Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8   New York, NY, USA

Listening tests of the localization performance of Stereodipole and Ambisonic systems

Andrea Capra[1,2], Simone Fontana[1,3], Fons Adriaensen[1] , Angelo Farina[1,2] , and Yves Grenier [3]

[1] LAE GROUP, Parma, 43100, Italy
andrea.capra@mac.com

[2] Industrial Engineering Department, Università di Parma, Viale Usberti 181/A, 43100, Parma, Italy

[3] Département TSI, Ecole Nationale Supérieure des Télécommunications, Paris, France
fontana@tsi.enst.fr

ABSTRACT

In order to find a possible correlation of objective parameters and subjective descriptors of the acoustic of theatres, auditoria or music hall, and so to perform meaningful listening tests, we need to find a reliable 3D audio system which should give the correct perception of the distances, a good localization all around the listener and a natural sense of realism. For this purpose a Stereo Dipole system and an Ambisonic system were installed in a listening room at La Casa Della Musica (Parma, Italy). Listening tests were carried out for evaluating the localization performances of the two systems.

1. INTRODUCTION

Correlation between subjective descriptors and objective parameters is fundamental for the design of new theatres and concert halls ([1], [2]): if we knew that a positive judgment on the acoustic of a theatre corresponds to a set of values of principal acoustical parameters we would be able to project new theatres close to the listeners expectations. This is one of the aims of research for reliable systems able to reproduce the right acoustical perception of a room. Our acoustical memory is quite short and the only way for an accurate evaluation of different acoustics is placed in auralization technique: in fact a real-time switch between theatres can give evidence to small shades, evidence that will be even clearer if the audio system has the capability of replicates a given sound field with some detail without loosing the naturalness and the overall respect of other perceptual parameters.

Stereo dipole [3] and Ambisonic [4] systems are promising techniques for sound field reproduction. If the systems and the techniques underlying are quite known nowadays, namely thanks to the studies at ISVR (see for instance [11]) and the Ambisonic Community (see for instance [6], [12]), only few studies compared the two approaches. In a first study the two systems have been compared in terms of subjective parameters ([7]). In this study we focus on objective parameters, such as localization, and in this sense we extend the study in [8], using a 3D first order Ambisonic System and a Triple Stereo Dipole. These systems allow for 3D reproduction, and make possible a comparison of localization accuracy not only in term of azimuth, but also of elevation. In the literature a number of localization tests have already been performed on each systems separately [11], [9].

The paper is organized as follows: in section 2 we briefly present the listening room and its acoustic treatment; in section 3 the physical equipment for performing the test. In section 4 we present the Ambisonic and Triple Stereo Dipole implementation. Section 5 describes the test, and section 6 presents the results and some observations.

2. THE LISTENING ROOM AND ITS ACOUSTIC TREATMENT

The room chosen for the test and situated in “La Casa della Musica” (Parma) is a not-perfect parallelepiped of 455x300x425cm (Figure 1).

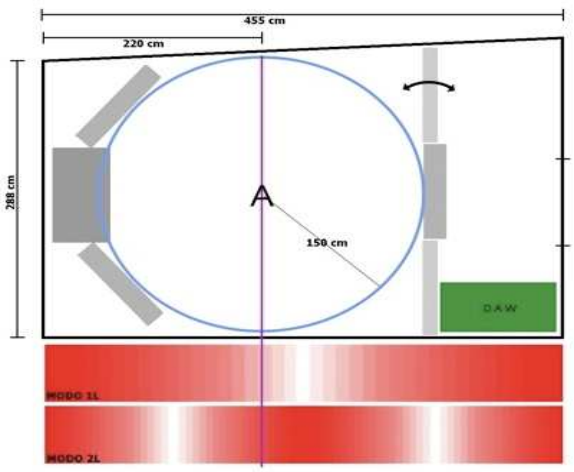


Figure 1 Dimensions of the room in relationship with two of the first modes

The position of the listener takes into account the modal distribution of the room. In the treatment of a small room like this, one of the main problems is the behavior of low frequencies. A high number of absorption panels made of polyurethane were placed along the walls but their action is effective only on the medium-high frequencies. In order to improve the absorption of low frequencies, panels of polyester fiber were placed on the ceiling, creating useful cavities for this aim (Figure 2).

Some Helmholtz resonators, made of big plastic water bottles, were built and placed in the bottom of the room in order to counteract the first longitudinal mode at 37.8 Hz; tuning of the resonance frequency was done by varying the dimension of the neck and adding absorbing materials inside the plastic bottle. Another kind of resonator, perforated wooden boxes filled with polyester fiber, is used for absorption in the range 100-300 Hz.

