Perception of low frequencies in small rooms

Fazenda, BM and Avis, MR

Title	Perception of low frequencies in small rooms
Authors	Fazenda, BM and Avis, MR
Publication title	Proceedings of IV Congreso Ibero-Americano de Acustica
Publisher	
Type	Conference or Workshop Item
USIR URL	This version is available at: http://usir.salford.ac.uk/id/eprint/9451/
Published Date	2004

Perception of Low Frequencies in Small Rooms

B.Fazenda[a,b], M.R.Avis[b]

[a] *Acoustic Research Centre, University of Salford, Salford, UK*
[b] *Acoustic Technology, Technical University of Denmark, Lyngby, Denmark*

ABSTRACT: Critical listening rooms are spaces carefully designed to provide the listener with an accurate listening environment. It is well known that in small rooms, resonances may occur that will cause some amplification or attenuation of sound at certain frequencies. These problems are known to cause a distraction from the optimum perception and the correct judgement of the reproduced sound.

A study has been carried out on various factors that affect how “bass” is perceived in rooms. Parameters like the modal distribution and the Q-factor of modes have been studied using a binaural technique that employs virtual representation of different sound fields. This technique allows different room conditions to be compared relatively quickly without actually having to change the conditions in a real room. The purpose of this study is to identify the subjective perception of these factors and to attempt to classify its perceptual importance with regards to time and frequency domain effects introduced by resonances.

Results indicate how changes to these factors are subjectively perceived by users. In practice these results may be applied to a more efficient use of acoustical treatment and aid the design of better rooms.

1. INTRODUCTION

Low frequency resonances in rooms have been identified as a common problem in critical listening applications [1]. An efficient effort to avoid or ameliorate these problems should rely not only on a comprehensive understanding of the mechanisms behind resonances in rooms, but also on a correlation between these mechanisms and their subjective perception. Previous studies have shown that the perception of resonances is highly dependent on the type of excitation stimulus [2,3,4]. Furthermore, it has been identified that high Q resonances are the most detectable and pose the greatest problem due to their long resonant decays [4,5]. Furthermore Fazenda et al carried out a study on the perception of modal distribution and the detection of changes dependent on room aspect ratios [6], which shows little correlation between a rank order of “good aspect ratios” and its perception. This study also demonstrated that the detection of changes is much dependent on the frequency content of the excitation signal and how this matches a particular point in the frequency response of a room. Decades earlier on another study of the perception of resonances [2], Bucklein concluded that for common music signals, resonances are only detectable if the excitation signal lies close in frequency. Following from previous work, the significance of this paper is on the effects of exciting room resonances using a single tone stimulus controlled to excite specific points of the frequency response. A distinction is made between resonance, normal level and anti-resonance excitation. This is done in order to understand further the effects of amplitude, frequency and time on the perception of single tones in the presence of resonances.

The cases under study are plotted in *Figure 1*, which describes three situations that may be found in common rooms. All these are centred at the frequency of the input stimulus and the overall gains have been adjusted so that the gain of each filter at the driving frequency is the same. The resonant frequencies are 80 Hz for the single resonance and 70 Hz, 90 Hz for the dual resonances. The anti resonance is obtained by inverting the phase of one of the filters. The cases under study will henceforth be referred to as *Peak* for the single resonance at 80 Hz, *Valley* for the smooth response when the two 70 Hz and 90 Hz add to produce a smooth region at the excitation frequency, and *Dip* to represent the anti-resonance at 80 Hz.

