ELSEVIER

applied acoustics

Applied Acoustics 62 (2001) 889–916

Acoustic quality of theatres: correlations between experimental measures and subjective evaluations

Angelo Farina *

Dipartimento Ingegneria Industriale, University of Parma, Via delle Scienze, 181/A I, 43100 Parma, Italy

Received 4 January 2000; received in revised form 5 September 2000; accepted 15 September 2000

Abstract

This paper deals with the psychoacoustic correlations between physical parameters and subjective aspects of perception of music in theatres. Many acoustic parameters were experimentally measured in eight Italian theatres and halls, including the Milanese Teatro Alla Scala, the Teatro Comunale in Bologna and the Teatro Comunale in Florence. A questionnaire was distributed, gathering almost 200 responses from well-known musicians (such as Riccardo Chailly, Severino Gazzelloni, Claudio Scimone and Uto Ughi). The results were compared and statistically analysed. Some interesting linear correlations were found between the physical acoustic parameters and the subjective evaluations. Finally, a short questionnaire, suitable for further psychoacoustic analysis was obtained. © 2001 Elsevier Science Ltd. All rights reserved.

1. Introduction

The aim of this research is to find those acoustical parameters, which are strongly related to the subjective judgement of the "acoustic comfort" in opera houses employed for symphonic music (as is common in Italy).

The knowledge of the relationships between objective parameters and subjective perception forms the basis for any intervention on existing theatres and for the design of new ones [1]. On the other hand, nowadays, we have dozens of proposed objective parameters, and their relevance and mutual interdependence is not clearly established [2,3].

E-mail address: farina@unipr.it

PII: S0003-682X(00)00082-7

This work does not pretend to give the final answer which would be a mathematical framework capable of predicting the complete subjective response from a reduced number of objective parameter values. Instead, it is aimed at a preliminary selection of the subjective questions, on one side, and of the objective parameters, on the other side, so that further studies can be restricted to a lower number of possible interrelationships. The collected information, however, is not sufficient yet to derive the definitive relationships with reasonable confidence.

The problem was dealt with by following these fundamental steps:

- carrying out experimental measurements of the *impulse response* (stage–audience) in many points of a number of well-known Italian theatres;
- setting up a questionnaire containing a number of items based on compared analysis techniques. A 6-steps numerical scale whose extremes were opposite subjective terms was used;
- gathering up almost 200 fully compiled questionnaires produced by highly qualified listeners (musicians, musicologists and music critics);
- carrying out a statistical analysis of objective and subjective data, by using multiple linear regression techniques.

The research was aimed mainly at assessing the acoustical quality of the musical perception from the listener's point of view: no explicit consideration was given to other important aspects, such as support for singers on stage, contact between stage and orchestra pit, balance between stage and orchestra, and the capability of the conductor or the musicians to listen each other.

This means that the theatres, although most of them are typical opera houses, were treated as if they were pure concert halls. It must be noted that in Italy, concert halls are quite unusual, and it is common here to use opera houses for symphonic music.

From the results obtained, it is possible to hypothesise the future realisation of a simple mathematical model which would make it possible to predict the subjective quality of a theatre from a reduced set of objective parameters. Such a subjective assessment will be very important in order to predict the effects of architectural interventions on theatres, so that the required collaboration between artists on one hand, and acoustic designers on the other, can be established.

Another relevant result was the explanation of some semantic errors, which usually make it difficult for musicians and acousticians to communicate: from the subjective responses to the questionnaires it was clear that some words are better than others for transferring information between the two worlds, and thus it was possible to end with a new simplified framework to be employed in further studies, made of a new set of questions and a reduced set of objective parameters.

2. Measurement procedure

The experimental measurements were conducted with instrumentation and data analysis methods compliant with the ISO3382/1997 standard [4].

The procedure followed during the phase of data acquisition and subsequent computation can be summarised in the following points:

- Binaural impulse responses were recorded in 18 seats of each room. The measurements were taken in empty conditions, but the seats were heavily upholstered in all the theatres considered here. The measurements were carried out using an impulsive source (pistol shot with spherical diffuser), binaural microphones (Sony DRW70-C) and a DAT recorder (Sony TCD-D3), according to ISO 3382/1997.
- The sound source was always placed at the centre of the proscenium, just under the fire curtain. The curtains were open and the stage/orchestra pit arranged for an opera performance.
- The number and placement of measuring points was chosen according to the guidelines for measurements in opera houses recently proposed by CIARM [5]: this means 9 points in the stalls and 9 in the balconies (or boxes).
- the DAT waveforms were digitally transferred to hard disk WAV files by a proper digital interface installed in a PC; the files were processed with the AURORA shareware software [6] for the subsequent computation of the objective parameters, following the ISO3382/1997 standard.
- Reverberation Times (EDT, T15 and T30) were computed in octave bands and linear (unweighted) band, by linear regression of the steady decay reconstructed through Schröder's backward integration of the impulse response; a noise reduction algorithm was applied to extend the range of the decay.
- The values of the reverberation times T15 were also checked by processing the DAT recordings with a real-time 1/3 octave analyzer (B&K2133). This ensured that the Aurora software performed correctly, within the tolerances of the ISO3382 standard.
- Some physical descriptors (clarity values C_5, C_{50}, C_{80}, reverb-to-direct ratio R/D, centre time t_s, rapid speech transmission index RASTI, strength G) were computed from the energetic impulse response (squared pressure), averaged among the two ear microphones, in the unweighted (linear) band.
- Evaluation of the sound-level distribution, recomputed from the strength values with a hypothetical sound power level of 100 dB (wide band). This means that the strength values can be recomputed back by subtracting 69 dB from the SPL values.
- Calculation of the wide-band normalised inter-aural cross-correlation ($IACC_E$) between the right and left impulse responses, using a special software implemented on a PC. This software, actually under development, will be made publicly available soon as part of the Aurora package [6].

The definitions of the above parameters are briefly reported in Appendix A.

For the Sala Europa, the Teatro alla Scala and the Teatro Comunale of Bologna, the measurement of the impulse responses was also performed with the MLSSA apparatus: this makes use of a pseudo-random MLS stimulus, emitted through an omnidirectional loudspeaker. In this case the binaural microphones were directly connected to the MLSSA board.

It was verified that the sound spectrum emitted by the omnidirectional loudspeaker is similar to that emitted by the pistol-shot: therefore the only significant difference between the two measuring techniques is the background noise, which is rejected efficiently with the MLSSA technique, but not with the pistol shots [7]. However, it must be noted that in all the theatres the background noise was very low (below 40 dBA), and thus its effect is substantially negligible, even with the pistol shots. Furthermore, the pistol shots are free from typical artifacts which affects MLS measurements, due to slew-rate limitation of the analogue electronics and to non-linear distortion inside transducers [8]. The new exponential sine sweep measurement method [9] was not available at the time of the measurements.

The pistol employed for firing blank ammunition was equipped with a spherical holed diffuser screwed on the exhaust hole of the pistol. This, connected with the use of special ammunition with calibrated powder charge (Fiocchi), made it possible to obtain a constant power level (within ±1 dB) and a reasonably uniform directivity pattern (much better than a dodechaedron loudspeaker).

A selection of the impulse responses measured in these 8 theatres is available for download on the Internet [10].

2.1. Experimental results

Measurements for acoustic qualification have been carried out in the following theatres:

– Teatro Alla Scala in Milan;
– Sala Poggio Imperiale in Florence;
– Teatro Comunale in Bologna;
– Teatro La Pergola in Florence;
– Sala Ridotto del Comunale in Florence;
– Teatro Comunale in Florence;
– Sala Europa in Bologna;
– Teatro Verdi in Florence;

Table 1 summarises the most important data of these theatres.

For each of these, 12 wide-band acoustic parameters were measured, which were:

- Reverberation time (R.T.) (s)
- Early decay time (E.D.T.) (s)
- Centre time (t_s) (ms)
- Initial Time Delay Gap (I.T.D.G.) (ms)
- Clarity (C80) (dB)
- Early/late ratio (C5) (dB)
- Sound pressure level (SPL) (dB)
- Rapid speech transmission index (RASTI)
- Equivalent reflection amplitude (Aeq)
- Inter aural cross correlation ($IACC_E$)
- Clarity (C50) (dB)
- Reverberant/direct ratio (R/D)(dB)

For each room analysed, a single number was found for each objective parameter by averaging first the left and right channels (to obtain monoaural descriptors) and then the different measurement points. Table 2 and Fig. 1 show the average values of the 12 measured parameters for each of the 8 theatres.