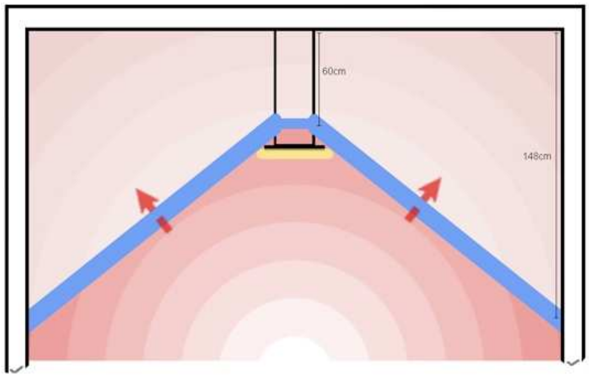


Figure 2 Cavities in the ceiling for low-frequency absorption

Considering all the acoustic treatments, medium-high frequencies Reverberation Time is 0.4-0.1 seconds, low frequencies RT is between 1.1 and 0.4 seconds in the range 31.5 Hz - 125 Hz (Figure 3).

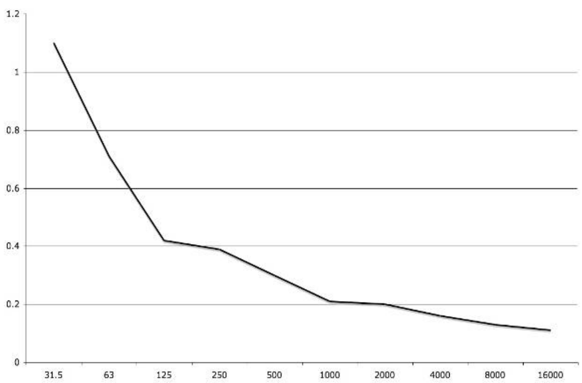


Figure 3 Reverberation Time (T30) of the room with the acoustic treatment.

3. AUDIO SETUP

The system is driven by an Asus S-Presso PC (Pentium IV, 2 Ghz) running Linux, with an RME HDSP MADI soundcard. The MADI protocol permits the use of a large number of channels; in our layout there are 24. The MADI soundcard is connected to RME ADI 648 that converts the MADI signal into ADAT; this digital signal is converted to analog by two converters, an Apogee 16 DA for the Ambisonic system and a Behringer ADA 8000 for the stereo dipoles. These drive three QSC multichannel amplifiers. For the Ambisonic system we use 16 Turbosound Impact 50 speakers, and for the triple stereo dipole two QSC AD82S (rear), two Impact 50 (above the listener) and a couple of Genelec S30D (front). In figure 4 we report the layout of the system.

The speakers were placed as shown in Figure 5.

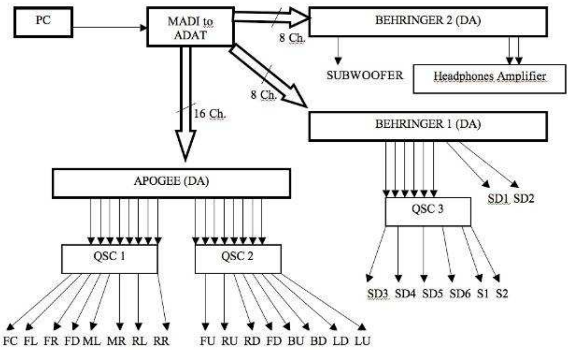
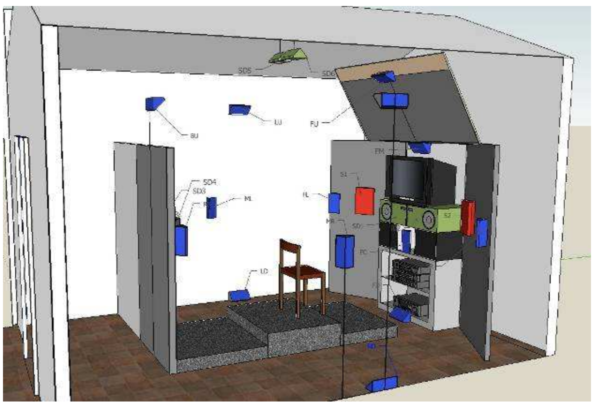


Figure 4 System Layout

The listening point is in a central position: the planar ring has a radius of 1.42 m, the top-ring has a medium distance of 1.60 m, the lower-ring has a medium distance of 1.80 m, the frontal stereo dipole is 1.44 m from the listener, the rear stereo dipole 1.30 m from the listener and the top stereo dipole is 1.30 m above the head of the listener.



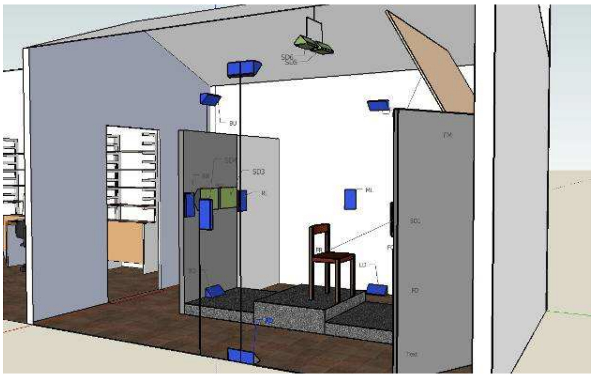


Figure 5 (both images) Ambisonic system (blue), triple stereodipole (green) and normal stereo (red) inside the model of the listening room.

