

Despite of these complexities, the objective loudness models such as the Dynamic Loudness Model of Chalupper and Fastl [10] and the Time-varying Loudness Model of Glasberg and Moore [11] effectively predict the loudness of both stationary and non-stationary sounds. According to Chalupper and Fastl [10], the Dynamic Loudness Model provides a good match with the results of psychoacoustic experiments for stationary sounds having a range of levels, bandwidth and durations. The discrepancies between the model predictions and the psychoacoustic data are mostly within the quartiles of the psychoacoustic data. With respect to non-stationary sounds, the authors represent that the model predictions are not entirely matched with a couple of previous studies [12, 13]. However good agreement is observed with an experiment by Grimm [14], which tested the level required for sinusoidally modulated sounds to sound equally loud as unmodulated sounds. For the Time-varying Loudness, the authors who developed the model remarked that the model outputs well accord with relevant psychoacoustic data of previous studies, although details of the comparisons were not provided in [11]. A detailed comparison of the two models has been made by Rennies *et al.* [15].

Hence previous studies of Lee and Cabrera [16, 5, 6] and of Lee *et al*. [7] tested the idea of using the loudness decay function, derived from such models, in deriving a reverberance predictor. The underlying concept is that reverberance should be related to the modelled loudness decay rate, as this aims to approximate what people hear rather than representing the physical decay of sound. The purpose of this approach is not to develop yet another room acoustical parameter, but instead is to better explain the concept of reverberance using a model based on human perception. (Indeed it would be impractical to apply the loudness-based reverberance parameters to auditorium qualification and design problems.) The fact that the loudness decay function is approximately exponential (at first) is helpful because we can define loudness-based parameters that are analogous to conventional decay parameters such as EDT and T. Figure 1 shows an example of the loudness decay functions derived from a room impulse response (RIR) having a range of L_{AFmax} levels (from 50 dBA to 80 dBA), derived from the Dynamic Loudness Model. As seen in the figure, the slope of the loudness decay functions varies with the L_{AFmax} value: as the RIR is louder (i.e., greater gain), the slope becomes less steep. By contrast, although it is not illustrated, the slope of the sound pressure level decay functions for the same RIR is independent of gain.

The loudness-based parameters were termed EDT_N and T_N (as the subscript '*N*' stands for loudness) depending on their evaluation range. According to Stevens [17], loudness is proportional to sound pressure raised to a power of 0.6 for tones of moderate frequency and moderate sound pressure level (this is consistent with the well-known rule-of-thumb that doubling or halving loudness corresponds to ±10 dB). Hence the EDT_N was calculated by measuring the time taken for a linear regression line of the loudness decay function from the peak loudness to half of the peak loudness, multiplied by 6. This evaluation range corresponds to the evaluation range of the conventional EDT. Like EDT_N, the T_N was calculated by measuring the time taken for a linear regression line of the loudness decay function over 0.708 of the peak loudness to 0.178 of the peak loudness, multiplied by 3. The evaluation range also corresponds to the evaluation range of the conventional T20. An example of the T_N calculation for a RIR that has a conventional $T20_{mid}$ of 2 s and a L_{AFmax} of 80 dBA is shown in Figure 1. In calculation of the loudness decay function, we used the Dynamic Loudness Model by Chalupper and Fastl [10] or the Time-Varying Loudness Model by Glasberg and Moore [11], both of which are implemented in PsySound3 [18]. The performance of the two models was similarly good.

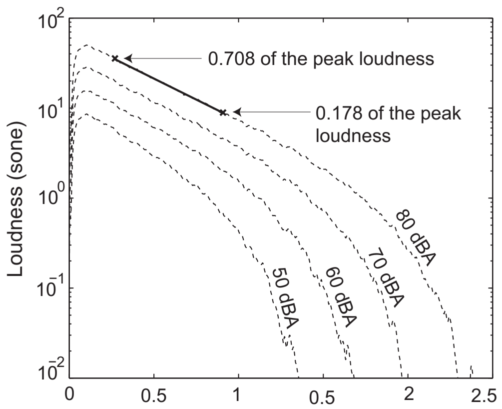

Figure 1. Loudness decay functions (dotted lines) of RIRs having a L_{Afmax} level of 50 dBA, 60 dBA, 70 dBA and 80 dBA. The straight line is a linear regression line between 0.708 and 0.178 of the peak loudness

