ACOUSTIC MEASUREMENT METHODS FOR OUTDOOR SITES: A COMPARATIVE STUDY

Mariana Lopez,
Department of Theatre, Film and Television
University of York
York, UK
mjl503@york.ac.uk

Sandra Pauletto,
Department of Theatre, Film and Television
University of York
York, UK
sandra.pauletto@york.ac.uk

ABSTRACT

Acoustic measurements of outdoor sites require researchers to carefully consider the appropriate method to ensure reliable results. This entails the consideration of the signal-to-noise ratio (SNR), the presence of visitors as well as restrictions that are specific to the site. The present paper presents the results of an experiment conducted in a controlled environment with the aim of comparing the reliability in the presence of different types of noise of three measurement techniques: the Exponential Sine Sweep (ESS) method using a 90-second sine sweep, this same method but with the application of time averaging of multiple measurements and the Inverse Repeated Sequence (IRS). The results are presented in relation to reverberation time and clarity parameters and demonstrate that under the test conditions the ESS method when used with a long sine sweep is the most dependable in the presence of the noise disturbances studied. These findings are of relevance for the application of convolution reverb in audio postproduction.

1. INTRODUCTION

The widespread use of convolution reverb for audio postproduction requires the constant expansion of impulse response (IR) libraries to meet the users' creative needs. A challenging task is the addition of IRs of outdoor sites, due to the difficulties in finding a suitable method for conducting the measurements.

Throughout the years several techniques have been developed with the objective of measuring the IRs of spaces accurately. Acousticians have used a variety of impulsive sounds as excitation signals, including starter pistols, balloon pops and handclaps, but these sound sources present limitations: they cannot produce sufficient energy at low frequencies, they are not omnidirectional in all frequency bands and repeatability cannot be assumed [1]. In recent years the use of deterministic excitation signals played through loudspeakers has become a popular choice among researchers, advantages include precise reproduction, repeatability of the measurements, and a better SNR [2]. Three popular methods are the Maximum Length Sequence (MLS), the Inverse Repeated Sequence (IRS) and the Exponential Sine Sweep method (ESS).

The MLS method uses a pseudo random noise signal created by a deterministic binary sequence characterised by an order (N) and its length is $l_1=2^N-1$ [3]. The length of the sequence needs to be the same length or longer than the reverberation time of the space being measured [2]. The IRS method uses as excitation signal two MLS sequences, the second one is an inversion of the first [4]. Consequently, the IRS method doubles the amount of time that it takes to complete a measurement but it is more resistant to speaker-induced distortion [5, 6]. In both methods deconvolution is accomplished through circular correlation [6]. The use of these two methods has been suggested for measurements in occupied spaces due to its resistance to noise [6]. The ESS method [7, 8, 9] employs an exponential sine sweep and the deconvolution is linear, using an inverse filter (a time-reversal of the test signal used). When this method is used on weakly nonlinear systems it can separate the harmonic distortion (caused by the loudspeaker) and turn it into pre-delayed signals at the start of the IR. However, recent research has suggested that the odd orders of distortion cannot be separated [10]. The ESS method is more robust in the presence of time-variance and presents a better SNR. The length of the sweep should be greater than that of the space's reverberation time multiplied by the number of octaves (taken as 10) [11].

When IR measurements are carried out in outdoor sites choosing one of these methods requires the consideration of numerous factors including the background noise level, the state of occupancy of the space, site restrictions and the optimum time for carrying out the measurements. The method selected has to provide a sufficient SNR requiring the decay curve to start 45 to 35dB above the background noise for T30 and T20 measurements [12], and has to be robust in the presence of time-variance.

If we assume a linear time-invariant (LTI) system, then the averaging of multiple MLS, IRS or ESS measurements will improve the SNR due to the fact that in every measurement the LTI component will remain the same, while the background noise will be random, resulting in an increase of the SNR by 3dB each time the number of measurements is doubled [2]. However, although time-invariance is assumed as a starting point, we must acknowledge that there will be time-variances introduced by changes in air temperature and more significantly by the presence of passers-by. When time averaging is applied these variances will produce major errors in the calculation of the acoustical parameters at different frequencies [2], which can be the result of phase cancellation effects. Another approach to improving the SNR is the use of a single, very long sine sweep, for instance, of 90 seconds of length. Doubling the length of the sweep increases the SNR by 3dB [2] and has been shown to provide more reliable results that the averaging technique [2, 9].