In order to have an equal level on all the speakers, a calibration was made using pink noise, reading the level on a phonometer with the microphone placed in the listening point and adjusting the gain of the amplifiers

4. 3D AUDIO SYSTEMS

4.1. Ambisonic system

Ambisonic theory bases the sound field reconstruction on the decomposition in spherical harmonics, in dependence of frequency and distance from the listening point. In contrast with the normal stereo technique, that has not the ambition of a faithful and physical reconstruction of the sound event, Ambisonic reproduction uses an array of loudspeaker to reconstruct the sound field in a limited “sweet spot” at the center of the array [9]. In order to have an Ambisonic system that permits a good localization of the sound source and the right perception of the distance, we have to consider the *metatheory* of auditory localization [5]. Gerzon pointed out that the best localization for an array of loudspeakers occurs when the magnitude of the reconstructed velocity vector is set to unity at low frequencies and the magnitude of the energy vector is maximized at middle frequencies, with the transition between the two regions somewhere between 300 Hz and 700 Hz [9]. This means that a proper Ambisonic decoder should either have shelf filters operating on the B-format input signals, or use crossover filters feeding two separate decoding matrices.

The two vectors should also have the same direction, as they have for a real source. While the velocity vector can be recreated exactly, a unity energy vector is possible only for a virtual sound source that coincides with a loudspeaker position. In order to avoid a marked 'speaker detent' effect, a decoder should have an energy vector magnitude that is a smooth function of direction, rather than trying to obtain the maximum value for certain preferred directions.

While for regular speaker layouts it is possible to derive decoder matrices according to the principles outlined above in an analytical way, no such systematic approach seems to exist for irregular setups such as the one used, at least not for energy vector optimization. To design the decoder matrices we used the *Makedec* tool , see [10], developed by one of the authors (Adriaensen) that provides a visual and interactive simulation of the performance of first and second order Ambisonic decoders (see Fig. 6). This tools computes the LF decoder matrix by pseudo-inversion of the speaker to Ambisonic components matrix. This LF matrix can then be modified if necessary, and is also the starting point for the manual optimization of the HF matrix. The first step in finding a good HF matrix is to adjust the per-order gains of the LF matrix, followed by tuning of the individual polar patterns corresponding to each speaker output.

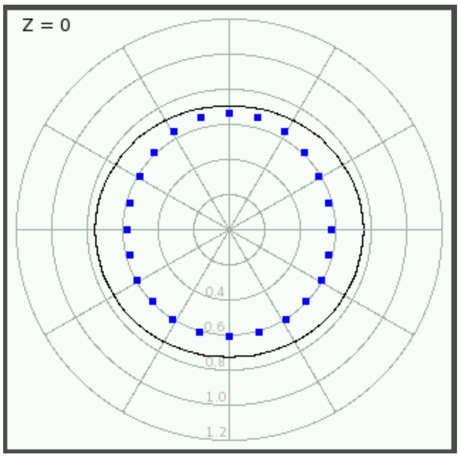


Figure 6. Makedec display of energy vector in the horizontal plane.

The output of the simulation program is a configuration file for the *Ambdec* decoder application. Apart from the implementing the phase aligned crossover filters and the decoding matrices, this program also performs near-field compensation and delay equalization for each individual speaker.

4.2. Stereo Dipole system

4.2.1.General principles

The basic idea of stereo dipole is to implement a transaural system (binaural over loudspeaker) with closed span loudspeaker (around 10 degrees). This arrangement is effective in order to guarantee increased robustness to little head movements ([3]). Recently ([11]) an enhancement of the basic idea has been studied ("Optimal Source Distribution systems") : it concerns the disposition of the loudspeakers, and states that the optimal loudspeaker span must vary continuously as a function of frequency in order to achieve the best performances in terms of system dynamics range and robustness. Namely, higher frequency sources should present little span, and vice versa for lower frequencies. Even more recently ([17]), loudspeaker pairs disposed in the frontal plane above the head have been found to provide better performances compared to pairs with zero elevation. As a general consideration, control loudspeakers with some elevation have been found to be more adapted to the signal processing involved in the SD or OSD implementation; so as frontal positions compared to back positions.

The transaural effect is achieved using a set of "crosstalk cancellation" filters, whose implementation can vary following the type of chosen multichannel inversion algorithm. The involved plant matrix which has to be inverted has been found to be better conditioned when OSD is implemented in the physical framework. From a signal processing standpoint, in order to limit the effect of ill-conditioned frequencies in the direct channels matrix, allowing for inversion robustness and dynamics preservation, least squares with regularization is often used. To operate selective regularization in frequency, the so-called frequency dependent regularization ([13]) can also be used.