Table 1
Main characteristics for the 8 theatres

No.	Name	No. of seats	Height (m)	Volume (m^3)	Main floor surface (m^2)	Stage surface (m^2)	Construction year
1	SCALA	2015	18	12500	700	400	1778
2	PERGOLA	430	14	7500	520	260	1657
3	VERDI	1538	16	10950	660	400	1854
4	COMUNALE FI	2126	17	14000	820	600	1860
5	COM. RID. FI	500–600	10	2900	300	100–150	1966
6	POGGIO IMP.	100–120	7	1800	250	80–100	1600?
7	SALA EUROPA	1480	8–10	12700	1400	250	1980
8	COMUNALE BO	1250 c.	17	11000	650	400	1764

Obviously, the acoustic parameters are not the same throughout the auditoria, especially for non-Sabinian halls; this means that some acoustic parameters (e.g. ITDG) are variable between the middle of an auditorium and the seats close to the rear wall. This variation was analysed applying the analysis of variance (ANOVA) technique to the experimental measurements: however, the goal of this paper is not to study the spatial variation of objective parameters inside a room, thus these data are not presented here. On the other hand, for those parameters/rooms which scored a high standard deviation, it is presumed that it would be difficult to establish high correlation between the spatially-averaged value of the parameter and the subjective responses.

The reverberation time T15 was also measured in 1/3 octave bands, as reported in Fig. 2.

3. Subjective evaluation

3.1. Preliminary test for setting up the questionnaire

In setting up the questionnaire, Wilkens' work was taken as the starting point [11]; infact, his questionnaire is the most complete, although it contains a large number of descriptive words.

Table 2
Average values of the measured parameters for each theatre

Theatre no.	Name	T30 (s)	EDT (s)	t_s (ms)	ITDG (ms)	C80 (dB)	$IACC_E$	SPL (dB)	Aeq	Rasti	C50 (dB)	C5 (dB)	R/D (dB)
1	SCALA	1.209	1.323	107.46	21.52	−0.11	0.22	68.1	1.00	0.50	−3.76	−13.7	−0.09
2	PERGOLA	1.251	1.259	97.89	16.26	0.80	0.17	75.2	3.05	0.52	−2.38	−13.6	9.22
3	VERDI	1.570	1.422	102.89	12.78	1.85	0.33	67.8	1.94	0.52	−0.64	−7.7	5.65
4	COMUNALE FI	1.624	1.242	69.23	21.99	4.17	0.36	66.3	1.25	0.60	1.66	−4.1	1.72
5	COM. RID. FI	1.080	1.010	62.96	15.09	4.01	0.28	72.2	1.56	0.61	1.41	−6.3	3.73
6	POGGIO IMP.	2.316	2.092	148.64	15.58	−1.21	0.21	76.1	2.83	0.39	−3.54	−10.3	8.95
7	SALA EUROPA	1.269	1.207	56.06	19.27	6.00	0.51	65.2	1.28	0.63	3.44	−6.5	1.71
8	COMUNALE BO	1.685	1.712	112.60	14.25	0.38	0.48	74.8	2.28	0.47	−2.45	−7.2	6.54

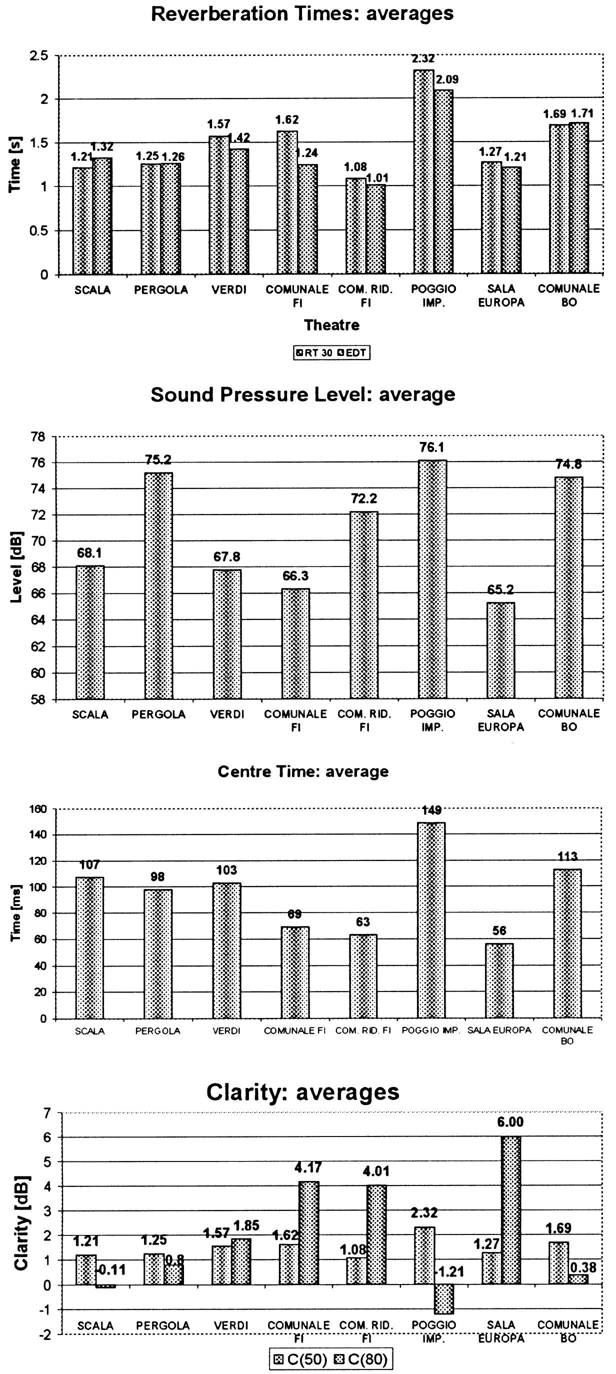


Fig. 1. Average acoustic parameters measured in 8 Italian theatres, as listed in Table 1 (*continued on next page*).

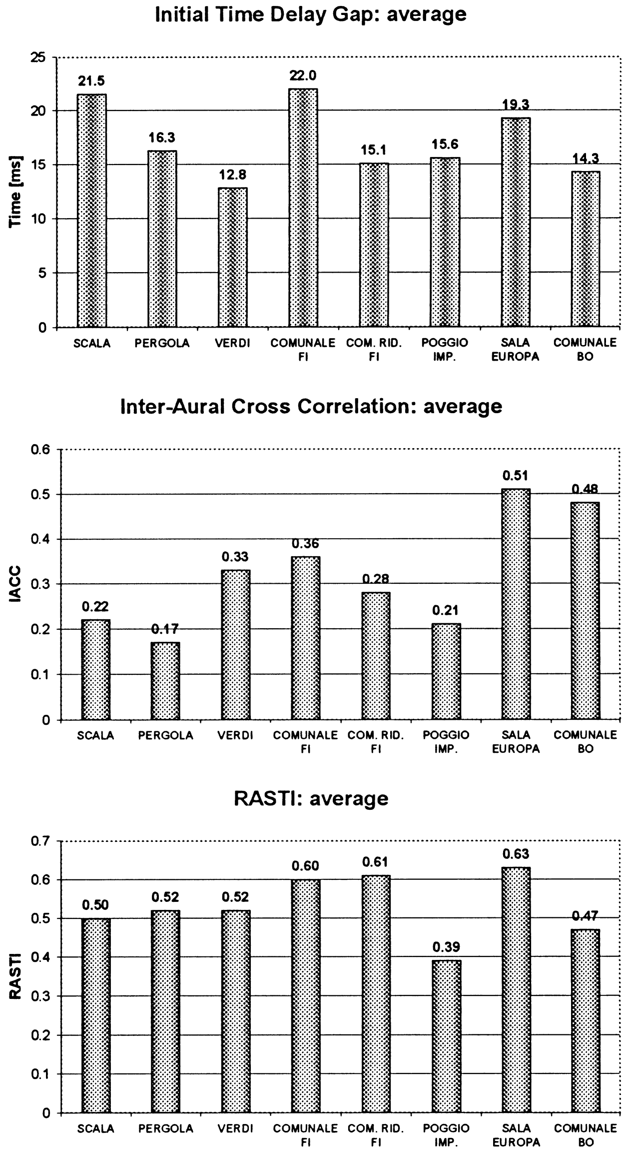


Fig. 1. (*continued*)

Wilkens' pairs, in Italian, German and English, were randomly sorted between the left and the right side, in order to avoid the risk of associating one side of the questionnaire to a particular quality judgement. It was decided to use a multi-lingual questionnaire in order to avoid misunderstandings on the part of musicians from different countries. Obviously this required an accurate choice for the translation of the German terms into English and Italian.