2. COMPARISON OF IR MEASUREMENT TECHNIQUES

An experiment was conducted in a controlled environment to determine the method that might be preferable for measurements in outdoor sites when noise disturbances are present. This experiment was focused on gaining further understanding on what the best method was for a particular site, Stonegate, a street in central York which is of importance for cultural heritage due to the fact that it was employed in the Middle Ages for drama performances including both spoken and sung extracts [13]. The fact that similar sources of noise are common to most outdoor sites makes the findings presented in this paper relevant to other researchers. The experiment was setup in the Black Box Theatre of the Department of Theatre, Film and Television (University of York) (Figure 1) and the methods tested are specified in Table 1.

Table 1 – *Methods tested in controlled experiment*

	Technique	Excitation Signal
Method 1 (90-second sweep method)	ESS	90-second exponential sine sweep. Frequency range = 22-22000Hz
Method 2 (Time averaging method)	ESS	8 15-second exponential sine sweeps, time averaged after the measurements. Frequency range = 22-22000Hz
Method 3 (IRS method)	IRS	N= 19, 10.92s

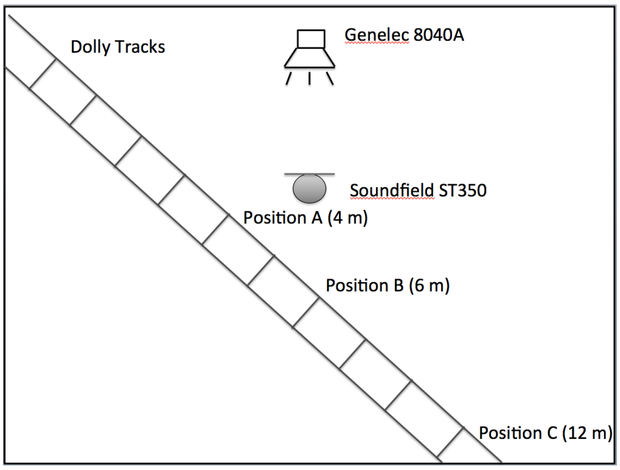

Figure 1: *Setup in BlackBox Theatre, Department of Theatre, Film and Television, University of York*

A Genelec 8040A Monitoring System, which presents omnidirectional characteristics at low frequencies but becomes increasingly directional at mid and high frequencies, was used to play the excitation signal. Although an omnidirectional source is indicated in [12], a directional source was chosen due to the fact that the space of interest (Stonegate) will be studied as a performance space for actors and singers. Therefore, the aim of the measurements is to study how the acoustics of the space respond to a source with similar directivity [14, 15]. A Soundfield ST350 microphone was employed to record the response and both source and receiver were setup at a height of 1.20 metres. The receiver was positioned 2 metres away from the source, and the position of the source and receiver remained fixed and constant during all measurements. A MOTU Traveler mk3 and an HP laptop were used for running Pro Tools 9 [16] and playing back and recording. A second Genelec 8040A was mounted on a dolly track and it was used to playback noise sources. A rope was attached to the speaker to allow the simulation of sound source movement.

The sources of noise chosen for the experiment were selected due to their presence in Stonegate: ambience, York Minster Bells, speech, footsteps and footsteps+speech (referred to here as F+S). The impact of these noise sources placed at four different positions was also tested. The critical distance calculated for the space was 8 metres, therefore positions A and B, which were 4 and 6 metres away from the source playing the excitation signal were both in the near field. Position C was in the diffuse field, approximately 12 meters away from the speaker. The fourth position was a mobile position: the speaker on the dolly tracks was moved from the diffuse field towards the microphone during the measurements.

In addition to the measurements in the presence of noise sources, measurements in optimum conditions (with no added noise sources) were also conducted. Furthermore, every measurement taken was repeated three times to check for consistency, the results below will present the mean of those three repetitions as well as the standard deviation.

Table 2 – *Noise Setups used for IR measurements. Greyed areas indicate the different positions in which the noise sources were located.*

Setups for Measurements with Noise Disturbances				
Noise Sources	Position A	Position B	Position C	Mobile Position
Ambience	■	■	■	
Speech	■	■	■	■
York Minster Bells			■	
Footsteps	■	■	■	■
Foot-steps+Speech	■	■	■	■

The reference output level for the excitation signals and the noise sources was recorded using a -20dBFS pink noise signal and it was 73.5 dBA at 1 metre from the source. The excitation signals were generated with the Aurora Suite [17] and the measurements based on the ESS method were deconvolved using Voxengo Deconvolver [18], whereas the IRS measurements were deconvolved with Aurora.

