Audio Engineering Society

Convention Paper

Presented at the 121st Convention
2006 October 5–8 San Francisco, CA, USA

Investigation of hearing loss influence on music perception, in auditoria, by means of stereo dipole reproduction

Andrea Capra[1], Marco Binelli[1], Daniela Marmiroli[2], Paolo Martignon[1] and Angelo Farina[1]

[1] Industrial Engineering Department, Università di Parma, Viale Usberti 181/A, 43100 Parma, Italy
andrea.capra@ied.unipr.it

[2] Faculty of Medicine, Università di Parma, 43100 Parma, Italy
daniela.marmiroli@studenti.unipr.it

ABSTRACT

The largest part of who seats in theatres or in auditoria has not an optimal perception of sound, because of hearing loss. So, in order to find out a correlation between objective parameters and subjective descriptors, executing therefore meaningful listening tests, we need first to study the perception of the customer. For this purpose, selected theatregoers, different for age, sex and degree, were chosen as subjects. The listening test was based on the virtual spatial recreation of several theatres, by means of an optimized stereo dipole technology. The test was repeated, for 30 subjects, with and without hearing aid, which was previously set to compensate the auditory loss. Some preliminary data analysis results are here shown.

1. INTRODUCTION

Getting older, our auditory system looses functionality: the level and the frequency range of the perceived sounds decrease. A large part of people who habitually assist to performances of lyric and classical music has an age older than forty years: it is therefore assumed that these persons have a loss of perception in high frequencies, beside other typologies of auditory disturbances. In most of cases, who has consciousness of his deficiency wears hearing aids, but they are set for a better intelligibility of speech. Because all of these things the largest part of who seats in theatres or in auditoria don't perceive correctly the sound generated from the performer (singers, orchestra, single musician, etc...). In order to find a possible correlation between objective parameters and subjective descriptors, in order to perform meaningful listening tests, we need to comprehend what is the perception of the medium customer and take it into account. The first step of our research was the distribution of a questionnaire in Teatro Regio (Parma) and Auditorium Paganini (Parma) to know the age, study degree, preferences, and degree

of music knowledge. The second step was a listening test on the subjects recruited with the mentioned questionnaires. The tests were performed in "La Casa della Musica" (Parma – Italy), in a listening room with an acoustic treatment, using Stereo Dipole as reproduction system. The complete test was composed by three parts: an audiometric test, a listening test without hearing aids and the same test with hearing aids, set with a correction only in the case of a hearing deficit. Therefore this paper shows the preparation, the development and the results of this session of listening tests.

1.1. Hearing Loss

It's an assumption that, with the increasing in the age, our auditory system looses functionality: this process is called presbycusis. There are many causes of it: most commonly it arises from changes in the inner ear of a person as he or she ages, but it can also result from changes in the middle ear or from complex changes along the nerve pathways leading to the brain. Presbycusis most often occurs in both ears, affecting them equally. Its level depends also from the life conditions: for example it is accentuated in the inhabitants of towns, whilst it is lower in people who live in the country. The cumulative effects of repeated exposure to daily traffic sounds or construction work, noisy offices, equipment that produces noise and loud music can cause sensorineural hearing loss. It means that the increasing of sound pollution is one of the most important causes of deficit in auditory capacity.

The main problem has to be found in the incapacity of people to recognize their auditory problems: the loss is gradual, the brain gets used to its new equalization and the person is convinced of his capability in hearing sounds.

A hearing aid could compensate the deficit, but the correction must be improved gradually in order to habit the ear in hearing newly the lost frequencies. The aid is essentially a Lo-Fi amplifier because the signal processing is done in a limited frequency band, which is the band typical of the speech. The basic schematic for an aid is made of a microphone, an amplifier, a receiver, a manual volume control and electric alimentation. The new generation of auditory aid permits the setting of multiple equalizations in the internal memory and has a higher sound fidelity. To better support also music listening.