Even if theoretically a pair of loudspeaker can be employed to provide a full 3D binaural effect, in practice it has be found to be quite difficult to simulate, for instance, back positions with a frontal only stereo dipole, due to the lack of emitted energy from the back. Back sources are believed to be better perceived when using personalized HRTFs ([14]), but the effect could not be easily generalized to every listener. A possible solution to this problem can be to use multiple stereo dipoles or OSD systems, in order to support back positions. Placing a second stereo dipole behind the listener has been found to enhance perception of back sources. The study of multiple stereo dipoles has not been carried out in a rigorous way yet, but several practical applications ([15], [16]) gave evidence to the usefulness of this kind of system, which could be obviously generalized to support, for example, above-the-head virtual positions.

4.2.2.Implementation

In the system at the Casa della Musica, we implemented a triple stereo dipole, consisting in three stereo dipoles, one in the front, one behind and one above the listener.

The two planar stereo dipoles are composed by horizontally-placed loudspeakers, with bigger span for woofer, and smaller for tweeter, is in accord with "Optimal Source Distribution". The back dipole presents 10 degrees of elevation (higher than the frontal one, which is in the azimuthal plane), in order to deal with the fact that the response at the rear has numerous dips and generally is weaker. Providing some elevation to the control transducers allows for a better plant, as indicated in ([17]).

To perform the plant inversion we used frequency-dependent regularization, with regularization coefficient equal to 0.05 and frequency dependent profile equal to 10 under 200 Hz and 100 above 16000 Hz.. The target function is a delta for the direct paths and the zero function for the crosstalk, the filter length 4096 samples. The inversion algorithm is used to find all the three stereo dipole plant inverse filters, and solve any synchronization problem between the three couples of signals, the resulting deltas being aligned in time.

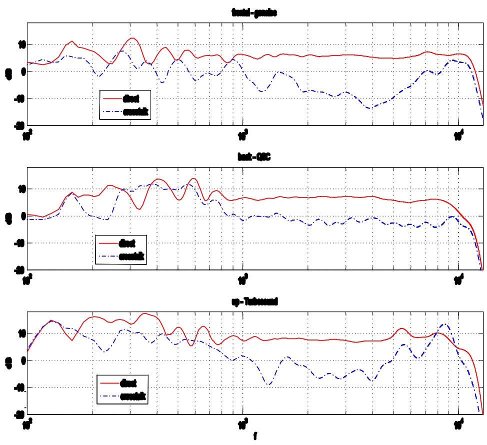


Figure 7 Real crosstalk cancellation result

In figure 7 we show the results for crosstalk cancellation and direct path equalization, which reveals effective namely in the range 700-10000 Hz, with a mean crosstalk cancellation of 10 dB. Remark simulated results shown in figure 8, where crosstalk cancellation reach -80 dB and the mean value around -40 dB.

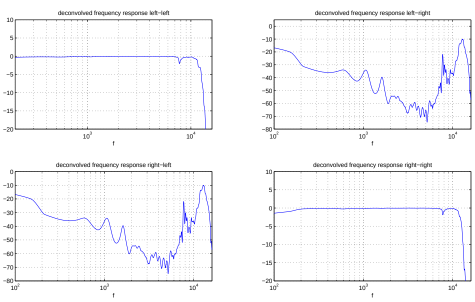


Figure 8 Real crosstalk cancellation result

The real results are affected by noise amplification and spectral shift due to the light time variance of the acoustic channel, as shown in [23]; ongoing studies on this subject carried in our team try to investigate the effect of channel non linearity on real time dereverberation performances.

5. LOCALIZATION TEST

The test consists in the listening of a sound coming from different directions: the subject under test must write azimuth and elevation perceived. For each listener we tested 25 directions (see Table 1) equal for each subject, chosen on a virtual grid with steps of 45° for the azimuth and 30° for elevation, this one limited to -30° under the 0° (Table 1). 21 positions were reproduced by the two audio systems in a pseudo-random way, divided into two sequences: in the first one 10 positions were played by a system and 11 by the other system, vice versa in the second one. Of the 20 subjects chosen, 10 listened at sequence A and 10 at the sequence B. We also inserted in this evaluation the sound coming directly from the speakers placed in 8 different positions (4 in sequence A and 4 in sequence B), with the aim of testing the capacity of the subjects to discriminate in a real sound field, and not in a simulated one, the right direction of arrival; the second aim, not less important, was the reliability of the listeners. At the beginning of the test the subject had the capability to adjust his position inside the sweet spot using a reference signal, positioned at 0° of azimuth and 0° of elevation, played by both the systems. A small plastic sphere, with azimuth and elevation signed on the surface, helped the subjects in finding and writing the angles perceived.

The test signal was a pink noise filtered in the band between 300 Hz and 16 kHz, with duration 130 ms, repeated 12 times in each virtual direction; between one direction and the next one there were 10 seconds of silence. The directions were synthesized for the Ambisonic system using the *VST Gerzonic Panorama* [18] inserted in *Plogue Bidule* [19] host, obtaining 4 channels (B-format) for each position; for the Stereo Dipole they were generated using *Voxengo Pristine Space VST* [20], inserted in *Plogue Bidule* host, using it for the cross-talk cancellation and for the convolution with one set of HRTF chosen in the *Listen* sets [21]. All the positions, according to the sequences played, were placed in the multi-channel DAW Ardour [22].