The analysis of the results is focused on the W channel of the B-format microphone. Three acoustic parameters related to sound decay were calculated: T20, T30 and EDT. T20 expresses the reverberation time considering the decay curve from -5 to -25dB, whereas T30 considers the decay curve between -5 and -35dB. Both parameters were calculated to consider how each of them was affected by a low SNR. EDT (Early Decay Time) considers the first 10dB of the curve. Clarity was calculated using C50, which is an early-to-late arriving energy ratio expressed in decibels and that considers the division between early and late energy as 50 ms.

The changes in the results of the acoustical parameters, as a consequence of the addition of noise sources at different positions, were considered in relation to the Just Noticeable Difference (JND) for the parameters studied. In this paper measurements were considered accurate if the variations between the results in noisy conditions and in optimum conditions were smaller than the JND and inaccurate if they were equal or larger. The JND for T20, T30 and EDT was considered as 5% [19]

for values larger than 0.6 s, and as an absolute value of 0.03 s with parameter results smaller than 0.6 s, based on the findings of [20]. The JND for C50 was considered to be 1.1dB [21]. In the measurements under noisy conditions not all parameters were calculated by the methods tested, this was due to low SNR. In these cases the results have been considered inaccurate and the percentage of results in which this occurred was indicated under the tag NR (No Results) in Tables 3-5.

2.1. Analysis of Results

In the analysis of the results of the measurements with no noise disturbances the 90-second sweep and the time averaging methods presented very similar results regarding all parameters, all differences being smaller than 1 JND, with the exception of the EDT at 2kHz in which the difference was of 1 JND. The IRS method presented greater differences in relation to the values calculated with the other methods. It presented lower values in the results of T20, T30 and EDT in most frequency bands, whereas the C50 results were higher than the ones calculated with the other two methods. When calculating the T20, differences ranged from 1-5 JNDs, while the EDT differences ranged from 1.11-15 JNDs, neither of these parameters presented differences smaller than 1 JND. When considering T30 there was only one result with a difference smaller than 1 JND, and the rest presented differences ranging from 2.75-5.5 JNDs. C50 results presented variations ranging from 1.13-2.13 JNDs, with differences smaller than the JND at 250Hz, 8 and 16kHz. These results demonstrate the greater reliability under the test conditions of the measurements with both ESS methods. Furthermore, when considering the mean of the three measurements taken with each technique, standard deviation presented very low values for all parameters for the three techniques.

2.1.1. Effects of Noise Sources on T20 results

At 125Hz noise sources had the highest negative impact on the IRS method (Table 3). The greatest deviation for this method was of 5.83 JNDs with the introduction of the F+S sample at position C. The greatest fluctuation in the time averaging method was of 10.6 JNDs with the introduction of the speech sample at position A, whereas the greatest fluctuation in the 90-second sweep method was of 4 JNDs, caused by the speech sample in the mobile position. The IRS method presented the highest values of standard deviation (Figure 2).

Table 3 – *Percentage of measurements of T20 that were affected by the introduction of noise sources. NR (No Results calculated).*

Frequency Band	90-second sweep	Time Averaging	IRS
125Hz	25%	31.25% + 6.25% (NR)	56.25% +6.25% (NR)
250Hz	0%	31.25%	56.25%
500Hz	56.25%	43.75%	12.5%

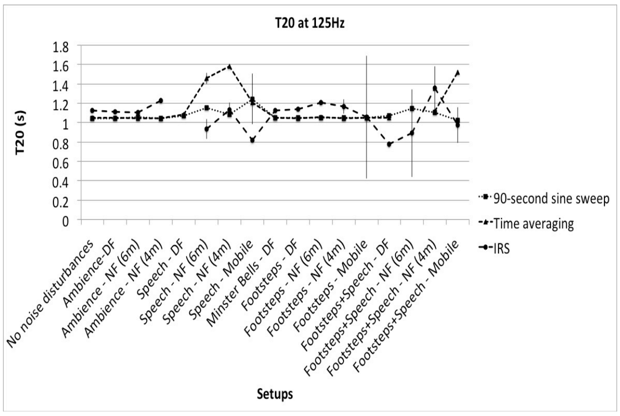

Figure 2: *Mean values and standard deviation of T20 at 125Hz measured with different noise setups.*

At low mid frequencies the time averaging method was the most affected technique in the presence of noise (Table 3). The greatest deviation from the results in optimum conditions was of 2.75 JNDs, caused by the speech sample in position A at 250Hz. The 90-second sweep method presented inaccurate results only at 500Hz, and the greatest fluctuation was of 7 JNDs, recorded with the introduction of the F+S sample at position A (Figure 3). The greatest variation with the IRS method was of 4.67 JNDs at 250Hz with the introduction of the speech sample in the mobile position. At low mid frequencies the highest recorded standard deviation was at 250Hz with the IRS method.