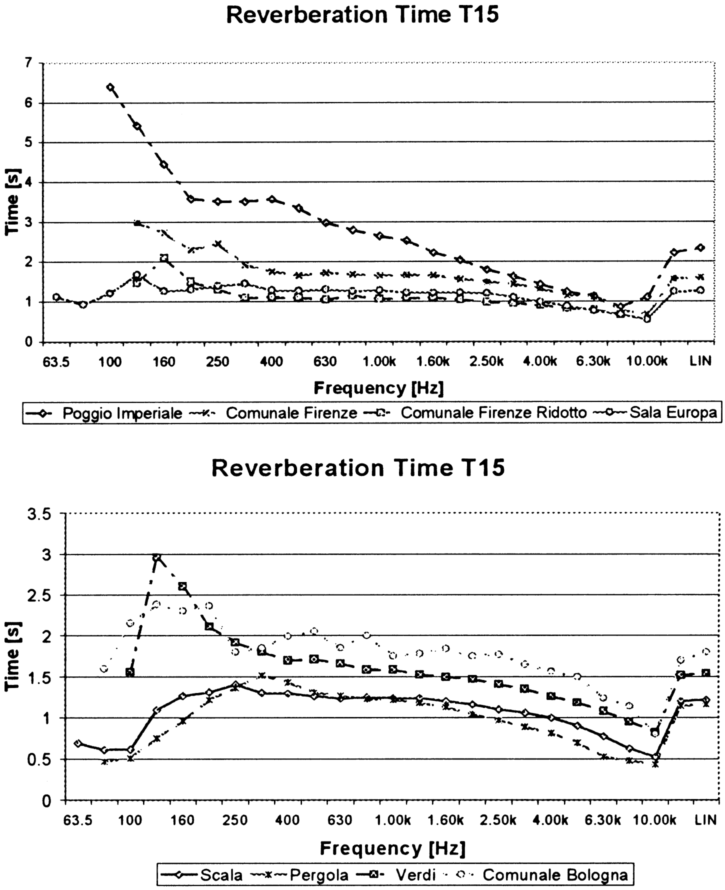


Fig. 2. Reverberation time T15 measured in 8 Italian theatres, as listed in Table 1.

The questionnaire was distributed among the members of the "Orchestra Teatro alla Scala di Milano". The aim was to judge the readiness of those musicians to answer the different questions and, if necessary, modify or improve them.

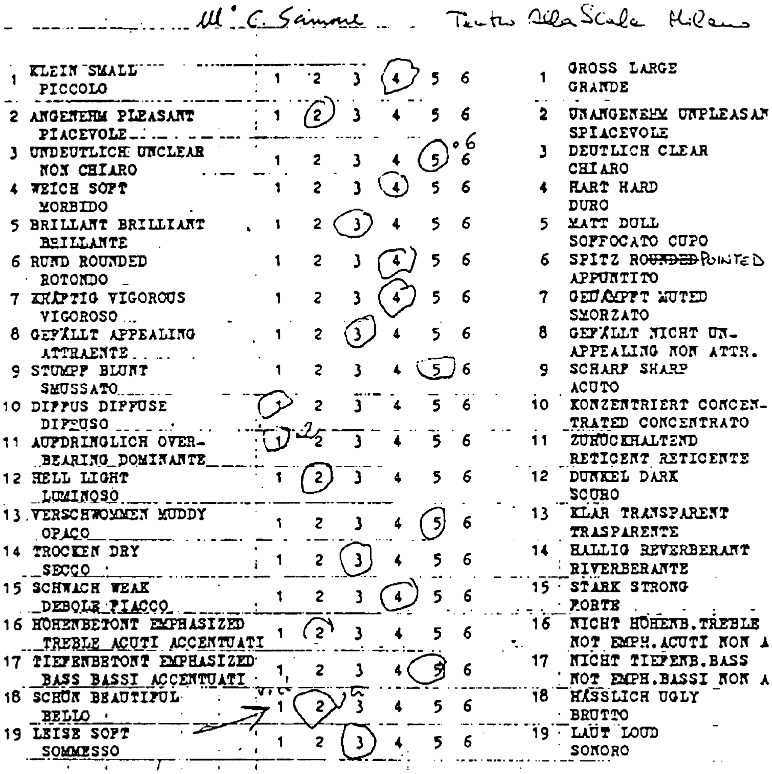
Fig. 3 reports the preliminary questionnaire compiled by Mr. Scimone.

Finally, when the completed questionnaires had been returned, the opinions were analysed. It came out that certain questions had generally been answered in the same way showing a very high correlation, while others showed "random" responses.

Due to the need to simplify the raw data, along with the need expressed by the musicians to be able to complete the questionnaire quickly, the text was re-written, reducing the number of terms.

3.2. Final version

In the next version of the questionnaire, some pairs of terms were deleted, e.g. "small–large"; other pairs were re-translated with a meaning closer to the language


M° C. Scimone — Teatro alla Scala Milano

1	KLEIN SMALL PICCOLO	1	2	3	4	5	6	1	GROSS LARGE GRANDE
2	ANGENEHM PLEASANT PIACEVOLE	1	2	3	4	5	6	2	UNANGENEHM UNPLEASANT SPIACEVOLE
3	UNDEUTLICH UNCLEAR NON CHIARO	1	2	3	4	5	6	3	DEUTLICH CLEAR CHIARO
4	WEICH SOFT MORBIDO	1	2	3	4	5	6	4	HART HARD DURO
5	BRILLANT BRILLIANT BRILLANTE	1	2	3	4	5	6	5	MATT DULL SOFFOCATO CUPO
6	RUND ROUNDED ROTONDO	1	2	3	4	5	6	6	SPITZ ~~ROUNDED~~ POINTED APPUNTITO
7	KRÄFTIG VIGOROUS VIGOROSO	1	2	3	4	5	6	7	GEDÄMPFT MUTED SMORZATO
8	GEFÄLLT APPEALING ATTRAENTE	1	2	3	4	5	6	8	GEFÄLLT NICHT UN-APPEALING NON ATTR.
9	STUMPF BLUNT SMUSSATO	1	2	3	4	5	6	9	SCHARF SHARP ACUTO
10	DIFFUS DIFFUSE DIFFUSO	1	2	3	4	5	6	10	KONZENTRIERT CONCEN-TRATED CONCENTRATO
11	AUFDRINGLICH OVER-BEARING DOMINANTE	1	2	3	4	5	6	11	ZURÜCKHALTEND RETICENT RETICENTE
12	HELL LIGHT LUMINOSO	1	2	3	4	5	6	12	DUNKEL DARK SCURO
13	VERSCHWOMMEN MUDDY OPACO	1	2	3	4	5	6	13	KLAR TRANSPARENT TRASPARENTE
14	TROCKEN DRY SECCO	1	2	3	4	5	6	14	HALLIG REVERBERANT RIVERBERANTE
15	SCHWACH WEAK DEBOLE FIACCO	1	2	3	4	5	6	15	STARK STRONG FORTE
16	HÖHENBETONT EMPHASIZED TREBLE ACUTI ACCENTUATI	1	2	3	4	5	6	16	NICHT HÖHENB. TREBLE NOT EMPH. ACUTI NON A
17	TIEFENBETONT EMPHASIZED BASS BASSI ACCENTUATI	1	2	3	4	5	6	17	NICHT TIEFENB. BASS NOT EMPH. BASSI NON A
18	SCHÖN BEAUTIFUL BELLO	1	2	3	4	5	6	18	HÄSSLICH UGLY BRUTTO
19	LEISE SOFT SOMMESSO	1	2	3	4	5	6	19	LAUT LOUD SONORO

Fig. 3. The first version of the questionnaire, completed by C. Scimone.

of music; some redundant pairs were reduced to one single pair, since they appeared very close in meaning: an example is the pair "rund–spitz" and "stumpf–scharf", in English "rounded–pointed" and "blunt–sharp", which, in Italian, became the single pair "rotondo-spigoloso".

Some pairs of words with almost identical meaning were still included, but were well separated and inverted, such as "vigorous–attenuated" and "weak–strong", in order to check the reliability of the judgements expressed.

Finally, in the margin of the questionnaire, under the space reserved for the overall judgement, there was a summary of the details regarding the person completing the questionnaire.