6. RESULTS

6.1. Presentation format

The way results are presented are inspired on [24]. On the y axis, the 21 virtual positions are plotted with their explicit name, and, on the x axis, the 33 possible positions. The order chosen for the tick labels (‘Back Left’, ‘Back Left Low’,etc) follows the azimuth of the positions, so that it is possible to divide the presentation grid in three vertical and three horizontal slices, as indicated in figure. Three **sectors** have been considered: left and right hemisphere, median plane.

Errors can then be firstly classified in (see figure 9)

1. Right/Left Left/Right Confusion, which is expected to be absent.

2. Median/Right, Right/Median, Median/Left, Left Median, which is expected to be scarcely present

Note that here and in the following, the XX term in a XX/YY confusion is always referred to the simulated position and the YY is the perceived one.

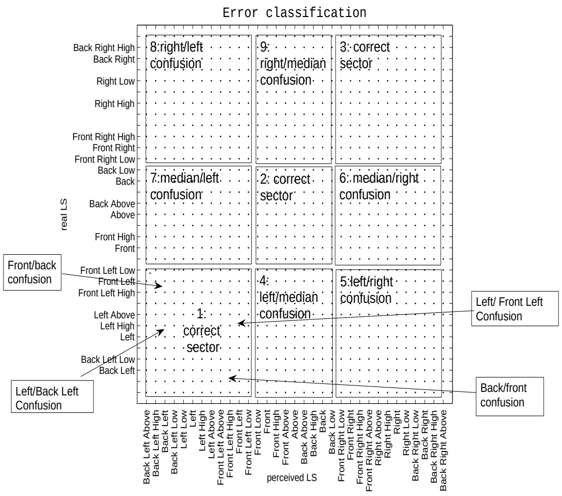


Figure 9 Error classification zones

Errors in each sector can then be further classified in Front/Back Back/Front errors. In figure 9 we showed 2 **zones** corresponding to Front/Back Back Front confusion

A slightly different classification is related to the possibility to group some positions in so-called 'cones of confusion', which can be defined as 'a number of points in three dimensional space that produce the same ITDs and ILDs'. Virtual sources positioned on a cone of confusion can be difficult to distinguish because they need accurate spectral precision in reproduction; on the other hand, sources belonging to different cones are in theory easier to identify. An example of cone of confusion is the median plane, where ITD and ILD are equal to zero. The other cones of confusion are contained in table.

135	-30	'Back Left Low';	left hemisphere	cone 1
135	30	'Back Left High'	left hemisphere	cone 1
45	-30	'Front Left Low';	left hemisphere	cone 1
45	30	'Front Left High';	left emisphere	cone 1
90	-30	'Left Low';	left hemisphere	cone 2
90	30	'Left High';	left hemisphere	cone 2
45	0	'Front Left';	left hemisphere	cone 3
135	0	Back Left';	left hemisphere	cone 3
135	60	'Back Left Above'	left hemisphere	cone 4
45	60	'Front Left Above'	left hemisphere	cone 4
0	-30	'Front Low'	median plane	cone 5
0	0	'Front'	median plane	cone 5
0	30	'Front High';	median plane	cone 5
0	60	'Front Above'	median plane	cone 5
0	90	'Above';	median plane	cone 5
180	60	'Back Above';	median plane	cone 5
180	30	'Back High'	median plane	cone 5
180	0	'Back';	median plane	cone 5
180	-30	'Back Low';	median plane	cone 5
315	60	'Front Right Above'	right hemisphere	cone 6
225	60	'Back Right Above'	right hemisphere	cone 6
225	0	'Back Right';	right hemisphere	cone 7
315	0	'Front Right';	right hemisphere	cone 7
270	30	'Right High';	right hemisphere	cone 8
270	-30	'Right Low';	right hemisphere	cone 8
315	30	'Front Right High';	right hemisphere	cone 9
315	-30	'Front Right Low';	right hemisphere	cone 9
225	30	'Back Right High'	right hemisphere	cone 9
225	-30	'Back Right Low'	right hemisphere	cone 9

Table 1. Simulated positions

A possible error classification following cone of confusion is then between Inter-cone/Intra-cone errors. The first ones include wrong localization of virtual sources in cones of confusions different from their own (ex. Left/Back Left confusion); the second ones include wrong localization of virtual sources in their own cone of confusion (ex. Front Left/Back Left confusion).

6.2. Test results

In figure 10 we plot the results for the true loudspeaker localization. We observe good agreement with most of the positions: five of them are exactly localized by 80% of the subjects, and by 100% within a range of 30 degrees in azimuth or elevation. We believe that this residual error could be eliminated using a more natural way of selecting localization (as, for instance, laser pointing). The back source is badly localized, which is in accord with known perceptual difficulties in localization in the median plane. Back low source presents an abnormal localization for one subject, which nevertheless has not been discarded because of the coherence in the other virtual sources localization. Left high source is strangely badly localized: this effect has been attributed to some critical reflection, but not clearly explained.