Performance of the loudness-based parameters was substantially better than conventional parameters in predicting subjectively matched reverberance. Both T_N and EDT_N were better predictors of reverberance for both music stimuli and an impulsive stimulus than the conventional parameters [5, 6, 7]. For the reverberance of music, the parameters were tested for both overall reverberance and running reverberance. According to Morimoto and Asaoka [19], the reverberance of music is categorized into two parts; (1) running reverberance and (2) stopped (or terminal) reverberance. The former refers to the reverberance given while a stimulus is running and the latter refers to the reverberance after a stimulus is stopped. As there are few opportunities to hear stopped reverberance when audience is listening to music (except when there are large temporal gaps between notes), our previous studies of music did not test stopped reverberance (although impulsive reverberance is similar). T_N has a relatively long evaluation range, which proved to be less suitable than EDT_N in assessing the reverbeance of a running stimulus, and this relates to the rationale for conventional EDT [7]. Lee and Cabrera [6] explored the reverberance of an impulsive stimulus (specifically, the perceived decay rate of RIRs listened to directly) in relation to a level variation of ±5 dB, using RIRs with a small range of reverberation times (2.0 s to 2.7 s). The RIRs in that experiment were recorded from real auditoria, and so had natural irregularities in them, including gross features such as a L_{AFmax} level ranging from 70 dBA to 75 dBA (before applying additional gain of ±5 dB), as well as frequency-dependent decay rates that were not well-controlled.

Hence the present study tests the perceived decay rate of a synthetic RIR, when it is directly listened to (rather being than convolved with anechoic signals), over a wide range of reverberation times and sound pressure levels, so that equal reverberance contours can be derived from the experiment results. The experiment results were also converted into both the loudness-based parameters and the conventional parameters so as to find the best predictor of the perceived decay rate over the wider range of reverberation times and sound pressure levels of previous studies. The details of the experiment are described in the following section.

METHOD

The synthetic RIRs were generated from white noise signals filtered into the ten octave bands centred on 32 Hz to 16 kHz. In order to make the synthetic RIRs more realistic, the decay rate of each white noise band was separately adjusted by applying Equation 1. Here, $p'(t)$ is sound pressure of the synthetic RIR and $p_n(t)$ is sound pressure of the filtered white noise in the n^{th} octave band. For example, n = 1, 2, 3, 4 correspond to the 31.5 Hz, 63 Hz, 125 Hz, 250 Hz octave bands respectively. T_n is the octave band reverberation time of the n^{th} band, t is time in seconds and d is a decay adjustment value.

$$p'(t) = \sum_{n=1}^{10} \left[p(t) \times \exp\left\{ \frac{-\log_e(1000) \times t}{T_n \times 1.04^d} \right\} \right] \ldots(1)$$

As seen in Equation 1, once the decay rate of the each white noise band had been separately adjusted, they were combined by summation to form the broadband synthetic RIR. After this, the direct sound impulse followed by the initial time delay gap (ITDG) of 0.02 s were added to the synthetic RIR. Although the decay rate of the each white noise band was adjusted with an exponential function (as seen in Equation 1), the synthetic RIR does not have a perfect exponential decay rate because of the summation of the different decay rates of noise. The T_n for each octave band was chosen so that the synthetic RIR has octave band T values similar with those of a RIR measured in a real auditorium. Table 1 shows octave band T values of the synthetic RIR when d = 0. Figure 2 shows an example of a synthetic RIR.

Table 1. Octave band T values of the synthetic RIR

Centre Freq. (*Hz*)	*31.5*	*63*	*125*	*250*	*500*
T (*s*)	2.6	2.0	1.7	2.1	2.1
Centre Freq. (*Hz*)	1000	2000	4000	8000	16000
T (*s*)	1.9	1.8	1.4	0.9	0.6

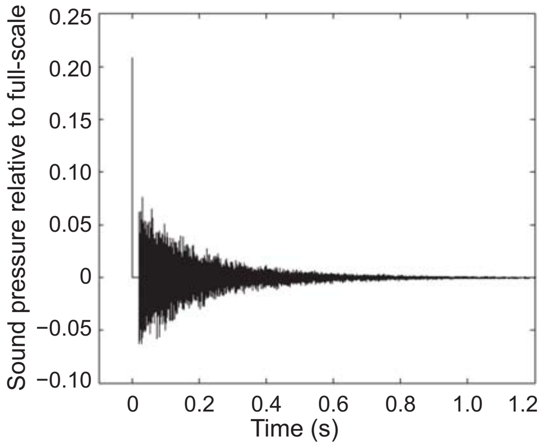

Figure 2. Sound pressure of the synthetic RIR

The experiment consisted of two parts. The first part (PART I) tested the perceived decay rate of the synthetic RIRs in relation to a L_{AFmax} level variation (we use this as the independent variable, rather than loudness in sones, as this allows us to construct easily interpretable equal reverberance contours straightforwardly). As seen in Figure 3, the reference stimuli in PART I have a L_{AFmax} level ranging from 50 dBA to 80 dBA, while T_{mid} is fixed at 2 s. The subscript 'mid' means the average of octave band parameter values in the 500 Hz and 1 kHz octave bands. Apart from the gain change, comparison stimuli in PART I were same as the reference stimuli, and the physical decay rate (which is quantified by T_{mid}) of the comparision stimuli was adjusted by experiment participants, so as to match the perceived decay rate of the corresponding reference stimulus.