1.2. Binaural Technique

The binaural technique is based on recordings made with dummy head or omnidirectional microphones placed into the one's ears: the principle is to replicate the way our ears capture the sounds and replay those sounds directly in the corresponding ears. In origin, the only way to reproduce this type of recordings was wearing headphones. Headphones method is fairly realistic and straightforward, but, as well known, suffers from the following main problems: head locking, front-rear indetermination (strongly related to the previous one), and of course the artificial feeling of physically wearing the headphones themselves. In order to overcome part of these problems, the stereo dipole seemed to be a good alternative, for the presence of two loudspeakers in spite of the headphones. This technique consists in deleting the cross talk: the sound on left monitor only arrives on the left ear, the sound on the right monitor only reaches the right ear and the cross-path sound is mostly canceled out.
Recent studies of D. Cabrera [2] point out through listening tests the strong realism of the stereo dipole in comparison with headphones and normal stereo: the listener seems to appreciate this binaural reproduction system, finding in it an accurate spatiality. Unfortunately the setting of the stereo dipole isn't so simple: the cross talk cancellation has to be implemented by an accurate inverse filtering showed in next chapters.

2. METHOD

2.1. Auditoria and impulse response measurements

For this particular session of listening tests we used impulse responses recorded by Angelo Farina and his staff. Measurements had been made using a dodecahedron with a subwoofer for the bass frequencies, as the source for the test signal: it was placed on the left and on the right on the stage. The test signal was an exponential sine sweep; equalization was applied to the signal for a constant spatially averaged output power from the loudspeaker. A Neumann KU100 dummy head was used as the binaural microphone. The distance between source and receiver was about 14 m in each room and the receiver was 1 m off the longitudinal axis of the room. [1]
For this test we used two opera houses, Teatro Regio (Parma, Italy) and Teatro Valli (Reggio Emilia, Italy),

two comparable for dimension auditoria, Sala Petrassi (Parco della Musica, Roma, Italy), Auditorium Paganini (Parma, Italy), and an historic theatre, Olimpico Thaeatre (Vicenza, Italy).
From the impulse responses were extracted objective parameters in according with ISO 3382: reverberation time, clarity, early decay time, definition, bass ratio, baricentric time, tonal balance, lateral fraction and inter aural cross correlation.

2.2. The listening room

The place chosen for the listening test is a room with a floor 4.5 m x 3.2 m, with a ceiling height of 4.2. It was treated with absorbing panels placed on wooden or iron structures all around the listener. Absorbing panels in polyester fiber with high density were placed below the ceiling at a medium distance of 1 meter, with oblique orientations, mainly for wide band absorption on bass frequencies. With this set up the reverb time of the room is less than 0.1 sec above 1 kHz and comprised between 0.5 sec and 1.3 sec below 125 Hz (one should speak about "space averaged decay time" in this case). The ceiling treatment has given a sensible improvement to the problem of low frequency resonances.

Two monitors, Genelec S30D, were used for the stereo dipole (Figure 1) with a distance between the tweeters of 18 cm, a distance between the mid-ranges of 40 cm and a distance between woofers of 76 cm; the opening angles were respectively α=7°, β=14° and γ=30°.

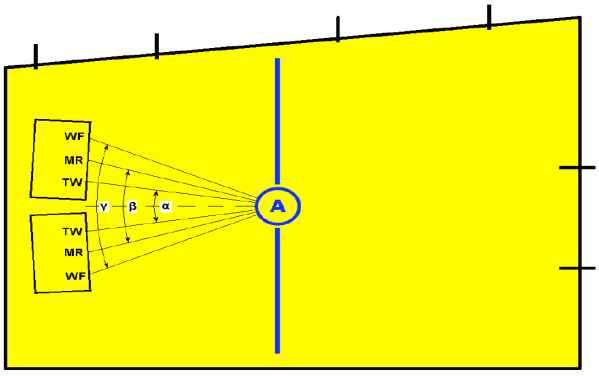

Figure 1 - Layout of the stereo dipole reproduction system

2.3. The stereo dipole

Stereo dipole aim is to recreate the correct sound signal at the two listener ears through a system of two loudspeakers, each fed with a processed version of the original binaural recordings, exploiting the technique of cross talk cancellation.
This technique (see Figure 2) uses a two by two matrix of (four) filters, calculated so that the system cancels the contribution of the left speakers to the right ear and viceversa, that is eliminates the cross-talk,. This matrix H is obtained inverting the original matrix of transfer function from speakers to ears C previously measured.