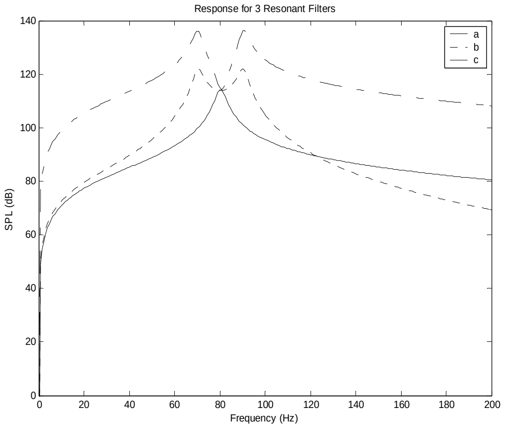

Figure 1 – Three resonant cases under study (a)Peak;(b)Valley;(c)Dip

In order to study the effects of resonances on single tones, filter models based on the IIR bi-quad design were used to generate one or two resonances at specific frequencies, with determined Q-factor values. Tests were carried out at a normalised amplitude level in order to identify effects that are solely associated with time or frequency characteristics of the systems under study.

Subjective tests concentrated on the audibility of peaks, valleys and anti-resonances, as shown in *Figure 1*. The results are important in identifying the characteristics of resonances most likely to produce audible effects when employing treatment methods to reduce the problem of room modes.

2. OBJECTIVE EFFECTS OF RESONANCES ON SINGLE TONES

2.1 The Temporal Aspects of Input Stimulus

There are a number of types of input stimulus that have been used to test effects of room resonances. In previous tests, the authors have used carefully selected musical signals as these represent a closer match to real applications [5,6]. On this study, the specific interaction between excitation signal and single or dual resonances is studied using a common test signal known as a tone burst [7].. The use of a fixed frequency signal provides a more physical and clearer understanding of the problem. Tone bursts are signals that contain cycles of a single frequency tone, and are shaped to produce the input or 'source velocity' of the system to be studied. In this study, two different input stimuli were, differing on their temporal characteristics. The stimuli are designated as *long stimulus* for the smooth sine burst shown in *Figure 2*, and *short stimulus* for the exponentially decaying sine wave shown in *Figure 3*.