The second part (PART II) tested the effect of the physical decay rate on the perceived decay rate of the synthetic RIRs. The level of reference stimuli was fixed at a L_{AFmax} of 60 dBA and T_{mid} varied almost logarithmically from 1 s to 3 s (as seen in Figure 3). The comparison stimuli for PART II were same as those in PART I and, again, the physical decay rate of the comparison stimuli was adjusted in the experiment. Although four pairs are common to PART I and PART II (when the reference stimulus has a T_{mid} of 2 s and a L_{AFmax} of 60 dBA), and so they were tested only once so as to shorten the time taken for the exepriment. Hence a total number of pairs tested in the experiment was twenty-eight. However the results from common pairs were included in the analyses of both parts of the experiment.

The experiment took a form of a magnitude-matching task. The decay rate of the comparison synthetic RIRs was adjusted by pressing 'More' or 'Less' buttons on the MATLAB GUI, which changed the *d* value of Equation 1 by ±1. According to ISO 3382-1 [1], the just noticeable difference (JND) of reverberance corresponds to a 5 % change of EDT_{mid}. Hence the equation was designed to change T_{oct}, T_{mid} and EDT_{mid} of the synthetic RIRs by approximately 4 % by incrementing and decrementing *d*. The available *d* adjustment in the experiment was from *d* = -36 to *d* = +18, which corresponds to T_{mid} of a synthetic RIR from 0.5 s to 4.0 s. The initial value of T_{mid} for each comparison stimulus was randomly chosen over a range of *d* = ±7 from the *d* of the corresponding reference stimulus. For example, a T_{mid} of 1 s corresponds to *d* = -19. Therefore, if a reference stimulus has a T_{mid} of 1 s, the initial *d* value of corresponding comparison stimuli was randomly chosen between *d* = -26 and *d* = -12. By doing this (rather than using the full available range for randomising the initial *d* value of the comparison stimulus), we could avoid the need for a subject to the 'More' or 'Less' buttons a frustratingly large number of times in matching the perceived decay rate between two given stimuli. Once the decay rate of the comparison stimulus was matched to that of the reference stimulus, the subjects moved to the next pair by pressing 'Next' button on the GUI and repeated the process.

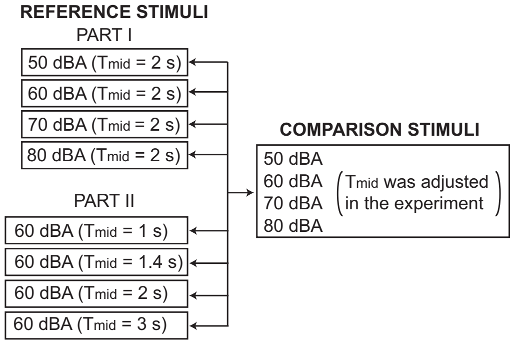

Figure 3. Stimuli parameters for PART I and PART II

The experiment was conducted in an anechoic chamber, which has a background noise level below the threshold of hearing specified in ANSI S12.2 [20]. The stimuli were listened to via circumaural headphones (Sennheiser HD600). Eleven subjects participated in the experiment and ten of them had an educational background in acoustics including room acoustics. None of the eleven subjects self-reported any hearing loss. Prior to the actual experiment, a training experiment was carried out so that the subjects could experience the process of matching the decay rate of two synthetic RIRs. For this, we used stimuli that could be matched physically identical, providing feedback to the subject. (However in the actual experiment after training was completed, many of the stimuli could not be adjusted to be physically identical due to their gain difference). In the training phase, if a subject mismatched the physical decay rate between two given synthetic RIRs within a range of *d* = ±1 (which is less than one unit of the JND of reverberance), the subject moved to the next pair when the 'Next' button on the MATLAB GUI was pressed. If a physical decay rate was mismatched by greater than *d* = ±1 and the 'Next' button was pressed, the words 'Press the *more* button at least one more time' or 'Press the *less* button at least one more time' appeared so as to provide some direction on how stimuli should be matched. It should be noted that these words did not appear in the actual experiment and the subject moved to the next pair regardless of the extent to which their response matched or mismatched the physical decay rate of two stimuli when the 'Next' button was pressed.