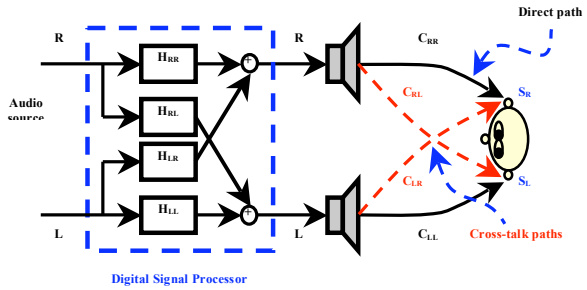

Figure 2 – Scheme for the stereo dipole concept.

The formula employed for such an inversion is a multichannel inversion proposed by Ole Kirkeby:

$$\mathbf{H}(k) = \frac{\mathbf{C}^H(k)\mathbf{A}(k)}{\left[\mathbf{C}^H(k)\mathbf{C}(k) + \beta\mathbf{I}\right]}$$

Where C is the 2X2 matrix containing the four measured direct field complex transfer functions from speakers to a dummy head ears, H is the 2X2 matrix of calculated inverse filters spectra (also complex), A is a 2X2 diagonal matrix of desired final transfer function from speakers to ears; β is a "regularization" parameter. For each single frequency considered, the regularization parameter helps to get a smoother inversion when the linear system represented by the matrix C is "ill conditioned". This in practice tends to limit ringing artifacts, hence to obtain causal and not too long filter impulse response. An inverse Fourier transform provides for the definite FIR form of the inverse filters.
In our case, a particular version of the inversion technique has been implemented which provides for cross talk cancellation but not for any equalization. This allows to separate two targets, cross-talk cancellation

and equalization, which, even taken individually, are quite complex.
Hence, in this case, the filtering aims at maintaining unchanged the measured transfer function left to left and right to right, and to eliminate the cross-talk; this formally means

$$\mathbf{A} = \begin{vmatrix} C_{LL} & 0 \\ 0 & C_{RR} \end{vmatrix}$$

Two main problems which limit the possibility of a perfect cross talk cancellation can be lighted:

- At low frequency the phase difference between the elements of C are very little, which is a clear case of "ill conditioning".
- At high frequency a correct cancellation may occur, but only in a very precisely fixed geometrically condition, precision which is definitely not achieved when a person is placed instead of the dummy head.

The former problem can lead to a huge power demand on the speakers, with consequent speaker bad working and too much energy injected in the room (which will never be well absorbing in this frequency range). The latter can lead to spatial comb filtering artifacts, perceptible with little movement of the head, phenomenon which partially corrupts the listening experience.
A tool which can limit these effects is to make the regularization parameter frequency-dependent, and "pump" it in the extreme part of the spectra.
This kind of inversion was done thanks to the Aurora plug-in "Invert Kirkeby", written by Angelo Farina: the in-band and out-band regularization frequency limits were 200 Hz and 10kHz; the length of the calculated FIR were 256 taps.

A physical means, which helps in improve both the problems listed above, is the speaker disposition described at the previous paragraph, that is to have low frequency transducers in wider positions than high frequency ones.

The four cross-canceling FIR convolutions were software implemented by means of the plug-in Voxengo "Pristine Space", hosted and routed inside "Plogue Bidule".

At this point the "left to right separation" sounded good, but still remained the problem of equalization.
In particular, the need of equalization is due to

- The room effect, which is not well absorbing at low frequencies as much as at medium-high ones
- The frequency-dependent regularization effect, which can cause the in-band and out-band ranges to be unbalanced.