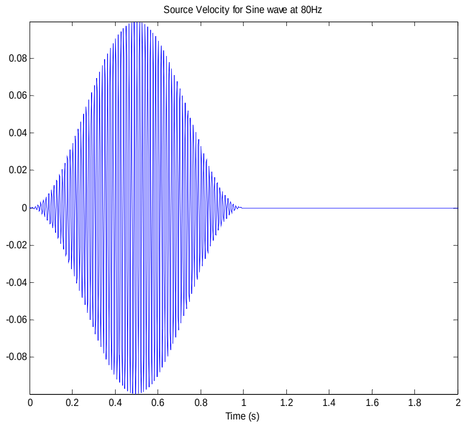

Figure 2 – Long input stimulus

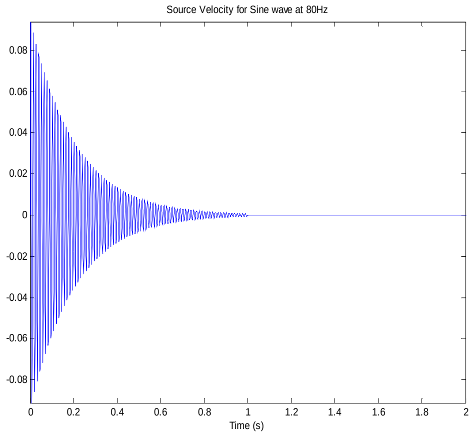

Figure 3 – Short input stimulus

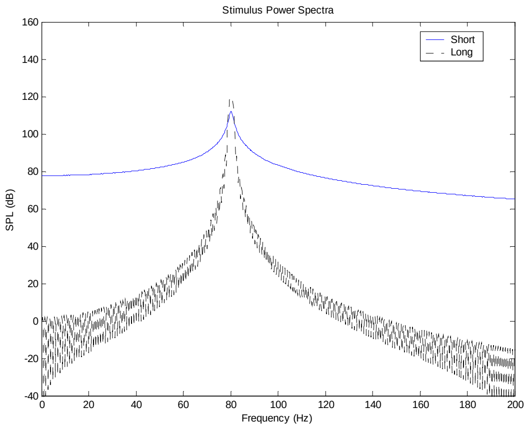

Figure 4 – Power Spectra for short and long input stimulus

The main difference between these two inputs is their temporal response. *Figure 4* shows the power spectra for each signal. It is clear that the short exponentially decaying sine wave shows much higher levels of energy outside the main frequency component when compared to the longer more continuous tone. The reason for this difference between the signals comes from the fact that fast transients in the time response result in a smear of energy in the frequency response. The constant magnitude frequency response for a delta function is an extreme example of this. The smooth onset and offset portions of the sine burst concentrate the energy in the main component of the signal. The side lobes visible in the response result from the Fourier transform of the window applied to the long stimulus with the aim of introducing the tapering on the onset and offset sections.

2.2 System Responses to Stimulus

The responses of the three systems shown in *Figure 1* are modelled through a simple filtering process. Some of the most interesting cases are presented here.

The response of a resonance, like the one plotted in *Figure 1*(a) may be determined for the two types of input. *Figure 5* and *Figure 6* show the response of the filter system to the long and short impulses respectively. *Figure 5* does not seem to show apparent distortions on the input signal. However, more careful analysis shows that the output is somewhat longer, extending beyond 1 second, which is the original length of the input stimulus. This long decay is associated with the resonant properties of the filter, and directly related to its Q-factor. This effect is sometimes described as ’ringing’, and for very high Q resonances ($Q>15$) it may be clearly audible [5]. However because the effect happens at the same frequency as the input signal, it is only audible after the input has been switched off. A longer rise time on the onset of the waveform is also evident. This indicates an onset delay until the system reaches full amplitude. The resonant behaviour of the filter imposes a gain on the output level, as expected. The same features may be observed in *Figure 6* although with a more visible effect. It is clear that the fast transients have been ‘smoothed’.

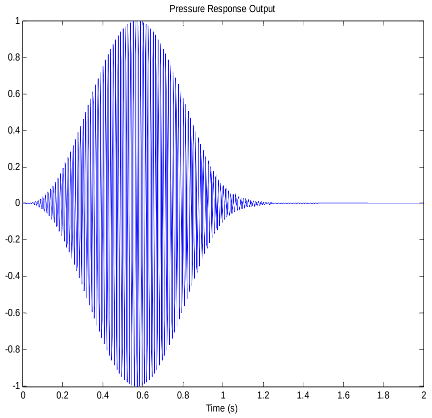

Figure 5 – Peak response to long stimulus

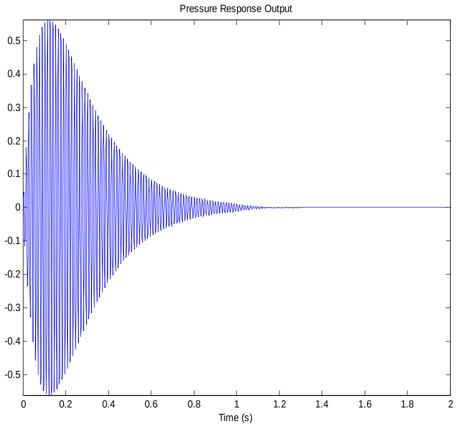

Figure 6 – Peak response to short stimulus

Given that this is a linear system, the output spectrum of these waveforms should not present any frequency components other than those present at the input.

Perhaps a more interesting case is one where the input stimulus falls on a frequency region between two resonances (see *Figure 1* (b)). This is shown in Figure 7 and Figure 8. It can be seen that the response to a long input stimulus does not show any clear effects of altering the input waveform, apart from a very low amplitude decay after 1 second. In contrast, the response of this system to a short stimulus is quite dramatic. Observation of Figure 8 clearly shows some interesting effects in the time domain. The expected amplification of levels is evident. There is also an increase on the attack time. The resonant behaviour has reduced the fast transient present in the input stimulus. Nonetheless, the short time variance introduced here does not appear likely to cause perceptible effects. Perhaps, more intriguing is the amplitude fluctuation clearly marked by the peaks and troughs of the response. It is interesting to note that this fluctuation has a period of 0.1 second, corresponding to a frequency which is half the frequency difference between the two resonances in the system.