RESULTS

The reliability of each subject's responses can be gauged from the degree to which they matched the seven pairs (of reference and comparison stimuli) that had the same L_{AFmax} for reference and comparison. Like those in the training phase, these pairs can be physically matched, and so give us an indication of each subject's ability to do the experiment task. The results are visually represented in Figure 4. The vertical axis is the unsigned average of the *d* discrepancies and the horizontal axis is the subject number. As seen in the figure, subject 8 yields an average *d* discrepancy greater than

2, which corresponds to a modified T_{mid} of 1.9 s to 2.2 s for an unmodified T_{mid} of 2.0 s (this corresponds to approximately 2 times the JND of reverberance). Hence, subject 8 was excluded from the further analyses.

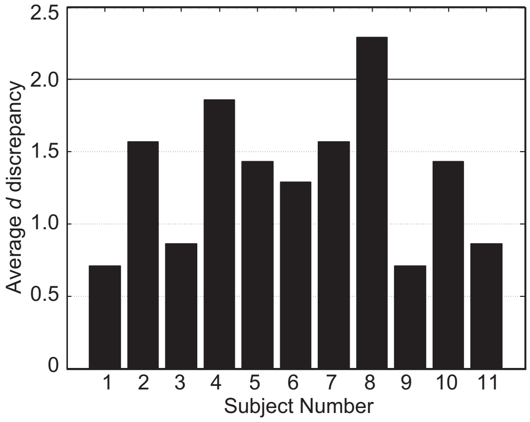

Figure 4. Unsigned average of *d* value discrepancies between the reference and comparison stimuli for which the L_{AFmax} of reference and comparison stimuli were identical.

The subject responses to PART I and PART II were separately averaged and synthetic RIRs were generated from the averaged subject responses. This was done to calculate both the corresponding conventional parameters and the loudness-based parameters from the subject responses expressed in *d*. Figure 5 shows the equal reverberance contours as a function of the level of the comparison stimuli constructed from the synthetic RIRs possessing the averaged subject responses of PART I (the upper figure) and PART II (the lower figure). The idea of Figure 5 is to graphically represent the extent to which the conventional T_{mid} adjustment is required so that the perceived decay rate of the comparison stimuli is matched to that of the reference stimuli. The symbols are the T_{mid} derived from the synthetic RIRs possessing the averaged subject responses, and the trends are shown by linear regression lines. The averaged root-mean-square (r.m.s) deviation between the linear regression lines and the raw dataset of the $T20_{mid}$ is 0.06 s for PART I and 0.05 s for PART II. The error bars on the symbols indicate a 5 % error (i.e., one JND) around each point. In order to disentangle the error bars, the symbols and the error bars shown in the upper figure are slightly offset horizontally. As seen in Figure 5, it is obvious that a reduced T_{mid} is required to match the reverberance of a RIR that has a greater sound pressure level than the reference stimulus. The regression lines for PART I are not far from parallel (apart from the 80 dBA line, which is a relatively poor fit), whereas those from PART II are clearly not parallel. With respect to PART II, this indicates that the effect of sound pressure level on reverberance is greater when the reverberation time is longer.

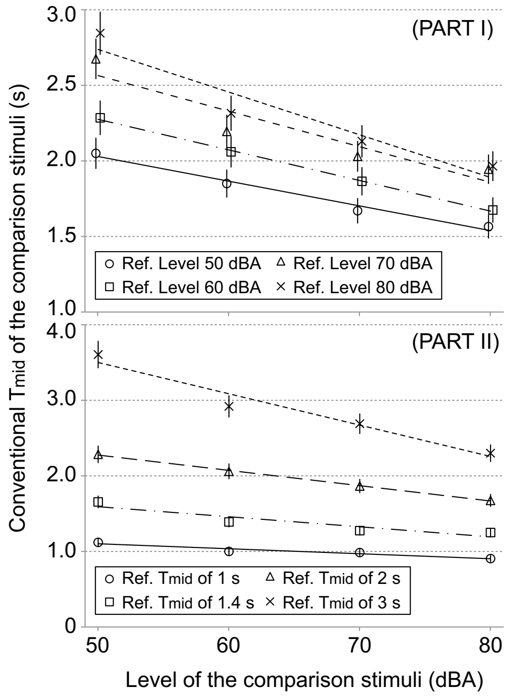

Figure 5. Equal reverberance contours as a function of the level of the comparison stimuli derived from the subject responses for PART I (above) and for PART II (below).

Table 2 shows the results of an ANalysis Of VAriance (ANOVA) executed on the subject responses for PART I. An ANOVA groups a given dataset and compares *variation between groups* with *variation within groups*. If the former is significantly greater than the latter, this indicates that the tested dataset (before being grouped) is significantly affected by factors that group the tested dataset. In the present study, an ANOVA was performed in order to test if the trends seen in Figure 5 were given by a chance or from the significant effect of listening level. The extent to which the different RIRs affect to the subject responses is represented in Table 2. Using a confidence level of 95%, values of *Prob*>F less than 0.05 means that the corresponding variable has a significant effect on the subject responses. As seen in the table, the two variables (the reference level and the comparison level) significantly affect the responses. However there is not a significant interaction effect between the two variables.