Such equalization was done manually, with a half-octave equalizer (Voxengo "MarvelEq"), feeding the stereo dipole with pink noise and monitoring the dummy-head ears signal with a half octave spectrum analyzer (Voxengo "SPAN"). All this was done inside the same single Plogue Bidule session. The technique employed, frequently used in recording studios or live situations, implies a stationary equalization, that is an action on the whole impulse response (not only the part integrated by the ear), and does not provide for phase equalization. It is a quite approximated but stable approach (since bases on half octave band-pass filters), and in this case, as well as in many others, gives a straightforward compensation for macroscopic spectral lacks; our ear tests have confirmed the definite improvement. A target for the future is to find a solid method for inverting magnitude and phase of a suitable portion of the impulse response, perceptively speaking and depending on frequency.

2.4. The session

2.4.1. The recruiting

To have a representative sample of the theatergoers a questionnaire was prepared in collaboration with Teatro Regio Foundation: the questions were about age, number of concerts or operas attended per year, degree and possible musical studies. There was also the request of a judgment on own auditory capacity and the availability for the listening test. The questionnaires were distributed during opera and concert season, between 13 January 2006 and 18 March 2006, in Regio Theater and Paganini Auditorium (Parma); the compilation was made principally inside the theater. 912 questionnaires were collected with a large availability for the test: we chose 30 subjects distributed in age in order to be representative of the listener population (Figure 3).

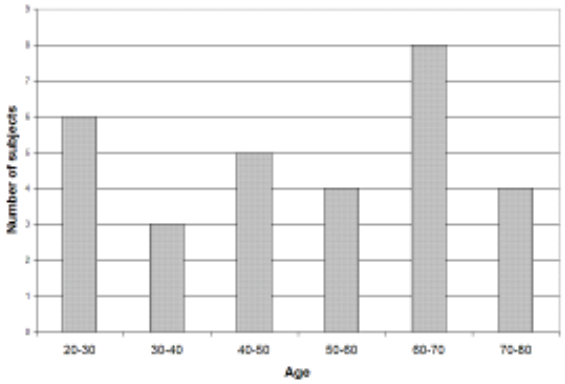

Figure 3 - Distribution in age of the tested subjects

2.4.2. The test

The test was divided into three parts. The first was an audiometric test: its aim was the characterization of the subjects to find out possible auditory problems. It was performed from an expert technician, one of the authors of this paper, using a portable audiometer and isolated headphones. The data collected were used for the setting of auditory aids that have been wearing in the third part of the test. The auditory test had the duration of 10 minutes for each person that was tested in the range of frequencies from 125 Hz to 8 KHz.

The second part of the test was the evaluation of the acoustical characteristics of the five rooms previously mentioned. It was performed through the listening of an anechoic track (overture of "Le nozze di Figaro" performed by Osaka Philarmonic Orchestra, Denon) convolved with the impulse responses of the rooms. The test was made using a dedicated software (Figure 4).