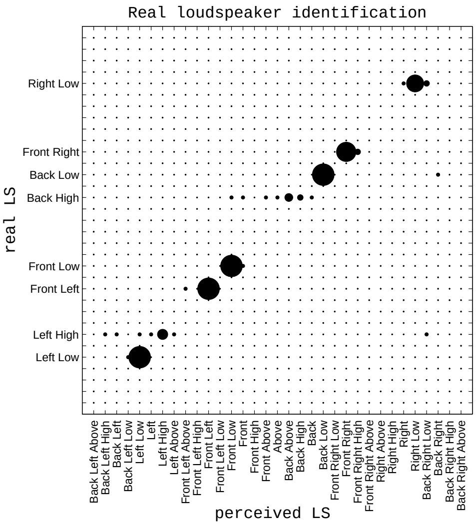


Figure 10 Real loudspeaker localization

We examine now the virtual sources localization test results, as reported in figure 11 and figure 12 , and extract some useful statistics.

First of all the localization of virtual sources in correct sector happened for 93,23% of stimuli for stereo dipole, and for and for 89.05% for Ambisonics (sector 1,2,3). No Left/Right or Right/Left confusion is present in the results, as expected (sector 8,5). Other global results are contained in table 2

Confusion type	**Stereo dipole**	**Ambisonics**
Right/median (sector 9)	0.9%	2.38%
Median/right (sector 6)	0.9%	0.4%
Median/left (sector 7)	0%	0%
Left/median (sector 4)	5.74%	2.85%

Table 2 Test global results

Based on observation of the left/median confusion, these data seem to confirm that there is evidence of a higher difficulty in localization in the left hemisphere, namely for stereo dipole. However the right/median confusion value for Ambisonics, which is similar to the one of left/median confusion could mean a generalized (but not dominant) trend in azimuth localization errors, without distinction left/right. Another result that can be useful in order to deny a bias in the left hemisphere is contained in figure 13, where we plot the number of exact estimations for positions in sector 1,2,3. It results that left sources tend to be better localized. We conclude that there is no evidence for an artificial bias in the left hemisphere, and that the strange result in the real left/high localization is due to a problem of limited number (10) of the subjects that listened to that particular stimulus and a known difficulty to localize non azimuthal sources.

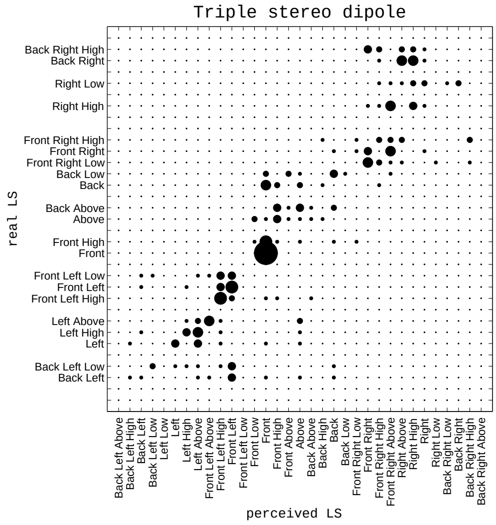


Figure 11 Localization for Triple Stereo Dipole

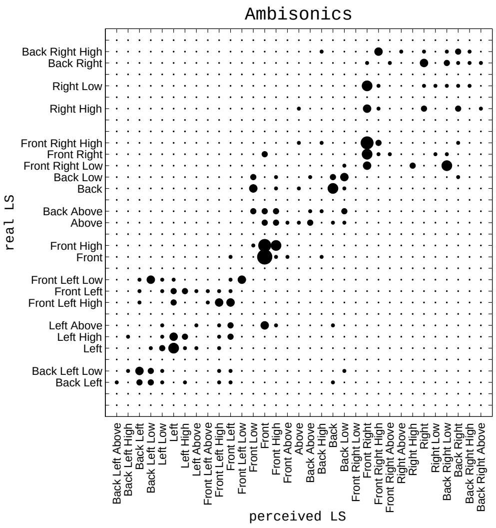


Figure 12 Localization for Ambisonics

In figure 14 we plotted the up/down confusion and the down/up confusion for the two systems. In figure 15 we show the results of Back/Front and Front/Back confusion. In figure 16 we plot the number of correct judgments, on 10 total judgments, for each azimuth, and zero elevation, in figure 17 the same thing, but varying elevation (for all values of azimuth).

6.3. Discussion

From the reported results, the first observation is that precise localization is not very accurate, for both systems. The three main sectors are hopefully well distinguished (table 2), but from figure 13 we can observe that an average of only 21% of sources is well localized.