Table 2. ANOVA for PART I

Variable	*Sum Sq*	*d.f*	*Mean Sq*	*F*	*Prob>F*
Ref Level	1130.4	3	376.8	50.82	0
Comp Level	1546.9	3	515.633	69.55	0
*Ref * Comp Level*	33.5	9	3.722	0.5	0.8711
Error	1067.6	144	7.414		
Total	3778.4	159			

Although an ANOVA confirms that the listening levels significantly affect the subject responses, this does not mean that there are significant differences between the subject responses for particular listening level conditions. Hence, a Tukey/Kramer's *post hoc* test (often referred to as Tukey's HSD) was performed to investigate if there are significant mean differences between the subject responses for different listening levels. This was done by calculating a confidence interval at a given confidence level with taking into account the group averages, the *Mean Square Error* (MSE), the sample size and the critical value (from the *Studentised Range Distribution*). If a value of zero is not within the confidence interval, this indicates that the mean difference between the two group means is not zero at the chosen confidence level.

In other words, there is a *significant mean difference* between the tested two groups. A Tukey's HSD test tests all possible pairs of groups. In Table 3, *CI Low* refers to the lower bound of the confidence interval at the confidence level of 95%, and *CI High* refers to the higher bound of the confidence interval at the same confidence level. As seen in the table, a significant *mean difference* is observed between all the reference levels, except between 70 dBA and 80 dBA.

Table 3. Multi-comparison test for effects of level of the reference stimuli on subject responses of PART I

Comp Level	*CI Low*	*Mean Diff.*	*CI High*
50 dBA, 60 dBA	-5.0667	-2.7000	-0.3333
50 dBA, 70 dBA	-8.0667	-5.7000	-3.3333
50 dBA, 80 dBA	-9.1667	-6.8000	-4.4333
60 dBA, 70 dBA	-5.3667	-3.000	-0.6333
60 dBA, 80 dBA	-6.4667	-4.1000	-1.7333
70 dBA, 80 dBA	-3.4667	-1.1000	1.2667

Table 4 shows the results of an ANOVA executed on the subject responses for PART II. This was also calculated with the confidence level of 95%. As seen in the table, the two variables (T_{mid} of the reference stimuli and the level of the comparison stimuli) have a significant effect on the subject responses. The *F* values in the table indicate that the T_{mid} of the reference stimuli (F = 991.18) has stronger effect on the subject responses than the level of the comparison stimuli (F = 81.54). Table 5 shows the results of a Turkey's HSD test for the reference T_{mid}. The table shows that there is a significant *mean difference* between all the tested reference T_{mid} values at a confidence level of 95 %.

Table 4. ANOVA for PART II

Variable	*Sum Sq*	*d.f*	*Mean Sq*	*F*	*Prob>F*
Ref. T_{mid}	17331.1	3	5777.04	991.18	0
Comp Level	1425.8	3	475.27	81.54	0
*Ref * Comp Level*	112.4	9	12.48	2.14	0.0296
Error	839.3	144	5.83		
Total	19708.6	159			

Table 5. Multi-comparison test for effects of T_{mid} of the reference stimuli on the subject responses of PART II

Ref. T_{mid}	*CI Low*	*Mean Diff.*	*CI High*
1 s, 1.4 s	-11.2503	-8.9000	-6.5497
1 s, 2 s	-20.4503	-18.1000	-15.7497
1 s, 3 s	-30.3253	-27.9750	-25.6247
1.4 s, 2 s	-11.5503	-9.2000	-6.8497
1.4 s, 3 s	-21.4253	-19.0750	-16.7247
2 s, 3 s	-12.2253	-9.8750	-7.5247

To derive the loudness-based parameters, we used the Dynamic Loudness Model by Chalupper and Fastl [10] (which is implemented in PsySound3 [18]) for calculating the loudness decay functions. Figure 6 compares the coefficient of variation of the tested parameters derived from the synthetic RIRs possessing the averaged subject responses of PART I. The coefficient of variation is the standard deviation divided by mean of a data set. Since the examined parameters always have positive values, the coefficient of variation removes mean-related biases that would be found in the standard deviation (i.e., we would expect smaller standard deviations as means approach zero). A smaller coefficient of variation represents a better prediction of the perceived decay rate, because a decay rate of each set of comparison stimuli was perceptually matched to a decay rate of a single reference stimulus. Figure 6 shows that the two loudness-based parameteres significantly outperform the conventional reverberance predictors for all the tested levels, and that T_N is the best predictor.