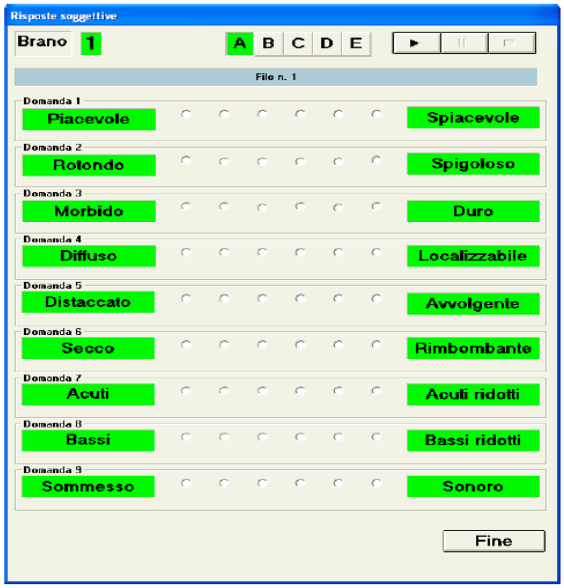

Figure 4 - Electronic form for the test

Watching the monitor placed in front of the listening point, the subject under test had the possibility to switch in real time from a room to the another, without interrupting the play of the track; this is very important because of a too scant acoustical memory of our brain. The evaluation was made through nine couples of adjectives, identified in literature as the most suitable for this type of tests: "pleasant-unpleasant", "round-sharp", "soft-hard", "diffuse-localizable", "detached-enveloping", "dry-reverberant", "treble boosted-treble reduced", "bass boosted-bass reduced", "quiet-loud". The vote was from 1 to 6 in order to give a weighted preference to the left adjective (from 1 to 3) or to the right adjective (from 4 to 6). The listener had no limit in time, modifying the judgment how many times he preferred and finally saving in a file his result. Only for the oldest persons with big difficulties with the mouse of the computer, was prepared a paper questionnaire. In order to avoid conditioning, the test was "blind", therefore the subject under test was unaware of which room was listening.

The third part of the test was the same evaluation of the previous part wearing an "Oticon Tego" (Figure 5) hearing aid on both ears, set to compensate the loss

evidenced by audiometric test. None of the subjects refused the aids.

Figure 5 - Oticon Tego hearing aid

3. RESULTS

3.1. Questionnaires

From the analysis of the 912 questionnaires emerges a photograph of the listeners who attend to the concerts and operas in Parma, a little town of 170,000 inhabitants with a high musical culture.

Females represent the 56% of theatergoers; the majority of the listener, the 78%, has no musical studies in their background; the 56% of them prefers the season ticket to the single one, synonymous of an assiduous frequentation. The percentages in Figure 6 show that the educational level of theatergoers is quite high: the 44.4% has a university degree in spite of a low percentage (7.8%) that attended only the primary school.

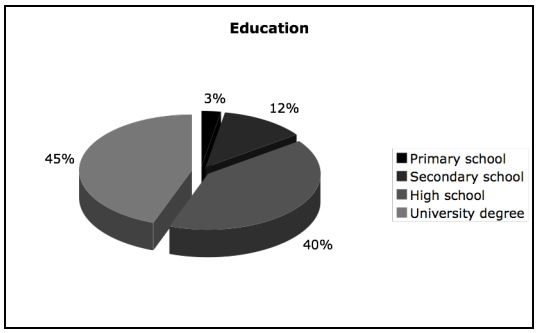

Figure 6 - Level of education in theatergoers

In Figure 7 and Figure 8 is shown how many operas and concerts are attended per year: from the data collected emerges a good frequentation of both.

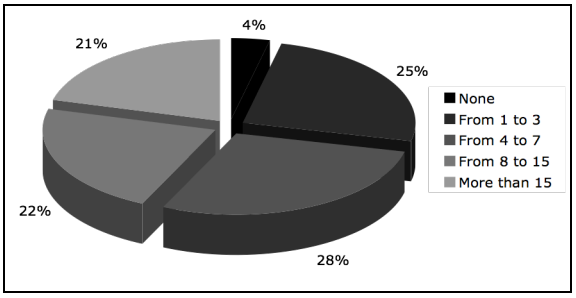

Figure 7 - Number of concerts attended from the medium listener per year

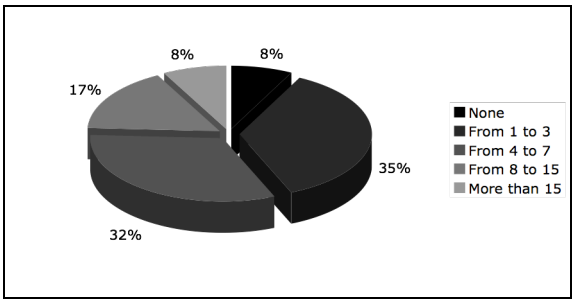

Figure 8 - Number of operas attended from the medium listener per year

3.2. Audiometric test

The results of each audiometric test were collected and averaged across subjects, taking into account the different ages (Figure 9).