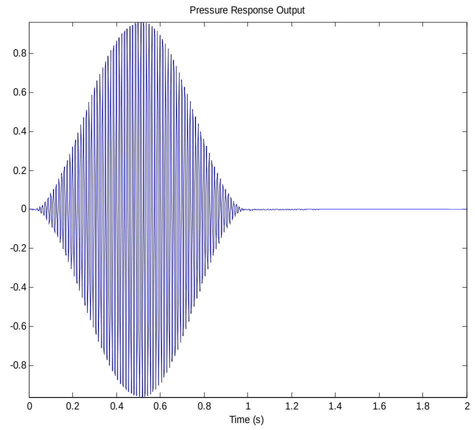

Figure 7 – Valley response to long stimulus

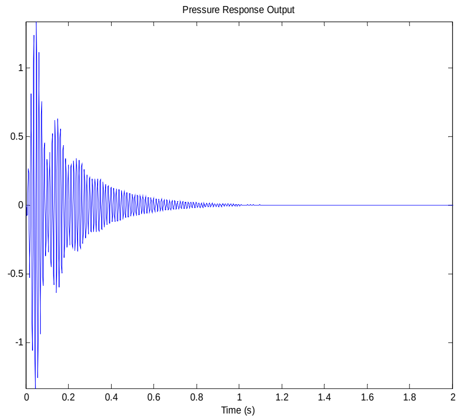

Figure 8 – Valley response to short stimulus

In order to explain the cause for this amplitude fluctuation it is useful to review the addition of two co-sinusoids.

$$A\cos(\omega_1 t) + A\cos(\omega_2 t) = 2A\cos\left(\frac{\omega_1+\omega_2}{2}\right)t \times \cos\left(\frac{\omega_1-\omega_2}{2}\right)t \qquad (1)$$

Equation 1 is helpful in understanding this effect. The addition of two harmonic waves of the same amplitude and different frequency is described. The resulting wave resembles a simple cosine wave of frequency equal to the mean of the two components. Due to the changing phase relationship between these two components, there is reinforcement and cancellation originating a fluctuation in amplitude. This fluctuation occurs at a rate equal to half the difference between the two components. This is represented by the second cosine on the right hand side of equation 1. *Figure 9* displays the effect of adding 70 Hz and 80 Hz cosine waves. It is clear that the resulting wave is comprised of a single frequency whose amplitude is modulated at a rate which corresponds to the difference in frequency between the two components interacting. Indeed, the amplitude peaks repeat at a rate which is twice the frequency predicted by equation 1. This is explained by the fact that the magnitude of the cosine wave reaches maximum amplitude twice per cycle.

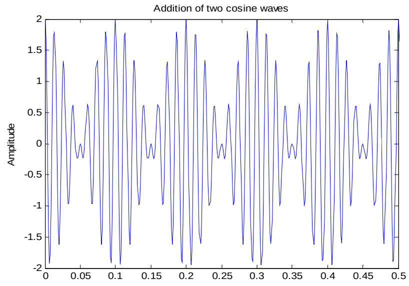

Figure 9 – Addition of two cosines

The subjective effects of this regular fluctuation in amplitude over time have been presented by Moore [8]. The interaction between the two frequency components results in amplitude modulation. If the difference in frequency between the components is small (~<5 Hz) this fluctuation is slow enough and may impose a loudness fluctuation in the perceived sound. This effect has been described as ‘beats’. Furthermore, it has been shown that as the frequency difference increases between the components, the sensation changes from “beats” to one of “roughness”. If, however the two components are well separated in frequency – usually more than one critical bandwidth – the ear discriminates two separate tones [8,pp81]. This suggests that these effects may be audible within the average spacing of low frequency room modes.