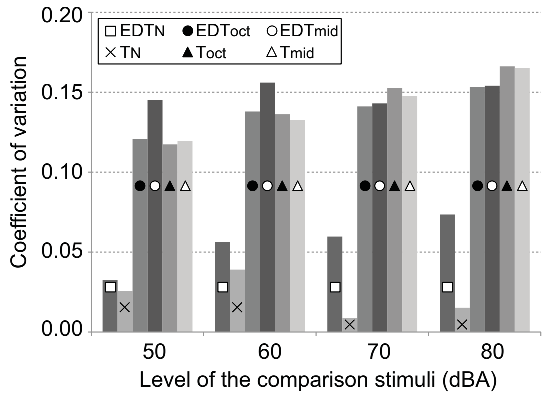

Figure 6. Coefficients of variation for the RIRs generated from mean subject responses in PART I as a function of listening levels of the reference stimuli

To assess the extent to which the T_N and EDT_N correspond to the perceived decay rate of the tested synthetic RIRs, Figure 7 shows the two parameters derived from the same synthetic RIRs used for Figure 6. The trends are shown by linear regression lines, which have an averaged r.m.s deviation of 0.02 s from the raw dataset for EDT_N (the upper figure) and of 0.03 s for T_N (the lower figure). A shallower slope of the linear regression lines represents a better match with the perceived decay rate, because the parameters should have a same value for the different comparison levels once the perceived decay rate of the synthetic RIRs are matched. As seen in the figure, EDT_N seems to exaggerate the effect of level, while T_N yields fairly flat slopes of the linear regression lines.

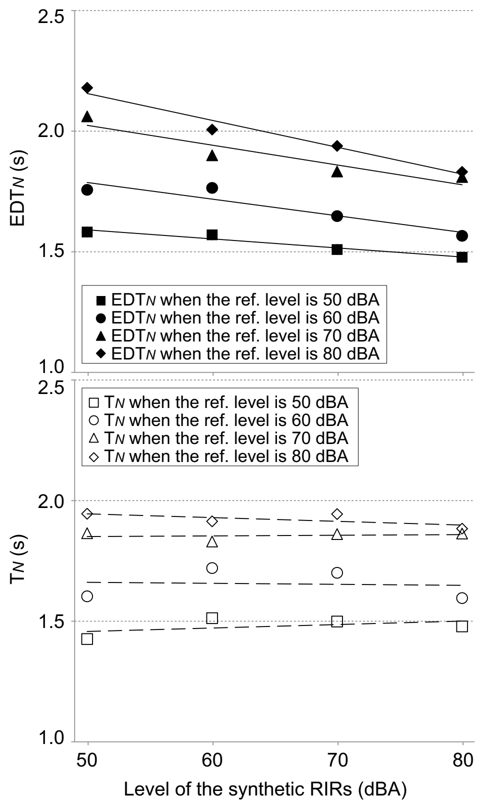

Figure 7. T_N and EDT_N derived from the synthetic RIRs possessing the averaged subject responses of PART I

Figure 8 shows the coefficient of variation of the tested parameters derived from the synthetic RIRs possessing the averaged subject responses of PART II. The results also show that the loudness-based parameters outperform the conventional parameters for all the tested reference T_{mid} values. Like PART I, T_N yields the best match with the perceived decay rate (except for the reference T_{mid} of 1 s).

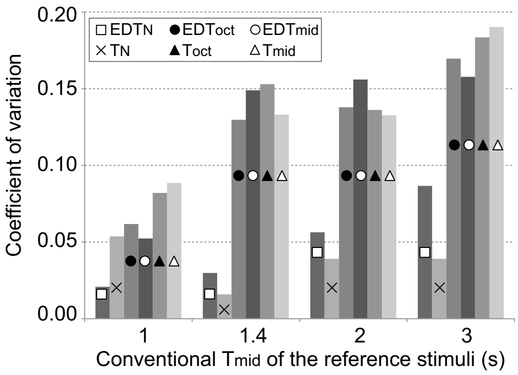

Figure 8. Coefficient of variation for the RIRs generated from mean subject responses in PART II as a function of T_{oct} of the reference stimuli

Figure 9 examines the extent to which the T_N and the EDT_N correspond to the perceived decay rate for PART II. The trends are shown by linear regression lines, which have averaged r.m.s deviations from the raw dataset (over the four lines for each parameter) less than 0.04 s. Like the results shown in Figure 8, the results also indicate that EDT_N exaggerates the effect of level in relation to the perceived decay rate, while T_N yields fairly flat slopes of the linear regression lines.