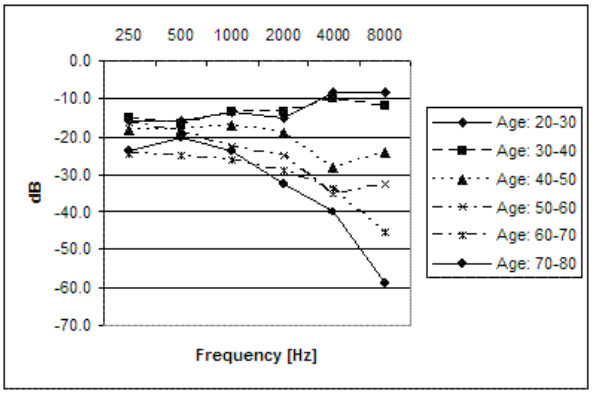

Figure 9 - Auditory loss for every class of age

From the analysis of these results emerges that the first two classes of age has no evident loss in particular band

of frequencies: they present only a small generic loss of level, lightly emphasized on basses. This could be explained thinking that the test was made in winter and part of the subjects might be victims of the colds of the season. Only one subject presented an auditory test perfectly linear. The high frequencies loss is proportional to the increasing in the age, with a heavy deficit for subjects from 70 to 80 years. This deficit is probably due to the combination of our natural loss in auditory capacity with the increase of sound pollution. This result, matched with most of the results in literature [xxx] shows that a big part of regular theatergoers, the 59.3%, doesn't ear correctly the sound of an orchestra: they are unable to ear with the right perception the notes of a piccolo, the last octave of a piano or violins. This could be a problem for the evaluation of the acoustic of a theater.

Most of the theatregoers were sure of their capability in hearing every frequency correctly: the 85.7% declared a normal auditory capacity, the 11.9% declared a slightly damaged capacity and only the 2.4% supposed an hard damage of his capacity. The real capacity, showed from the audiometric test (again Figure 10 and 11), is quite far from the previous percentages: the 16.7% of subjects has a hardly compromised auditory system on the 4000 Hz and even the 36.7% has a big deficit on the 8000 Hz. Is emblematic the case of a subject, suspended 15 years ago from his job for hypacusia due to an auditory trauma, who declares that he hears "every sound, also the highest notes of violins played very quiet"; this is impossible because his audiometric test reveals a big loss, from 40 dB to 70 dB, on medium-high frequencies.