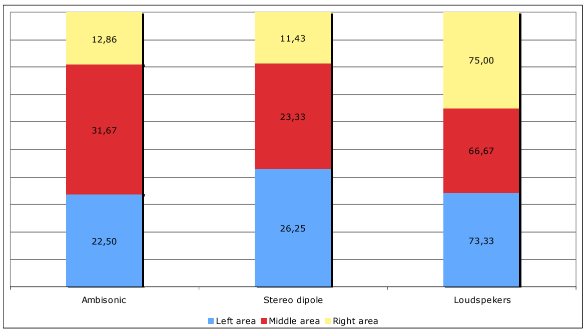


Figure 13 Rate of exactly localized positions for each sector and each system

Back/Front confusion (figure 15) (44% on average) is higher than Front/Back confusion (15% on average), meaning that sources not well localized in the rear sector are usually localized in the front sector more that the opposite. Up/down confusion (figure 14) is 6.5 % on average and 28% for down/up confusion, meaning that low sources are often localized in the upper hemisphere. Front sources azimuth is better identified than rear sources (figure 16) and zero elevation sources are better localized than sources with non-null elevation (figure 17).

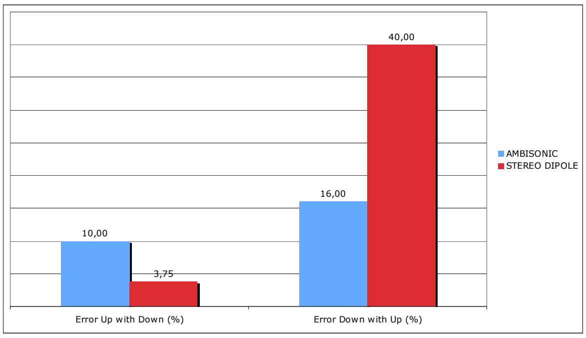


Figure 14 Upper and lower hemisphere confusion

The reason why we observe such bad performances is certainly due to the systems, but it is important to remark how back and non-azimuthal sources are in general more difficult to localize even in real life (a clue of this is given in the real loudspeaker localization figure). Moreover the results we report here do not allow for any tolerance: a source is considered as 'exactly localized' only if it has been correctly placed on the position grid: subjects were asked to force their answer indicating one position on the grid, wihtit any possibility of interpolation between the possible grid points. Allowing for some tolerance would surely result in a better performances report.

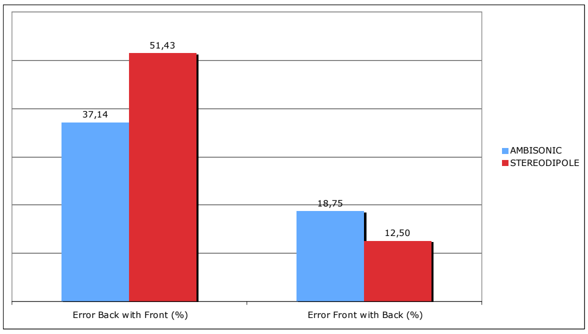


Figure 15 Front and back hemisphere confusion

Comparing the two systems, we can observe that low sources are badly localized with stereo dipole, which tends to a higher elevation perception (40% of down/up confusion against the 16% of Ambisonics). This can be due to the presence of the top triple stereo dipole. 51% and 37% are the back/front confusion rate with stereo dipole and Ambisonics, which results in better performances. Up/down confusion and front/back confusion are lower for the two systems, with lightly better performances for stereo dipole, which present better localization of frontal sources. Stereo dipole results in better elevation perception, except for the lower hemisphere, where its performances drop dramatically.

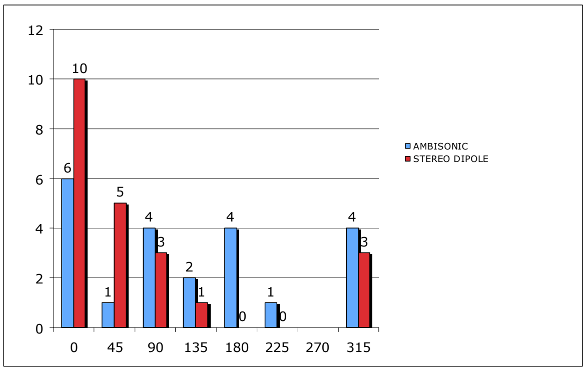


Figure 16 Number of exact localizations in function of azimuth (elevation 0)

From these data, it seems quite difficult to answer whether if stereo dipole or Ambisonics is better for localization. One reasonable remark is that the three stereo dipoles should be tuned on the basis of perceptual cues, more than considering the three pairs as equivalent.

Comparing our results with [8], we observe how Ambisonic results are much better in our system, while the opposite can be observed for Transaural system. This difference can be due to the implementation of the systems but most of all to the fact that in our tests subjects were not constrained to localize position on the azimuthal plane. However, restricting our analysis on the azimuthal plane (which anyway is not completely equivalent to the constrained test), result in stereo dipole predominance, especially in the front plane.