The second questionnaire was then distributed among Italian artists and musicians of international fame (conductors, soloists and orchestra members) asking them to express their opinions only in the case of personal performing experiences in the theatres proposed, taking into account their overall impression gathered from having been performers in these theatres, but with reference to what they feel was the audience response. They were asked to judge only the effect on the audience produced by the theatre, considering a general musical orchestral piece: this is an uncommon approach, as usually performers were asked to give their judgement from the performer's point of view, whilst the audience response was investigated with questionnaires distributed to the audience itself. This uncommon approach is justified by the fact that usually in Italy only the opinion of the performers, and particularly of conductors, is taken into account regarding the evaluation of the acoustics of a theatre. Furthermore, it is quite common that high-level performers and conductors do some listening tests in the theatre in various audience positions, for better tailoring their execution to the performance space. Only very recently, a paper from Hidaka and Beranek [12] was published based exactly on the same approach (employing, in that case, only conductors who had to judge the listening experience of the audience in opera houses).

Of course, each subject of the panel filled in a questionnaire only for those theatres where he happened to play several times, taking also into account other experiences as a listener in the same theatre.

A total of 192 questionnaires were analysed, coming from 47 subjects.

Fig. 4 reports the final questionnaire compiled by Mr. Alberti.

4. Statistical analysis

4.1. Data analysis

The questionnaires referred to the eight theatres analysed in part.3. Therefore, the data obtained belonged to two matrixes measuring 192×14 (questionnaires-questions) and 8×12 (theatres-parameters) respectively.

The "subjective data" matrix (192×14) was reduced to (8×14), averaging the values of each hall: the reduced matrix was used only for plotting the relation between each objective/subjective pair. The complete matrix was instead used for calculating the linear regression between objective parameters and subjective evaluations.

It should be noted that some interviewees suggested intermediate answers to the values proposed, and in such cases the decimal values proposed were taken as valid data.

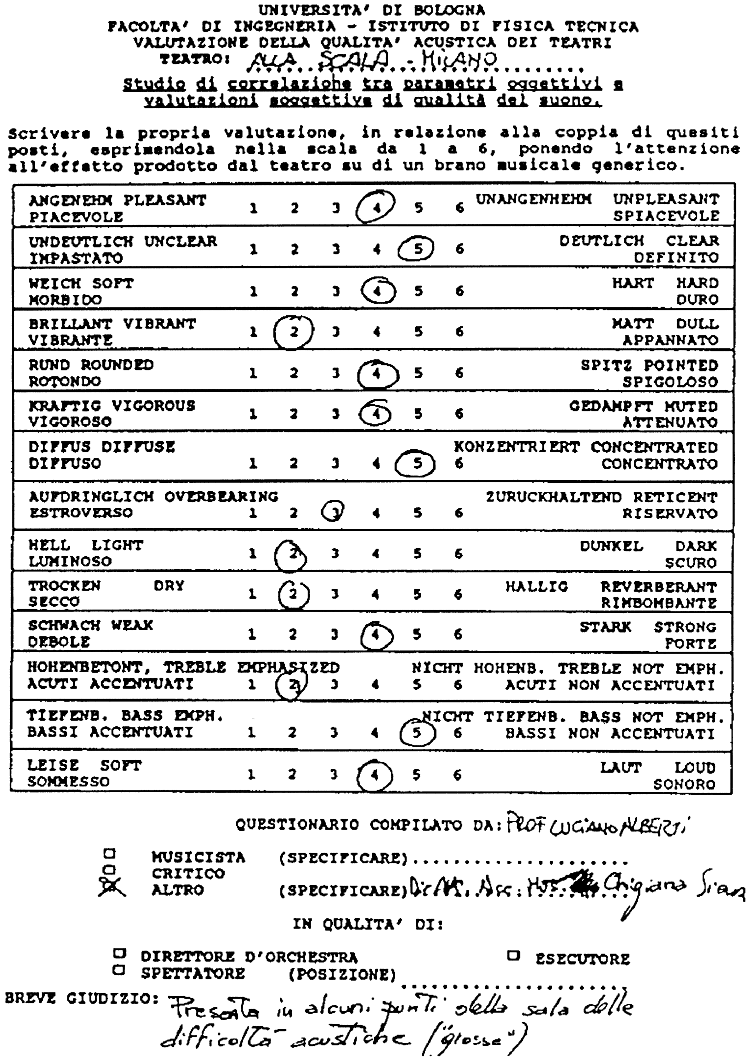
4.2. Statistical calculation procedure

For each of the 168 crossings of objective/subjective data (12×14), the linear regression and correlation coefficients, with the 192 data points available, were calculated. In fact, according to Ando, each subjective parameter is a function of only one objective descriptor: thus, the proposed method should be able to identify the most correlated objective parameter for each subjective response.

The linear regression line can be expressed in the following formula:

$$Y_{SUBJ} = A \cdot X_{OBJ} + B \tag{1}$$

in which X_{OBJ} = objective parameter; Y_{SUBJ} = subjective parameter.



UNIVERSITA' DI BOLOGNA
FACOLTA' DI INGEGNERIA - ISTITUTO DI FISICA TECNICA
VALUTAZIONE DELLA QUALITA' ACUSTICA DEI TEATRI
TEATRO: ALLA SCALA - MILANO........
Studio di correlazione tra parametri oggettivi e valutazioni soggettive di qualità del suono.

Scrivere la propria valutazione, in relazione alla coppia di quesiti posti, esprimendola nella scala da 1 a 6, ponendo l'attenzione all'effetto prodotto dal teatro su di un brano musicale generico.

ANGENEHM PLEASANT PIACEVOLE	1	2	3	(4)	5	6	UNANGENHEHM UNPLEASANT SPIACEVOLE
UNDEUTLICH UNCLEAR IMPASTATO	1	2	3	4	(5)	6	DEUTLICH CLEAR DEFINITO
WEICH SOFT MORBIDO	1	2	3	(4)	5	6	HART HARD DURO
BRILLANT VIBRANT VIBRANTE	1	(2)	3	4	5	6	MATT DULL APPANNATO
RUND ROUNDED ROTONDO	1	2	3	(4)	5	6	SPITZ POINTED SPIGOLOSO
KRAFTIG VIGOROUS VIGOROSO	1	2	3	(4)	5	6	GEDAMPFT MUTED ATTENUATO
DIFFUS DIFFUSE DIFFUSO	1	2	3	4	(5)	6	KONZENTRIERT CONCENTRATED CONCENTRATO
AUFDRINGLICH OVERBEARING ESTROVERSO	1	2	(3)	4	5	6	ZURUCKHALTEND RETICENT RISERVATO
HELL LIGHT LUMINOSO	1	(2)	3	4	5	6	DUNKEL DARK SCURO
TROCKEN DRY SECCO	1	(2)	3	4	5	6	HALLIG REVERBERANT RIMBOMBANTE
SCHWACH WEAK DEBOLE	1	2	3	(4)	5	6	STARK STRONG FORTE
HOHENBETONT, TREBLE EMPHASIZED ACUTI ACCENTUATI	1	(2)	3	4	5	6	NICHT HOHENB. TREBLE NOT EMPH. ACUTI NON ACCENTUATI
TIEFENB. BASS EMPH. BASSI ACCENTUATI	1	2	3	4	(5)	6	NICHT TIEFENB. BASS NOT EMPH. BASSI NON ACCENTUATI
LEISE SOFT SOMMESSO	1	2	3	(4)	5	6	LAUT LOUD SONORO

QUESTIONARIO COMPILATO DA: PROF LUCIANO ALBERTI

☐ MUSICISTA (SPECIFICARE).....................
☐ CRITICO
☒ ALTRO (SPECIFICARE) Dir. Art. Acc. Mus. Chigiana Siena

IN QUALITA' DI:

☐ DIRETTORE D'ORCHESTRA ☐ ESECUTORE
☐ SPETTATORE (POSIZIONE) ....

BREVE GIUDIZIO: Presenta in alcuni punti della sala delle difficoltà acustiche ("glosse")

Fig. 4. The final version of the questionnaire, completed by L. Alberti.

Coefficient A is calculated with the method of the least squares, according to well-known algorithms. The coefficient of correlation "r" represents the degree of approximation obtained in the calculation of the regression. It should be noted that it is always:

$$r \in [-1, 1] \quad (2)$$

"r" represents a positive or negative number, which shows how well the regression line approximates the input data. The majority of the 168 correlation coefficients "r" obtained are very small, ($\ll$ 0.1), indicating that the linear relationship (1) does not fit with the subjective results. Of course, it could well be that other relationships, for example the fractional-order parabolic functions suggested by Ando [1], are better suited for describing subjective responses which potentially express maximum preference for an intermediate value of the objective parameter, and reduced preference for lower or higher values.

Only a few correlation coefficients are greater than 0.30, which was assumed as the threshold for discarding insignificant linear correlations.