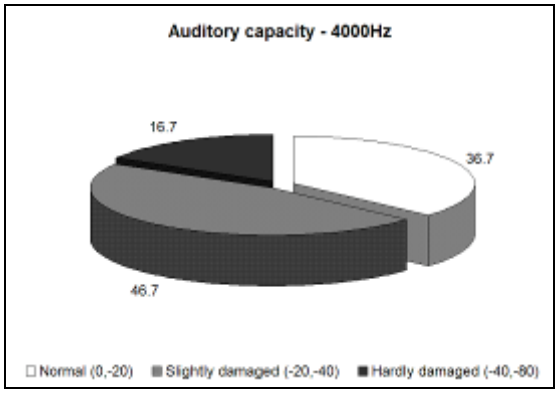

Figure 10 - Auditory capacity of the subjects tested on the band of 4 KHz

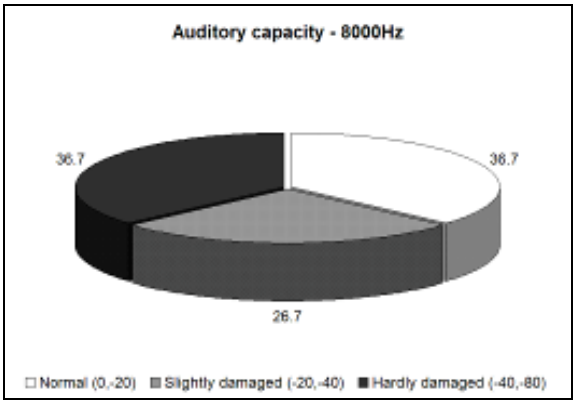

Figure 11 - Auditory capacity of the subjects tested on the band of 8 KHz

3.3. Listening test

For the analysis of possible correlations between subjective descriptors and objective parameters of acoustics, nine parameters for every room, showed in Table 1, were chosen: reverberation time (T20 and T30), clarity (C50 and C80), definition (D50), energy barycentric time (Ts), early decay time (EDT), bass ratio (BR), tonal balance (TB), inter-aural cross-correlation coefficient (IACC) and lateral fraction (LF). Prof. Angelo Farina and his staff measured all the values from 2002 to 2005.

	A.P.	T.O.	T.R.	R.700	T.V.
C50 [dB]	2.00	1.21	2.86	-1.51	0.51
C80 [dB]	3.92	2.94	6.49	1.30	6.37
D50 [%]	54.36	56.24	59.89	41.25	52.45
Ts [ms]	110.64	94.49	75.12	122.51	67.52
EDT [s]	2.12	1.77	1.34	1.71	1.15
T20 [s]	1.99	1.86	1.32	1.80	1.36
T30 [s]	2.03	1.85	1.28	1.84	2.08
LF	0.48	0.44	0.41	0.62	0.44
IACC	0.49	0.42	0.57	0.40	0.65
TB	0.08	-0.04	0.11	0.14	0.13
Bass Ratio	1.13	0.77	1.34	0.99	1.33

Table 1 - Objective parameters (A.P.=Auditorium Paganini, T.O.=OlimpicoTheatre, T.R.=Regio Theatre, R.700=Sala700, T.V.=Valli Theatre)

A first comparison between the two test (with and without auditory aid) on the descriptor "pleasant-unpleasant" shows as, in subjects with a "normal" auditory capacity, the aid doesn't modify the judgment: the only little gain doesn't seem to affect the right perception of sounds.

From a multivariate analysis, performed with software SIMCA P11, comes out a correlation of "pleasant-unpleasant" with four subjective descriptors (Figure 12). The four couples, "soft-hard", "round-sharp", "detached-enveloping" and "dry-reverberant", seems to be the most important in forming a judgment on the pleasantness. This result is strengthened from its observation in both cases, with and without auditory aid.

For the statistic analysis the Last Squared Method was used, finding out a linear regression coefficient for every objective-subjective parameters couple. The absolute value of 0.3 was chosen as a limit between a sufficient and a bad correlation.

The parameter Tonal Balance, the balance between middle and bass, presents a lot of correlations in the absence of hearing aids but, with the aids, some of those correlations, in particular with the couples "treble boosted-treble reduced" and "bass boosted-bass reduced", disappear. This could be caused by the compression of the dynamic inside the aid. One of the most significant results is that an increase of Tonal Balance seems correlated with the couple "round-sharp" in both cases, with aid (r = -0.39) and without aid (r = -0.36) (Figure 13). The same objective parameter affects the pleasantness (r = -0.33) wearing the aid: the overemphasizing of bass frequencies in subjects, which are not used anymore in such equalization, makes unpleasant the sound.

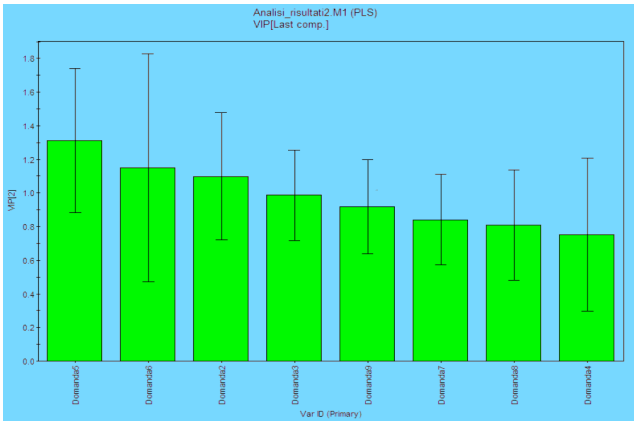

Figure 12 - Scale of importance of the subjective descriptors contributing to "pleasant-unpleasant" (with reference to Figure 4)