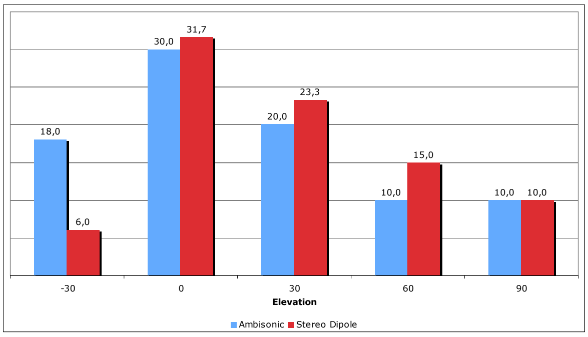


Figure 17 Rate of exact localization in function of elevation (all azimuths)

7. CONCLUSIONS, REMARKS AND FUTURE WORK

A comparison between a full 3D Ambisonic first order system and a triple stereo dipole has been done on the basis of localization tests on 20 subjects. 3D Localization is not very accurate for both systems, namely for not-azimuthal sources and back sources. even if no global statistical prove of the superior performances of one on the other have been found. However superiority of one system on the other (and vice-versa) has been found for particular positions. An ANOVA analysis of the results in order to extract significance parameters has not been done for the limited dimensions of our test results corpus.

The remarks of the subjects under test are quite important to improve this type of research. During this session we collected opinions, suggestions, comments on difficulties that will be helpful for future works. Here below a short list of the main notes:

- Most of the subjects pointed out the difficult in discriminating the elevation. The reason should be found in the habit of listening at stereo, and therefore planar, sources.
- Some subjects lamented the impossibility of a stop during the playback of the sequence, because they would re-listen at not well-recognized sources.
- Someone remarked that changing the orientation of the head, he should change his evaluation on the arrival direction. This could be caused by strange reflections inside the room, or simply to the limits of the stereo dipole.

In the future, we plan to investigate if a second order Ambisonic system and a customized triple stereo dipole will result in better performances.

8. REFERENCES

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

[2] A. Capra, M. Binelli, D. Marmiroli, P. Martignon, A. Farina, "Correlation Between Subjective Descriptors And Objective Parameters Of Theatres And Auditoria Acoustics Simulated With Binaural Sound Systems", ICSV13, Vienna, Austria, July 2-6, 2006

[3] O Kirkeby, PA. Nelson, H Hamada, "The stereo dipole: A virtual source imaging system using two closely spaced loudspeakers", Journal of the Audio Engineering Society, 1998

[4] Michael A. Gerzon. "Multidirectional Sound Reproduction", U.S. Patent 3,997,725, December 1976.

[5] Micheal A. Gerzon. "General Metatheory of Auditory Localization". In Preprints form the 92nd Auditory Engineering Society Convention, Vienna, number 3306, March 1992, AES-EL 3366.

[6] J. Daniel, "Représentation de Champs Acoustiques, Application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia", Ph.D. thesis, Universite Paris VI, 2000.

[7] A. Farina, E. Ugolotti, " Subjective comparison between Stereo Dipole and 3D Ambisonics surround systems for automotive applications", 16th AES Conference, Rovaniemi (Finland), 10-12 April 1999

[8] Catherine Guastavino, Véronique Larcher, Guillaume Catusseau and Patrick Boussard, "Spatial audio quality evaluation: comparing transaural, ambisonics and stereo", Proceedings of the 13th International Conference on Auditory Display, Montréal, Canada, June 26-29, 2007

[9] Eric Benjamin, Richard Lee, and Aaron Heller. "Localization in Horizontal-Only Ambisonic Systems", 121st AES Convention, San Francisco, CA, USA.

[10] http://www.kokkinizita.net/linuxaudio/

[11] Takeuchi, T. and Nelson, P.A. "Optimal source distribution for virtual acoustic imaging, ISVR Technical Report 288 ", Southampton, University of Southampton, 2000,

[12] D. G. Malham, "Higher order ambisonic systems for the spatialization of sound", in Proc. Int. Computer Music Conf., Beijing, China, 1999, pp. 484–487.

[13] O. Kirkeby, P. A. Nelson, "Digital Filter Designfor Inversion Problems in Sound Reproduction", J.AudioEng.Soc.,Vol.47,No.7/8,1999July/Augus

[14] S. Fontana, S. Campanini, a. Farina, "New method for auralizing the results of room acoustics simulation", ICA 2007, Madrid.

[15] C. Varani, E. Armelloni, A. Farina "Implementation of a double StereoDipole system on a DSP board - Experimental validation and subjective evaluation inside a car cockpit", 115th AES Convention, New York, 10-13 October 2003

[16] 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, pre-print n. 6485

[17] Takashi Takeuchi, Philip A Nelson, and Martin Teschl, "Elevated Control Transducers for virtual acoustic imaging" , 112th Convention 2002 May 10–13 Munich, Germany

[18] http://www.gerzonic.net

[19] http://www.plogue.com

[20] http://www.voxengo.com

[21] http://recherche.ircam.fr/equipes/salles/listen/

[22] http://ardour.org

[23] P.D Hatziantoniou, J.N Mourjopoulos "Errors in Real-Time Room Acoustics Dereverberation" J. Audio Eng. Soc, Vol. 52, No. 9, pp. 883 -899, 2004.

[24] H. Moller et all, "Binaural Technique: Do We Need Individual Recordings?", J. Audio Eng. Soc.,Vol. 44, No.6, 1996 June