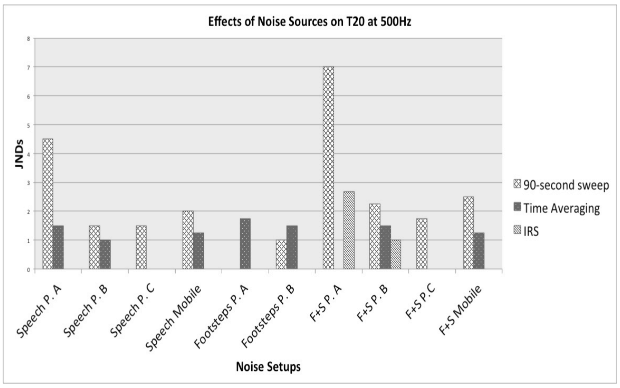

Figure 3: *Comparison of the effects of noise sources (expressed in number of JNDs) on T20 results at 500Hz on the three methods tested. The setups that presented effects below 1 JND were considered accurate and therefore are not included in the figure.*

No significant fluctuations were recorded with any of the methods from 1-8kHz and no values were recorded at 16kHz, very low values of standard deviations were present at all these frequency bands.

2.1.2. Effects of Noise Sources on T30 results

At 125Hz noise disturbances had the most unfavourable effects when the IRS method was employed (Table 4). The greatest deviation was of 349.17 JNDs at 125Hz with the introduction of the speech sample in the diffuse field. When using the 90-second sweep method the greatest deviation was of 2 JNDs with the introduction of the F+S sample in the mobile position.

The results calculated with the time averaging method had a maximum deviation of 1.4 JNDs with the introduction of the footsteps sample at position B. Very low measurements of standard deviation were found at this frequency band.

Table 4 – *Percentage of measurements of T30 that were affected by the introduction of noise sources.*

Frequency Band	90-second sweep	Time Averaging	IRS
125Hz	12.5% +43.75% (NR)	6.25% +50% (NR)	56.25% +31.25% (NR)
250Hz	75% +18.75% (NR)	68.75% +6.25% (NR)	56.25% +6.25% (NR)
500Hz	37.5% +56.25% (NR)	56.25% +12.5% (NR)	50% +12.5% (NR)
1kHz	37.5%	31.25% +12.5% (NR)	31.25%
2kHz	0%	6.25%	0%
4kHz	0%	37.5%	0%
8kHz	0%	12.5%	0%

At low mid frequencies the 90-second sweep method was the most affected (Table 4). The greatest deviation was of 9 JNDs at 250Hz with the introduction of the voice sample at position C, whereas the time averaging method presented the greatest deviation as 10.25 JNDs at 500Hz with the introduction of the F+S sample at position C. The greatest fluctuation for the IRS method was of 7.33 JNDs at 250Hz with the introduction of the F+S sample in the mobile position. T30 at 250Hz presented high values of standard deviation in some setups (Figure 4) whereas standard deviation at 500Hz presented low values.

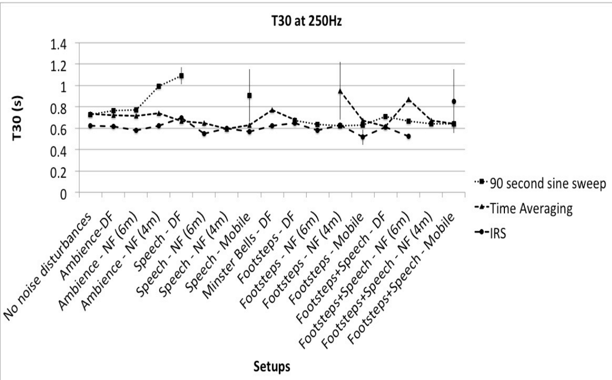

Figure 4: *Mean values and standard deviation of T30 at 250Hz measured with different noise setups.*

At 1kHz the time averaging method had as its greatest deviation 2.6 JNDs with the introduction of the speech sample at position B. The measurements with the 90-second sweep had deviations of up to 3.2 JNDs, with the introduction of the speech sample at position B. The greatest deviation in the IRS method was of 1.75 JNDs with the introduction of the F+S sample in the mobile position. The standard deviation calculated was very low for all methods.