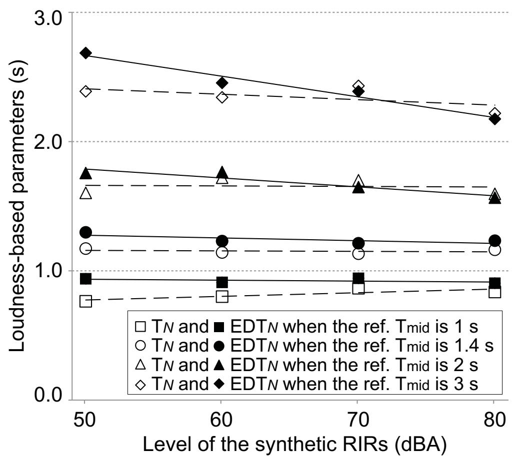

Figure 9. T_N and EDT_N derived from the synthetic RIRs possessing the averaged subject responses of PART II

DISCUSSION

The Dyanmic Loudness Model [10] was used for the analyses, so the accuracy of the model prediction is important in justifying the findings of the present study. As remarked in the Introduction section, the model predictions for non-stationary sounds do not perfectly match the psychoacoustic experiment of Bauch [12] and Moore *et al.* [13], while a relatively good match is observed with the study by Grimm [14]. However it should be noted that considerable discrepancies are also observed in certain cases between other psychoacoustic experiments (such as between the studies by Bauch [12], by Moore *et al.* [21] and by Zhang and Zeng [22]), which investigated the loudness for non-stationary sounds using similar amplitude-modulated pure tones. In our previous work we used both the Dynamic Loudness Model and Glasberg and Moore's Time Varying Loudness Model for a similar analysis of RIR reverberance [6]. We found that both models performed well in predicting reverberance from loudness decay functions. While there are, clearly, areas for improvement in computational loudness modelling, such models provide good working approximations of loudness decay functions.

As seen in the equal reverberance contours derived from the synthetic RIRs possessing the averaged subject responses of PART II (the lower figure of Figure 5), the effect of level on the perceived decay rate weakens as the reference T_{mid} decreases. That is, there is a smaller change of the perceived decay rate for a given level variation, as a reference T_{mid} is lower. Since loudness decay parameters provide a good model for our experiment results, we can also check this by applying loudness decay parameters to RIRs of various sound pressure levels that have T_{mid} matched (instead of their reverberance matched). Figure 10 shows the T_N derived from synthetic RIRs with L_{AFmax} of 50 dBA, 60 dBA 70 dBA and 80 dBA and a T_{mid} of 1 s, 1.4 s, 2 s and 3 s. To summarize the calculation method of T_N, it is the time period of a linear regression line over an evaluation range of 0.708 and 0.178 of the peak of the loudness decay function, multiplied by three. The symbols represent values from these RIRs, without any subjective decay rate adjustment, and trends are shown by linear regression lines. The average r.m.s deviation between the linear regression lines and the raw dataset is approximately 0.04 s. The bars on the symbols extend to the T_N from the synthetic RIRs possessing the averaged subject responses of PART II (in other words, from those having the physical decay rate adjustment in the experiment so as to match the perceived decay rate). As seen in the figure, the difference of the perceived decay rate needs to be compensated for matching the perceived decay rate (which is represented by a length of the vertical lines) is greater for a longer reference T_{mid}.

The figure also shows that the length of the vertical line is shortest for the synthetic RIRs having a L_{AFmax} of 60 dBA. Since the reference stimulus in PART II was fixed at a L_{AFmax} of 60 dBA, it means that the subjects matched the perceived decay rate very precisely when a reference and a comparison stimulus have the same L_{AFmax} level. A similar result is also observed in Figure 5. A comparison T_{mid} was adjusted close to 2 s when the comparison stimulus has a L_{AFmax} of 60 dBA and a corresponding reference stimulus also has a L_{AFmax} of 60 dBA and T_{mid} of 2 s.

The finding that a loudness-based parameter is a better predictor than conventional parameters is consistent with previous findings [5, 6, 7]. The slopes of loudness decay functions are sensitive to the sound pressure level (the slope is less steep as sound pressure level increases), which is in agreement with the subjective data on the effect of sound pressure level on reverberance. The rationale for using loudness based parameters is that the loudness decay function is an approximation of what people hear.

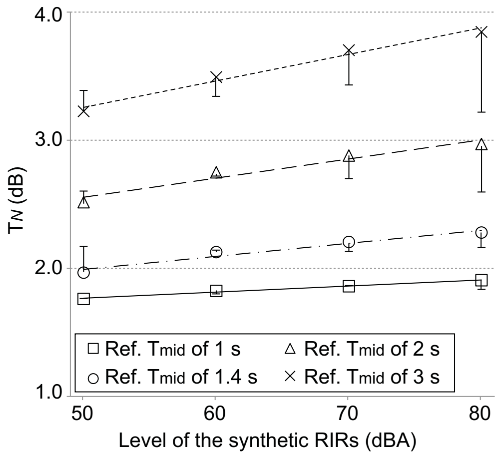

Figure 10. T_N derived from the synthetic RIRs *without the decay rate adjustment* (symbols) and from the synthetic RIRs possessing the averaged subject responses of PART II (bars)

In this study, T_N is the best predictor of reverberance, whereas in our previous study of running reverberance of music stimuli, EDT_N was the best predictor. This is not surprising, as listening to RIRs directly makes the full decay available for auditory evaluation, whereas only the start of the decay is audible in running reverberance. The one exception in our present results, where T_{mid} is 1 s (PART II) might be explained by the relatively rapid loudness decay that occurs at this relatively short reverberation time – which reduces the opportunity for the subject to concentrate on the full sound decay period.