Another relevant objective parameter is the Baricentric Time (Ts): its improving seems correlated with the sensation of bass accentuation (r = -0.39). The reverb time T20 shows a correlation only with "quiet-loud" (r = -0.31) and only in the case of absence of auditory aid.

We noted unexpected results as the lack of correlation between the two spatial parameters Lateral Fraction and IACC with the subjective descriptors "diffuse-localizable" and "detached-enveloping" or as the small influence of reverb parameters on subjective judgments.

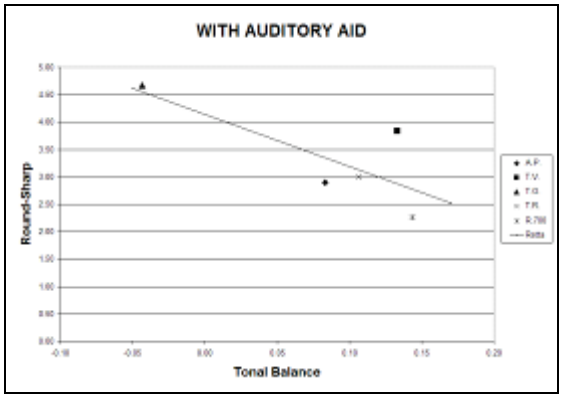

Figure 13 - Linear regression of Tonal Balance and "round-sharp", wearing auditory aid

4. CONCLUSIONS

For listening tests with a good relevance, the fidelity of reproduction system is essential. In the stereo dipole technique the treatment of the listening room and the accuracy of inverse filtering become fundamental for the realism of the recordings.

The listening test points out the importance of taking into account auditory capability of theatergoers for an accurate evaluation of theaters, concert hall and listening room: people are sure of their capability in hearing sound but, in most of the cases, their deficit, due to auditory traumas or natural aging, doesn't allow a correct perception.

The test must be improved with a more accurate statistical analysis because of a too scant linear correlation between single subjective and single objective parameters. The next step will be the research on these data, with a multivariate analysis, of a mathematical model representing the descriptor "pleasant-unpleasant" in dependence of objective parameters.

5. ACKNOWLDGEMENT

This work was also supported by Andrea Cammi, Andrea Azzali, "La Casa della Musica", "Fondazione Teatro Regio", the staff of Auditorium Paganini and Teatro Regio.

6. REFERENCES

[1] A. Farina, R. Ayalon, "Recording concert hall acoustics for posterity", 24th AES Conference on Multichannel Audio, Banff, Canada, 26-28 June 2003.

[2] A. Azzali, D. Cabrera, A. Capra, A. Farina, P. Martignon, "Reproduction of Auditorium Spatial Impression with Binaural and Stereophonic Sound Systems", 118th AES Convention, Barcelona, 28-31 May 2005.

[3] A. Farina, L. Tronchin, "Acoustic qualities of theatres: correlation between experimental measures and subjective evaluations", Applied Acoustics 62 (2001) 889-912.

[4] O. Kirkeby, P.A. Nelson, H. Hamada, "The stereo dipole – a virtual source imaging system using two closely spaced loudspeakers", J. Audio Eng. Soc., 46, 387-395 (1998).

[5] S. G. Norcross, G. A. Soulodre, M.C. Lavoie, "Evaluation of Inverse Filtering Techniques for Room/Speaker Equalzation", 113th AES Convention, Los Angeles, 5-8 October 2002.

[6] E.Alajmo – "Otorinolaringoiatria" - Piccin Nuova Libreria, Padova, p.19