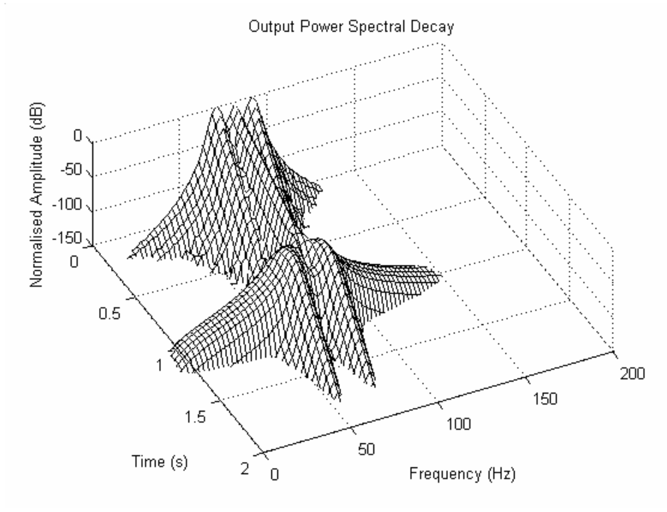

Figure 10 – PSD for short stimulus and Dip case

An interesting display of these effects combines the time and frequency responses in a single plot. This is usually referred to as the *power spectral decay* (PSD), or more commonly a *waterfall* plot. This shows the progress of the spectral components in time. *Figure 10* shows the response of system (c) in *Figure 1*. The time response shows evidence of three components in frequency – the input stimulus and the two resonant frequencies. The latter decay at a much faster rate than the former. When the input is removed, the resonant components reappear in the response to decay at a fixed rate. The level at which these components reappear after the input stimulus has been switched off dictates its perception.

3 SUBJECTIVE PERCEPTION TESTS

The previous section illustrated the mechanisms underpinning the interaction between input stimulus and resonant system. Consideration has been given to effects on quantifiable factors that may be measured and observed in time and frequency responses.

It has been suggested in literature that the sensitivity to resonances may be dependent on the temporal aspects of the input signal [3,4] Furthermore, as suggested by Bucklein [2] and further supported by Fazenda et al.[6] the alignment of the input stimulus and the centre frequencies of the resonances in a system also plays an important role in detection.

This section aims to identify the perceptibility of differences between these factors. Two main aspects have been studied. These are the effects of the temporal structure of the input stimulus and the effects of frequency alignment between the input stimulus and the resonant system. The test procedure utilizes a binaural technique previously used successfully for other tests. Samples are auditioned through headphones.

The tests utilise an automated ABX [9] technique in order to evaluate the differences between two audition samples which represent the cases under study. The task takes the form of a 2 interval forced choice where the subjects compare two samples. Comparisons are made between the original input and a modified version representing one of the resonant systems and between two systems representing a different room situation. All audition samples are normalised to the same amplitude. This is done in order to remove obvious perceptual effects associated with loudness.

The Q-factor of the resonances is fixed at Q=20. This was chosen as a representation of a small room with hard walls.

The tests were carried out on six subjects. All subjects reported no hearing disabilities and had previous experience on listening and psychoacoustic tests. For each test case the subjects were asked to perform 10 trials, after which the results were recorded and a new test was started for a different case. A total of six tests were performed by each subject, lasting no longer than 20 minutes in total. Tests were carried out in a quiet environment.

4. RESULTS AND DISCUSSION

4.1 Effects of Temporal Structure of Stimulus

Figure 11 shows the average results for all cases tested. It is noticeable the large difference of correct responses for short and long stimulus. A non-parametric *Friedman analysis of variance* further shows that there is a highly significant difference dependent on the type of stimulus. It is clear that the short stimulus induces the higher rate of correct responses. Furthermore, the larger deviations obtained in the results for the long stimulus reflect the higher level of difficulty in detecting differences for this type of signal. A chi-square test calculated for each individual data indicates the efficiency in detecting differences. This is shown under the value for *p* in *Table 1*. A result is statistically significant if the value is at least below 0.05.