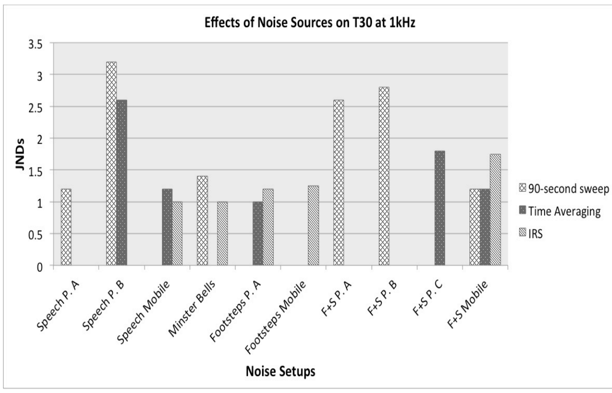

Figure 5: *Comparison of the effects of noise sources (expressed in number of JNDs) on T30 results at 1kHz on the three methods tested. The setups that presented effects below 1 JND were considered accurate and therefore are not included in the figure.*

At high mid and high frequencies neither the 90-second sweep nor the IRS methods presented any inaccurate results. The time averaging method presented its greatest variation with the introduction of the F+S sample at position B, which caused a fluctuation of 4 JNDs at 4kHz. At 8kHz the greatest variation was of 1.33 JNDs with the introduction of the speech sample at position B.

2.1.3 Effects of Noise Sources on EDT results

At low frequencies the IRS method was the most affected. The greatest deviation was of 3 JNDs at 125Hz with the introduction of the F+S sample at position C. The greatest fluctuation with the time averaging method was of 1 JND with the introduction at 125Hz of the speech sample at position A and the F+S sample at position B. The 90-second sweep method did not record any inaccurate results.

Table 5 – *Percentage of measurements of EDT that were affected by the introduction of noise sources.*

Frequency Band	90-second sweep	Time Averaging	IRS
125Hz	0%	12.5%	25%
500Hz	0%	6.25%	6.25%
1kHz	12.5%	50%	62.5%
2kHz	75%	81.25%	0%
4kHz	75%	68.75%	0%
8kHz	75%	68.75%	0%

At low mid frequencies the 90-second sweep method did not present any inaccurate results. The time averaging and IRS methods only presented a variation of 1 JND at 500Hz. With the time averaging method this deviation was introduced by the presence of the speech sample at position A, while with the IRS method this deviation was caused by the introduction of the F+S sample at the same position.

At 1kHz the most affected method was the IRS method (Table 5). The greatest deviation in the IRS and time averaging methods was of 1.33 JNDs. The largest deviation in the 90-second method was of 1.25 JNDs with the introduction of the footsteps sample at position A. At 2kHz the 90-second sweep and the time averaging methods were the only affected (Table 5). The greatest deviation for the 90-second method was of 2.75 JNDs with the introduction of the F+S sample at position A, whereas the greatest deviation with the time averaging method was 4.67 JNDs, caused by the footsteps sample at position A. From 2kHz upwards both the 90-second sweep and the time averaging methods had high percentages of inaccurate results. Standard deviation was low for all methods at all frequency bands.

2.1.4 Effects of Noise Sources on C50 results

All variations in the results of C50 with the introduction of noise sources in relation to the measurements in optimum conditions were smaller than 1 JND. This demonstrates that under the test conditions the results of this parameter remained reliable even in the presence of noise.

3. CONCLUSIONS

The results analysed demonstrate that under the test conditions the most reliable method for measurements without noise disturbances is the ESS, both using a long sine sweep and time averaging. The IRS method presented lower values of T20, T30 and EDT and higher values of C50 when compared to the other methods. In the presence of noise all methods presented deviations from the values calculated with no injected noise. This applied to all parameters calculated with the exception of C50. Moreover, the methods were affected by the noise sources in different ways depending on the frequency band. At low frequencies, the IRS method was affected the most. At low mid frequencies, the time averaging method was the most affected in relation to the T20 results, however, when considering the T30 results it was the 90-second sweep method that presented the largest percentage of deviations from the optimum results. When considering mid, mid high and high frequencies the time averaging method was the most unreliable. These results demonstrated that the 90-second sweep method is the most dependable method in the presence of the noise sources and positions here studied. Furthermore, the noise sources that had a greater impact on the results were the speech, footsteps and F+S samples at the near field positions and as mobile sources, demonstrating the small impact noises such as ambience and bell samples, as well as the position in the diffuse field have on the results. It was shown that noise sources did not have significant effect on the clarity values.