The standard deviation was also calculated, in order to have a figure regarding the dispersion of the data analysed.

4.3. Results

At the end of the calculation procedure, three 12×14 data matrixes were constituted, for the coefficients "A" and "B" of the linear regression line, and for the linear regression coefficients "*r*" respectively.

Table 3 reports the values of "*r*", with absolute values greater or equal to 0.30 marked bold.

5. Discussion of results

Just a quick look at Table 3 demonstrates that some objective parameters and some questions happen to correlate more easily than others. Particularly, it emerges that R/D is better than A_{eq} and C5 (which have substantially the same physical meaning), and that the "best" objective parameter is the sound pressure level SPL. On the other hand, $IACC_E$ revealed very poor correlation with all the subjective terms of the questionnaire.

Looking at the rows of the table, questions 1, 7, 10 and 11 were the ones with better linear correlation with objective parameters: question 1 is the preference

Table 3
Linear regression coefficient "*r*"

Parameters	1 T30	2 EDT	3 t_s	4 ITDG	5 C80	6 $IACC_E$	7 SPL	8 Aeq	9 Rasti	10 C50	11 C5	12 R/D
Pleasant – Unpleasant	−0.10	−0.23	**−0.31**	0.28	**0.34**	0.03	**−0.41**	**−0.37**	**0.33**	**0.31**	0.15	**−0.35**
Unclear – Clear	−0.16	−0.07	−0.03	0.04	−0.05	−0.06	0.10	0.03	0.01	−0.08	−0.12	0.01
Soft – Hard	−0.15	−0.20	−0.25	**0.35**	0.22	0.03	**−0.35**	**−0.40**	0.26	0.18	0.06	**−0.40**
Vibrant – Dull	−0.14	−0.24	**−0.30**	0.17	**0.32**	0.02	**−0.34**	−0.29	**0.31**	0.29	0.12	−0.27
Rounded – Pointed	−0.08	−0.14	−0.20	**0.34**	0.20	0.00	**−0.31**	**−0.35**	0.21	0.16	0.06	**−0.35**
Vigorous – Muted	−0.20	−0.22	−0.20	0.22	0.18	−0.03	**−0.31**	−0.29	0.22	0.14	−0.02	−0.29
Diffuse – Concentrated	**−0.30**	**−0.36**	**−0.30**	**0.31**	0.24	−0.14	**−0.40**	**−0.38**	**0.34**	0.18	−0.04	**−0.38**
Overbearing – Reticent	−0.25	−0.25	−0.20	0.17	0.14	−0.09	−0.20	−0.19	0.22	0.10	−0.07	−0.20
Light – Dark	−0.11	−0.17	−0.24	0.08	0.26	0.07	−0.24	−0.22	0.25	0.25	0.13	−0.20
Dry – Reverberant	**0.35**	**0.36**	0.29	**−0.34**	−0.19	0.17	−0.27	0.29	**−0.31**	−0.12	0.11	**0.31**
Weak – Strong	0.22	**0.31**	**0.30**	**−0.30**	−0.28	0.12	**0.34**	0.27	**−0.33**	−0.23	0.03	0.28
Treble enhanced – Tr. not en.	−0.07	−0.05	−0.08	−0.06	0.06	0.07	0.02	−0.02	0.06	0.06	0.06	−0.01
Bass enhanced – Bass not en.	−0.18	−0.09	−0.02	0.16	−0.04	−0.05	−0.10	−0.17	0.03	−0.08	−0.15	−0.19
Soft – Loud	0.26	**0.33**	0.27	**−0.31**	−0.23	0.20	**0.34**	0.28	**−0.31**	−0.18	0.10	**0.30**

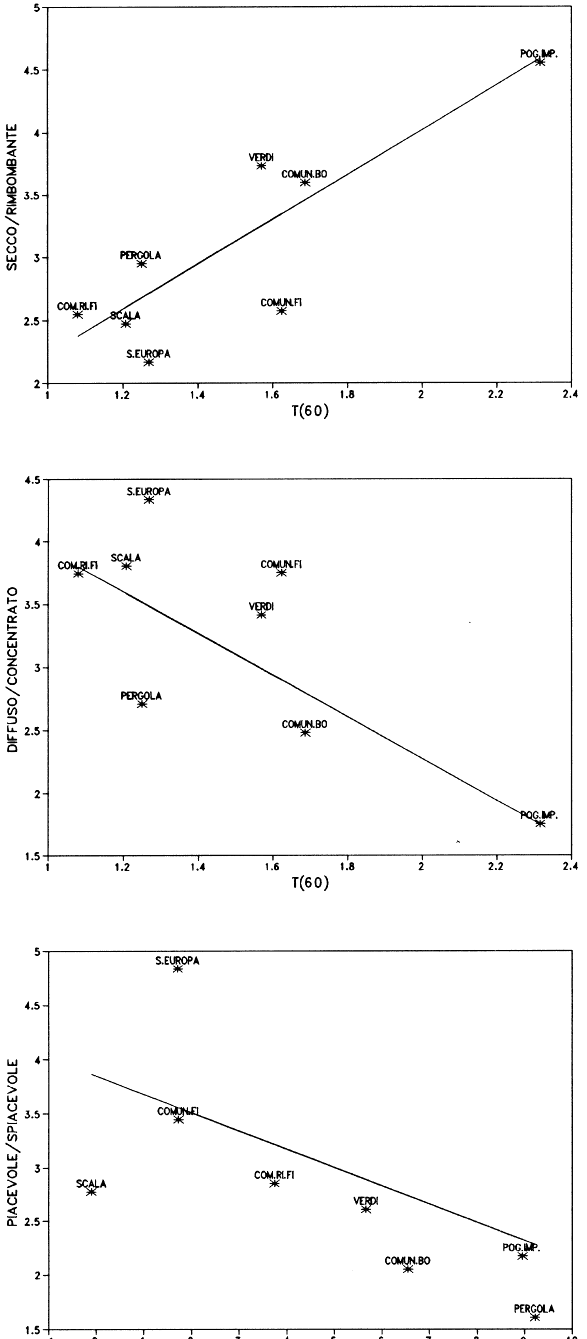


Fig. 5. Six correlation plots, chosen between pairs with $r > 0.30$ (*continued on next page*).

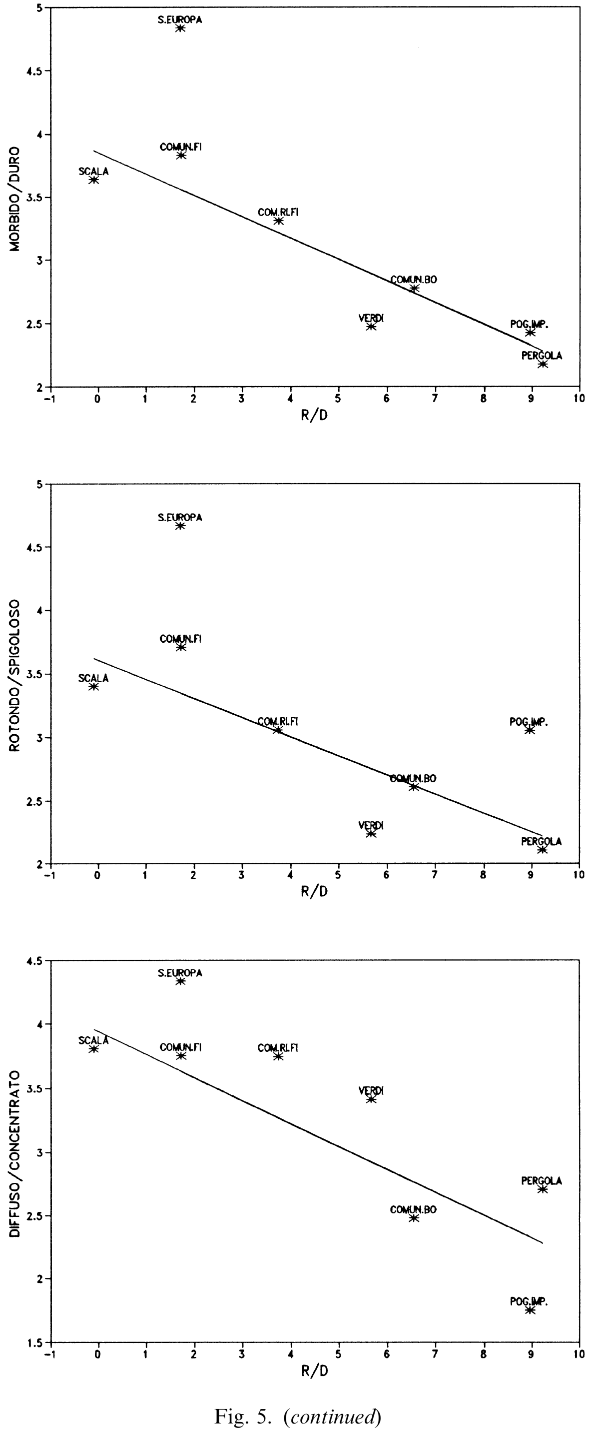


Fig. 5. (*continued*)