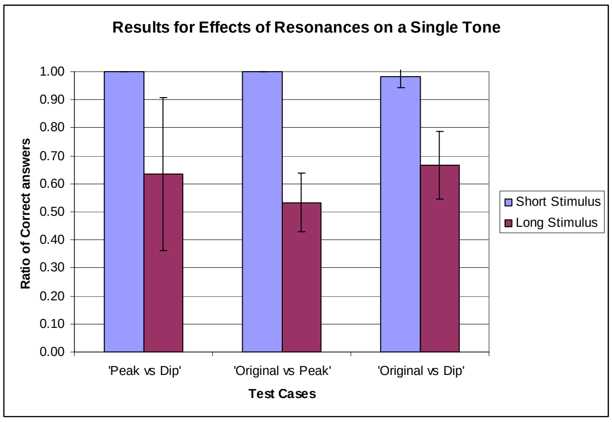

Figure 11 – Results for subjective tests

	Short Peak vs Dip	Long Peak vs Dip	Short Orig vs Peak	Long Orig vs Peak	Short Orig vs Dip	Long Orig vs Dip
p	0.000	0.087	0.000	0.945	0.000	0.441

Table 1 – Chi-square test for significance of difference detection

The results for the short stimulus are all highly significant ($p<0.001$), whereas those for the long stimulus are not significant ($p>0.05$). This clearly shows that subjects could reliably detect differences in samples when presented with the short stimulus, but this ability was lost when the same situations were presented with a long stimulus.

4.2 Effects of Resonance Response

This section analyses the results according to the effect of the resonant response in each system. It has already been shown that the differences are only reliably detected when in the presence of a fast transient stimulus. Three comparisons are studied.

The first comparison concentrates on the effects of a Peak, created by a single resonance, contrasted to a Dip, generated by two adjacent resonances. The cases are depicted in curves (a) and (c) of *Figure 1*.

In the presence of a short stimulus, subjects could reliably detect the differences between the two systems even though they are presented at the same amplitude. When using the long stimulus the rate of correct responses is much lower. The probability level for this case with the long stimulus is almost significant at p=0.087. This suggests that the differences between Peak and Dip may be detectable in the presence of a long stimulus, but the task is severely hindered by the temporal structure of the signal.

Original input stimuli were then compared to two modified cases (Orig vs Peak; Orig vs Dip). Again results show opposite outcomes depending on the type of input stimulus. When in the presence of a fast transient short stimulus, subjects are able to detect the effects of the resonances imposed on the original sample. Considering that all samples are presented at the same amplitude, it can be concluded that the perception of the effects of resonances on single tones is not level dependent. Hence, even when auditioned at the same amplitude, the resonances impose a character on the sound that is clearly perceived by listeners.

5. DISCUSSION

In view of these results, there is a strong dependence of the audibility of resonances on the temporal structure of the input. A clear effect confirms that detection is highly dependent on the time content of the input stimulus. In all cases studied, there is highly significant evidence to support that the effects of resonances are best noticed when fast transient inputs are present. Furthermore, it may be argued that the temporal characteristics of signals whose amplitude envelope varies slowly, interferes with the ability to detect effects introduced by resonances.

The different resonance situations commonly found in rooms (described here as *Peak*, *Valley* and *Dip*) do generate different perception, although this is most clearly revealed in the presence of impulsive signals.

The effects of resonant frequency on single tones may be categorised into two main areas.

When the input signal matches the resonant frequency, there is an increase of the attack and decay times of the output signal. The interaction between input and system is mostly noticeable by an extension of the decay time. However, in the presence of input stimuli with fast onset transients, a reduction in the attack time is also noticeable. The resonant effects, after the input has been removed appear at the same frequency of the stimulus which makes them less prone to detection.