It is important to point out that although this experiment indicates the level of reliability of different methods in the presence of the noise sources studied, further experiments need to be carried out in order to analyse whether these results are also valid for other spaces and setups. Furthermore, there are other important factors that can affect the results in outdoors sites. Particularly important is to consider the effect of environmental conditions and the unpredictable behaviour of the visitors to the site as having potential effects on the results.

4. REFERENCES

[1] D. Griesinger, ‘Beyond MLS – Occupied hall measurement with FFT techniques’ in AES 101st Convention, Los Angeles, USA, November 8-11, (1996) (Preprint)

[2] ISO 18233: 2006, Acoustics – Application of new measurement methods in building and room acoustics, ISO, (2006)

[3] M.R. Schroeder, ‘Integrated-impulse method measuring sound decay without using impulses’ in Journal of the Acoustical Society of America, 66 (2), pp. 497-500, (1979).

[4] N. Ream, ‘Nonlinear identification using inverse-repeat *m* sequences’ in Proceedings IEE, London, 117, pp. 213-218, (1970).

[5] C. Dunn and M. Hawksford, ‘Distortion Immunity of MLS-Derived Impulse Response Measurements’ in Journal of the Audio Engineering Society, 41 (5), pp. 314-335, (1993).

[6] G. Stan, J. Embrechts and D. Archambeau, ‘Comparison of different impulse response measurement techniques’ in Journal of the Audio Engineering Society, 50 (4), pp. 249-262, (2002).

[7] A. Farina, ‘Simultaneous measurements of impulse response and distortion with a swept-sine technique’ in AES 108th Convention, Paris, France, February 18-22, (2000).

[8] S. Müller and P. Massarani, ‘Transfer-Function Measurement with Sweeps’ in Journal of the Audio Engineering Society, 49 (6), pp. 443-471, (2001).

[9] A. Farina, ‘Advancements in impulse response measurements by sine sweeps’ in AES 122nd Convention, Vienna, Austria, May 5-8, (2007).

[10] A. Torras-Rosell and F. Jacobsen, ‘A New Interpretation of Distortion Artifacts in Sweep Measurements’ in Journal of the Audio Engineering Society, 59 (5), pp. 283-289, (2011).

[11] R. Kessler, ‘An optimised method for capturing multidimensional acoustic fingerprints’ in AES 118th Convention, May (2005).

[12] ISO 3382: 1997, Acoustics-Measurement of reverberation time of rooms with reference to other acoustical parameters, ISO, (1997).

[13] M. Lopez, S. Pauletto and G. Kearney, ‘Virtual Acoustics and performance spaces in Medieval English Drama’ in Conference on The Acoustics of Ancient Theatres, Patras, Greece, September 18-21, (2011).

[14] A. Farina, E. Armelloni and P. Martignon, ‘An experimental study of 20 Italian opera houses: measurement techniques’ in 147th Meeting of the Acoustical Society of America, New York, USA, 24-28 May, (2004). Available at http://pcfarina.eng.unipr.it/Public/Papers/194-ASA-2004.pdf [Last accessed on 12 July 2012]

[15] A. Farina and L. Tronchin, ‘New Measurement Technique for 3D Sound characteriszation in Theatres’ in Conference on The Acoustics of Ancient Theatres, Patras, Greece, September 18-21, (2011). Available at http://pcfarina.eng.unipr.it/Public/Papers/258-AncientTheatres.pdf [Last accessed on 12 July 2012]

[16] http://www.avid.com/US/products/Pro-Tools-Software

[17] http://pcfarina.eng.unipr.it/Aurora_XP/index.htm

[18] http://www.voxengo.com/product/deconvolver/

[19] ISO 3382-1: 2009, Acoustics-Measurement of room acoustic parameters. Part 1: Performance spaces, ISO, (2009).

[20] T.I. Niaounakis and W.J. Davies, ‘Perception of Reverberation Time in Small Listening Rooms’ in Journal of the Audio Engineering Society, 50 (5), pp. 343-350, (2002).

[21] J.S. Bradley, R. Reich and S.G. Norcross, ‘A just noticeable difference in C_{50} for speech’ in Applied Acoustics, 58, 99-108, (1999).