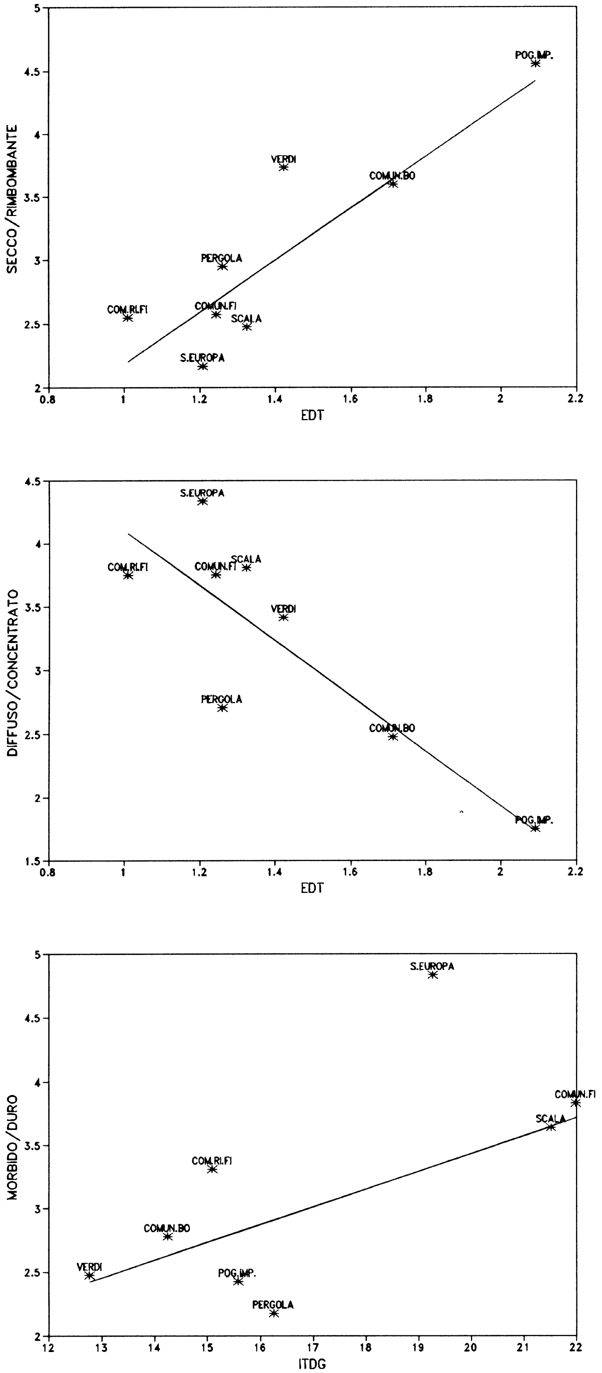


Fig. 6. Six correlation plots chosen between pairs with $r > 0.30$ (*continued on next page*).

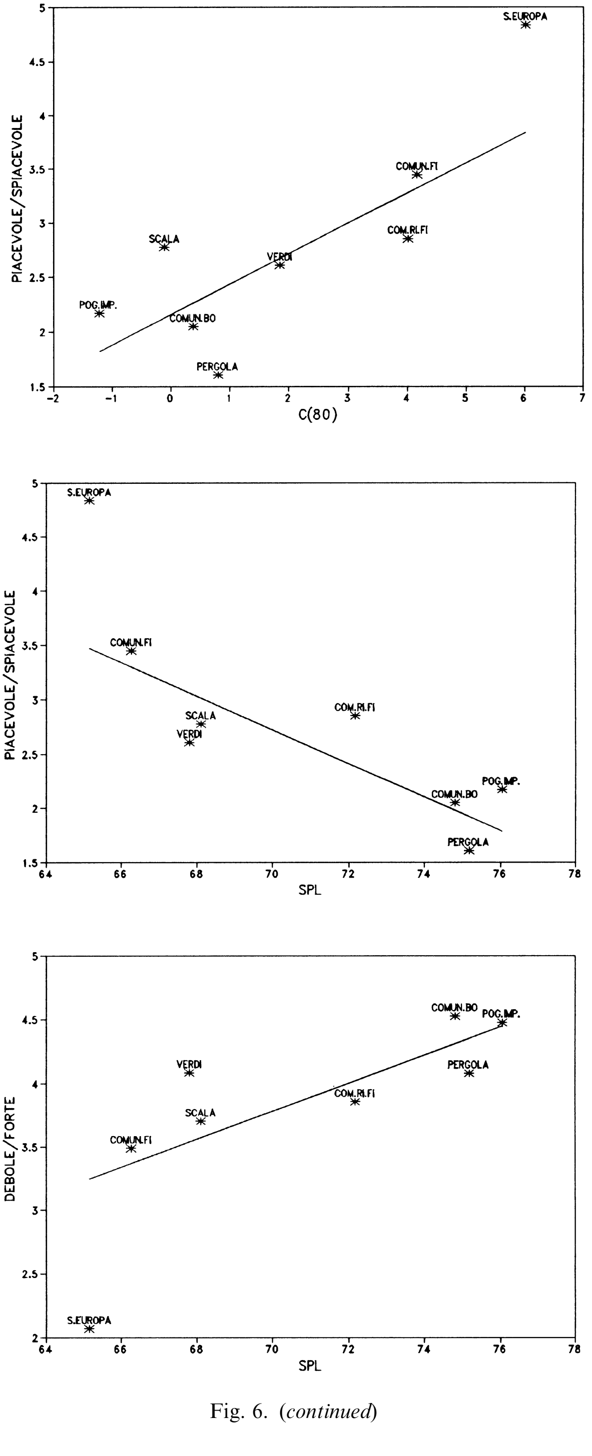


Fig. 6. (*continued*)

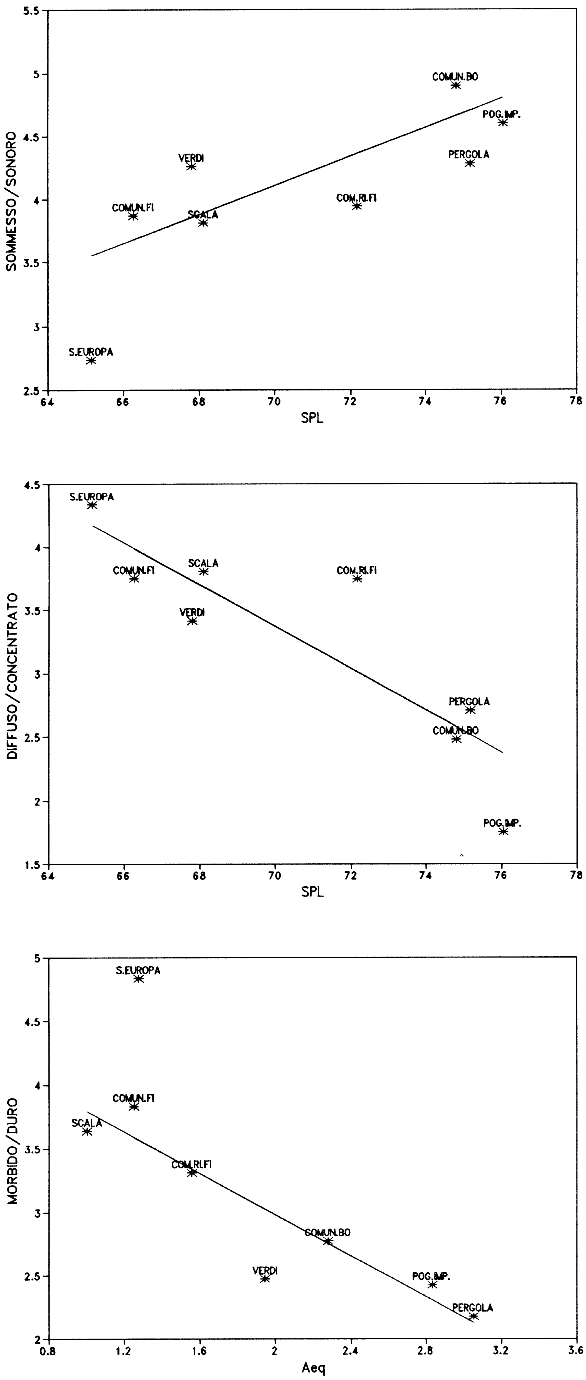
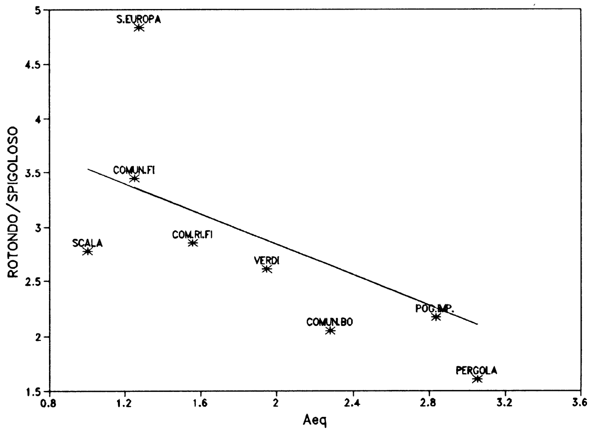
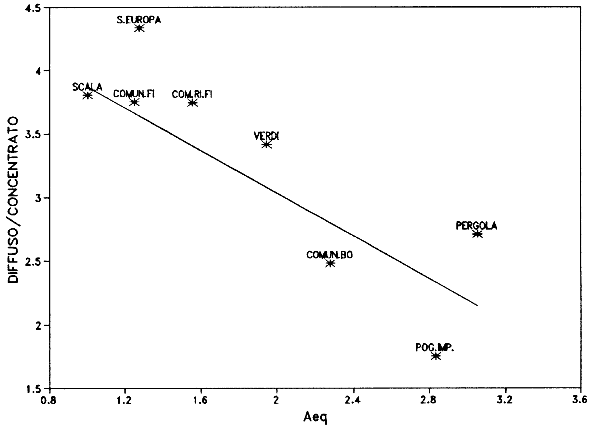