When the system is comprised of two resonances and the input stimulus falls near these, the result is an interaction between all components in the system. If the frequencies of these components are close enough in frequency, there is evidence of ‘beats’ or fluctuations in the amplitude. The audibility of these effects is dependent on the amount of energy present in the transient portions of the input signal and, especially in the presence of fast transient signals, the effect is clearly audible. This is explained by the following.

Signals with fast transients have inherent energy outside their main frequency component as it was shown in section 2. It is this energy that is capable of exciting resonances near its main frequency.

Smoothly varying signals in time have a much reduced level of energy at frequencies outside their main components. The interaction with the resonant components of the system does exist, but the energy input at the resonant frequencies is of such small amplitude that the effects are masked by the higher amplitude input stimulus. The decay time after the input is switched off is dependent on the decay characteristics of the resonances, however because it starts at such a low level its effects are inaudible.

The perception of differences in the resonant cases is independent of the amplitude level difference between them. This suggests that each resonant system imposes a characteristic sound which depends on the frequency relationship between the input and the resonant frequencies.

6. CONCLUSION

The results presented here indicate that the perception of resonances is highly dependent on the temporal content of the excitation signal. Considering that in most musical examples the signals are comprised of fast transients, the effects of untreated resonances will certainly be audible in most situations. Nonetheless, this experiment and previous work show that there are critical factors that determine the audibility of resonances. Shown in previous research was the fact that resonant problems are only evident when the excitation tone is near or at the resonant frequency(ies). As most music contains a multitude of frequencies, it is reasonable to argue that any distribution of resonances at low frequencies will present problems when excited by music signals. In addition, it was further shown here that if level differences are removed, subjects are still sensitive to effects introduced by resonances. This fact supports the idea that any low frequency correction method should concentrate on reducing the temporal resonant characteristic of the modes. A technique that solely addresses magnitude frequency irregularities is likely to leave evidence of the temporal effects, especially after the removal of the excitation stimulus. In some cases, in the presence of fast tonal transients, the effects may even be detected simultaneous to the input stimulus. Any attempt to reduce this effect is only likely to be successful if the temporal characteristic of the resonance is modified by reducing its decay time. The work presented by the authors on the difference limen of Q-factors may thus be used as a target for successful control techniques.

ACKNOWLEDGEMENTS

FCT – Programa Praxis XXI – Portugal
Marie Curie Fellowships – EDSVS

REFERENCES

[1] Fazenda, B. Avis, M. Davies, W. *The views of control room users* Proceedings of the Institute of Acoustics, U.K. Vol. 23, pt 8, 2001

[2] Bucklein, R. *The audibility of Frequency Response irregularities*. Journal of the Audio Engineering Society, Vol. 29, No. 3, March 1981

[3] Fryer, P. A. *Intermodulation listening tests*, presented at the 50th Convention of the Audio Engineering Society, Journal of the Audio Engineering Society (abstracts), Vol. 23, p. 402, June 1975

[4] Olive, S., P. Schuck, Ryan, J. G., Sally, S. L., Bonneville, M. E. *The Detection Thresholds of Resonances at Low Frequencies*, Journal of the Audio Engineering Society, Vol. 45, No. 3, March 1997

[5] Fazenda, B. Avis, M. Davies,W. *Difference limen for the Q-factor of room modes*, presented at the Audio Engineering Society Convention, New York, USA, October 2003

[6] Fazenda, B. Avis, M. Davies, Jacobsen, F *Perception of modal distribution in critical listening spaces* presented at the 11th International Congress on Sound and Vibration,St. Petersburg, Russia, July 2004

[7] Salava, T. *Low-Frequency performance of listening rooms for steady-state and transient signals*, Journal of the Audio Engineering Society, Vol. 39, No. 11, November 1991

[8] Moore, B. *An Introduction to the Psychology of hearing,* Academic Press, 5th Edition,USA 2003

[9] www.pcabx.com