CONCLUSION

This paper examined the effect of listening level and reverberation time on the reverberance of a synthetic RIR, which was expressed in terms of perceived decay rate. Listening level and reverberation time both significantly affect the perceived decay rate of the stimuli. We constructed equal reverberance contours from the experiment data. The experiment results show that loudness-based parameters outperform conventional parameters as predictors of equal reverberance.

REFERENCES

1 ISO 3382-1 Acoustics – measuremet of room acoustics parameters – Part 1: Performance spaces (2009)
2 B.S. Atal, M.R. Schroeder and G.M. Sessler, “Subjective reverberation time and its relation to sound decay” In: Proceedings of the 5th Int. Cong. Acoust. (1965)
3 H. Haas, “The influence of a single echo on the audibility of speech” *J. Aud. Eng. Soc.* **20**, 146-159 (1951, translated 1972)
4 G.A. Soulodre and J.S. Bradley, “Subjective evaluation of new room acoustic measures,” *J. Acoust. Soc. Am.* **98**, 294-301 (1995)
5 D. Lee and D. Cabrera, “Equal reverberance matching of music” In: Proceedings of Acoust. 2009 (2009)
6 D. Lee and D. Cabrera, “Effect of listening level and background noise on the subjective decay rate of room impulse responses: using time-varying loudness to model reverberance” *App. Acoust.* **71**(9), 801-811 (2010)
7 D. Lee, D. Cabrera and W.L. Martens, “Equal reverberance matching of running musical stimuli having various reverberation times and SPLs” In: Proceedings of the 20th Int. Cong. Acoust. (2010)
8 S. Hase, A. Takatsu, S. Sato, H. Sakai and Y. Ando, “Reverberance of an existing hall in relation to both subsequent reverberation time and SPL” *J. Sound. Vib.* **232**, 149-155 (2000)
9 E. Zwicker and H. Fastl, *Psychoacoustics: facts and models*, 2nd updated Edition, (Springer, Berlin, 1999)
10 J. Chalupper and H. Fastl, “Dynamic Loudness Model (DLM) for normal and hearing-impaired listeners” *Acta. Acust. United. Ac.* **88**, 378-396 (2002)
11 B.R. Glasberg and B.C.J Moore, “A model of loudness applicable to time-varying sounds” *J. Aud. Eng. Soc.* **50**, 331-342 (2002)
12 E. Bauch, “Díe bedeutung der frequenzgruppe für díe lautheit von klängen” *Acustica* **6**, 40-45 (1956)
13 B.C.J. Moore, D.A Vickers, T. Baer and S. Launer, “Factors affecting the loudness of modulated sounds” *J. Acoust. Soc. Am.* **105**, 2757-2772 (1999)
14 G. Grimm, V. Hohmann and J.L. Verhey, “Loudness of fluctuating sounds” *Acta. Acust. United. Ac.* **88**, 359-377 (2002)
15 J. Rennies, J.L. Verhey and H. Fastl, “Comparison of loudness models for time-varying sounds,” *Acustica* **96**, 383-396 (2010)
16 D. Lee and D. Cabrera, “Basic considerations in loudness-based analysis of room impulse responses,” *Building Acoustics* **16(1)**, 31-46 (2009)
17 S.S. Stevens, “The measurement of loudness” *J. Acoust. Soc. Am.* **27**, 815-829 (1955)
18 D. Cabrera, S. Ferguson, F. Rizwi and E. Schubert, “‘PsySound3’: a program for the analysis of sound recordings” In: Proceedings of Acoustics’08 (2008)
19 M. Morimoto and A. Asaoka, “Multi-dimentional analysis of *Reverberance*” In: Proceedings of the 18th Int. Cong. Acoust. (2004)
20 ANSI S12.2 *American national standard criteria for evaluating room noise* (1995)
21 B.C.J. Moore, S. Launer, D. Vickers and T. Baer, “Loudness of modulated sounds as a function of modulation rate, modulation depth, modulation waveform and overall level” in Psychophysical and Physiological Advances in Hearing, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).
22 C. Zhang and F.G. Zeng, “Loudness of dynamic stimuli in acoustic and electro hearing” *J. Acoust. Soc. Am.* **102**, 2925-2934 (1997)