Fig. 7. Six correlation plots chosen between pairs with $r > 0.30$ (*continued on next page*).





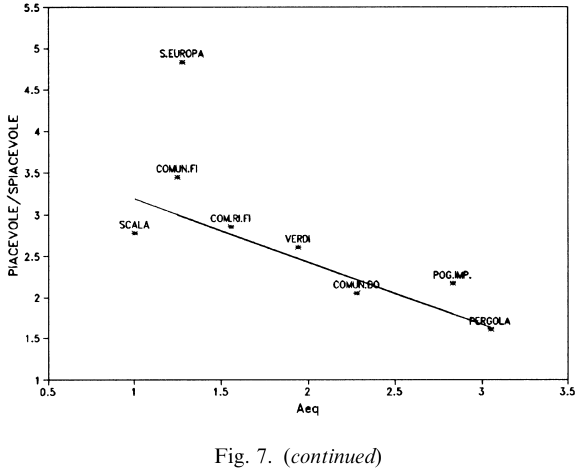


Fig. 7. (*continued*)

(pleasant–unpleasant), which of course is the most important subjective attribute, and which seems to be affected by many objective parameters; question 7 is the pair (diffuse–concentrated), which was expected to give high correlation with $IACC_E$, and instead correlated well with many other parameters, giving an awful −0.14 with $IACC_E$. Question 10 is the pair (dry-reverberant), and of course it is strongly correlated with objective parameters affected by the room reverberation (T30, EDT, R/D); question 11 is the pair (weak-strong), which of course correlates well with SPL, but also with objective parameters related to the reverberance, such as EDT, ITDG and t_s.

For the most significant subjective/objective crossings, a plot was made of the resulting regression line, superposed to the average values of the responses to the questionnaires for each of the 8 theatres analysed; these plots are visible in 5, 6 and 7.

It should be stressed how the average value of the subjective parameters of each of the theatres does not represent directly the data which have been used for the regression calculation: in fact as the number of questionnaires is different for each theatre, these data points have different weight. For example, the questionnaires relating to the Teatro Comunale di Bologna and to the Teatro alla Scala di Milano represent about 54% of the total number of questionnaires, while the data available for the Sala Europa di Bologna and the Sala Concerti di Poggio Imperiale in Florence were only 7% of the total number (14 out of 192). This means that the data points related to the latter two have very little weight compared to the first two.

The acoustic parameters having a higher correlation with the subjective pairs are:

- EDT with *diffuse–concentrated* and with *dry–reverberant*: $r = -0.36$, $r = +0.36$;
- T30 with*diffuse–concentrated* and with *dry–reverberant*: $r = -0.30$, $r = +0.35$;
- ITDG with s*oft–hard*: $r = +0.35$;
- C80 with *pleasant–unpleasant*: $r = +0.34$;
- SPL with *pleasant–unpleasant*, with *soft–hard*, with *diffuse–concentrated,* with *weak–strong* and with *soft–loud*: $r = -0.41$, $r = -0.35$, $r = -0.40$, $r = +0.34$, $r = +0.34$;
- A_{eq} with *pleasant–unpleasant*, with *soft–hard*, with *rounded–pointed,* with *diffuse–concentrated*: $r = -0.37$, $r = -0.40$, $r = -0.35$, $r = -0.38$;
- R/D with *pleasant–unpleasant*, with *soft–hard*, with *rounded–pointed,* with *diffuse–concentrated*: $r = -0.35$, $r = -0.40$, $r = -0.35$, $r = -0.38$.

In conclusion, even for these selected cases, the absolute values of the linear correlation coefficients are quite low. Of course, more advanced data analysis techniques could be used for extracting more information from the high statistical noise captured in the subjective data: the author did not attempt this because in other studies he demonstrated how, starting with subjective responses collected in much more controlled conditions, and making use of trained and selected subjects, it is possible to obtain much better results [13].

In practice, the above data are not reliable enough for deriving the mathematical relationship between objective parameters and subjective responses, although they can serve for improving the formulation of a new questionnaire and for discarding some of the objective parameters.

6. Conclusions

6.1. Proposal for a new questionnaire

Taking into account the presence of some semantic errors (particularly regarding the translation of terms in different languages), and removing the evident redundancy, it was possible to write a reduced questionnaire, more suitable to statistic elaboration and easier to complete.

The new questionnaire has been optimised in order to carry out further research in the field of acoustics and architecture, to measure a number of physical descriptors, and to evaluate their correlation with the most important subjective parameters.

Furthermore, it will be possible to check also non-linear correlations between objective parameters and subjective responses, as this was revealed to be the weaker point in the present work.

The final goal is to create a numerical formulation which enables the researchers to calculate the subjective preference (and other subjective criteria) in each seat of a given concert hall, based on the knowledge of some objective parameters. The new questionnaire is as follows:

Question no.	Left attribute	Right attribute
1	Piacevole (pleasant)	Spiacevole (unpleasant)
2	Rotundo (round)	Spigaloso (sharp)
3	Morbido (soft)	Duro (hard)
4	Diffuso (diffuse)	Localizzabile (localisable)
5	Distaccato (detached)	Avvolgente (enveloping)
6	Secco (dry)	Rimbombante (reverberant)
7	Acuti accentuati (treble boosted)	Acuti ridotti (treble reduced)
8	Bassi accentuati (bass boosted)	Bassi ridotti (bass reduced)
9	Sommesso (quiet)	Sonoro (loud)

Comparing the new questionnaire with the previous one, it can be noted how the redundant pairs have been removed, leaving for each group the terms more similar to musical language.

Pairs 4 and 5 have been added, closer to binaural descriptors such as $IACC_E$, which had been left out of Wilkens' questionnaire.

Another point was the number of values between the two opposite terms. As many subjects attempted to place the mark exactly in the middle of the scale, it was advisable to change the number of slots from an even one (6) to an odd one (5 or 7), so that there was one exactly in the middle.

Finally, the set of objective parameters has also been slightly reduced. As the result was that some parameters were substantially the same, in each of these groups the "best" parameter was chosen. The value of SPL was substituted with Strength, which is now more widely employed for room acoustic qualification. Furthermore, due to the failure of $IACC_E$, LE and LF were also introduced (also described in Appendix A, although not employed in the experimental part of this work), with the goal of better describing the spatial envelopment of the listener in the room. Finally, RASTI was substituted with STI, as nowadays the measurement of the latter poses no more problems than the first.

This is the reduced list of objective parameters:

Parameters	1	2	3	4	5	6	7	8	9
	Strength (dB)	EDT (s)	Ts (ms)	C80 (dB)	R/D (dB)	IACC	LE	LF	STI

6.2. Future experiments

The continuation of this research will profit by the recent developments in the field of virtual acoustics and real-time auralization [14,15]. In particular, as the impulse responses of the 8 theatres are available in standard WAV format, these can be used as numerical filters, applied by convolution to anechoic music samples [16]. In this way, it will be possible to prepare a huge set of sound samples, obtained with different original music pieces, convolved with the impulse responses of different positions of the 8 theatres.

The subjects will have to fill-in the questionnaire while listening to each music piece. As the impulse responses are binaural, the easiest way to obtain this is through headphone reproduction. On the other hand, the reproduction over a pair of loudspeakers in the stereo-dipole configuration (and through proper cross-talk cancelling filters) can give more natural listening conditions [17–19], although this requires a treated listening room.

Finally, for automating the process of playing the WAV files and simultaneously collecting the questionnaires, special software was developed, as shown in Fig. 8: this allows for interactive compilation of the responses, with the capability of instantaneously switching to any other sound sample in the collection for easy comparison. This method, together with a preliminary selection of the subjects, was capable of producing subjective responses very consistent and highly correlatable with objective measurement, as reported in [13] with reference to the subjective evaluation of the sound quality of car audio systems.

The availability of the responses directly in electronic format opens the possibility of more advanced analytical approaches. In particular, the frequency dependence of the most correlated parameters shall be evaluated, taking into account the values of the objective parameters in at least 8 octave bands (63 Hz–8 kHz).

Furthermore, as the convolution is made with the impulse response in a particular point in the theatre, it will be possible to analyse also the spatial dependence of both the subjective responses and the objective parameters.

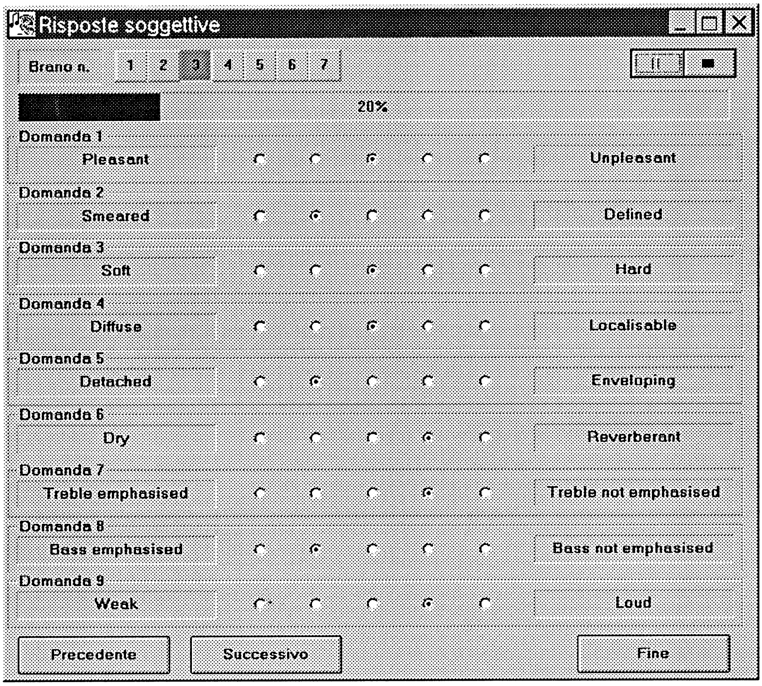


Fig. 8. Special software for listening tests with automatic collection of questionaires.

It is planned to analyse these data in a more complex way than in this study, making use of the factor analysis and other advanced statistical techniques.

Acknowledgements

The author wishes to express his gratitude to Lamberto Tronchin for his substantial support to the research. For their co-operation during the subjective studies, the author also wishes to thank: the permanent orchestra of the "Teatro alla Scala di Milano" (23 musicians); the permanent orchestra of the "Teatro Comunale di Bologna" (56 musicians); orchestra of the "Accademia S.Cecilia di Roma" (12 musicians). Among those who gave their contribution by compiling the questionnaires, the author would like to mention Mr. Alberti, Mr. Benini, Mr. Chailly, Mr. Gazzelloni, Mr. Scimone, Mr. Ughi and Professor Raul Farolfi.

Appendix A. Definition of the acoustical parameters

In the following the physical acoustical parameters employed in the work are described in detail. In general these definitions are in agreement with the ISO 3382

(1997) standard [4] and with the guidelines proposed by CIARM for measurements in opera houses [5].

A1. T_{15}, T_{30}

Reverberation time calculated from the decay range between −5 and −20 dB (T_{15}) and between −5 and −35 dB (T_{30}) on the integrated Schroeder curve, in seconds. Schroeder [20] found that the reverberant decay can be described by a backward integration of the impulse response:

$$\langle p^2(t)\rangle = \int_t^{\infty} h^2(\tau)\mathrm{d}\tau \tag{a1}$$

where:

$\langle p^2(t)\rangle$ = average of a infinite number of decay
$h^2(\tau)$ = impulse response.

Eq. (a1) can be written as:

$$\langle p^2(t)\rangle = N\int_0^{\infty} h^2(\tau)\mathrm{d}\tau\text{-}\int_0^{t} \mathrm{h}^2(\tau)\mathrm{d}\tau \tag{a2}$$

A2. Early decay time (EDT)

Since Jordan 21 demonstrated that the subjective perception of reverberation in correlated more strongly with the initial decay of the reverberant tail, he suggested to calculated the reverberation time from the decay range between 0 and −10 dB on the integrated Schroeder curve, in seconds.

A3.Centre time $\mathrm{t_s}$

It was defined by Kürer [22], as *Schwerpunktzeit*, in the following equation:

$$t_s = \frac{\int_0^{\infty} \tau h^2(\tau)\mathrm{d}\tau}{\int_0^{\infty} h^2(\tau)\mathrm{d}\tau} \tag{a3}$$

It is the first-order momentum of the squared pressure impulse response, expressed in milliseconds.

A4. Initial time delay gap (ITDG)

Defined by Beranek [23], it is the delay of the first reflection from the direct wave, expressed in milliseconds. It is usually calculated directly from the impulse response.

A5. Inter-aural cross correlation ($IACC_E$)

As suggested by Ando [1], it is the normalized correlation coefficient between the first 80 ms of the pressure impulse responses measured at the two ears of the binaural microphone.

From the definition of the cross-correlation function, given by:

$$p(\tau) = \frac{\int_0^{80ms} h_2(t) \cdot h_R(t+\tau)dt}{\int_0^{80ms} P_2^2(t)dt \cdot \int_0^{80ms} P_R^2(t) \cdot dt} \tag{a4}$$

the $IACC_E$ is defined as the maximum value $p(\tau)$, that is:

$$IACC_E = |p(\tau)|_{max} \qquad \text{where } -ms < \tau < 1\,\text{ms} \tag{a5}$$

A6. Strength (G)

It is the difference between the measured sound pressure level, and that produced by the same omnidirectional source in a free field, at 10-m distance from its center, and is expressed in decibels. It was defined in ISO 3382/1997, and expressed in the following equation

$$G = 10\log\frac{\int_0^{\infty} h^2(\tau)d\tau}{\int_T^{\infty} h_{10}^2(\tau)d\tau} \tag{a6}$$

In this work the Strength is not employed directly, but this parameter is easily recomputed from the values of SPL, as the impulse response were scaled correspondingly to a source's power level of 100 dB, and thus producing a free-field level, at 10m, of 69 dB. This means that the Strength values can be obtained back subtracting 69 dB from the SPL values reported in this work.

A7. Clarity C_{80}, C_{50} *and* C_5

Reichardt, Abdel, Alim and Schmidt [24] defined such an acoustic parameter in order to relate the "transparence" of the music to an energetic parameter. It is defined by the equation

$$C = 10\log\frac{\int_0^{T} h^2(\tau)d\tau}{\int_T^{\infty} h^2(\tau)d\tau} \tag{a7}$$

When the clarity is related to the musical perception, the time interval T is limited to 80 ms, whereas if the clarity is related to speech, the time interval T is set to 50 ms. Here it was even computed with a shorter integration limit of 5 ms.

A8. Reverberant-to-direct ratio R/D

The Reverberant-to-direct ratio is the level difference between the reverberant sound field and the direct sound. Thus the definition is:

$$R/D = 10\log\left[\frac{\int_{\varepsilon}^{\infty} h^2(\tau)\cdot d\tau}{\int_0^{\varepsilon} h^2(\tau)\cdot d\tau}\right] \tag{a8}$$

where ε is the duration of the direct sound: if this is equal to 5 ms, then *R/D* is simply equal to minus C_5; instead, in general, as the duration of the direct sound can vary for each impulse response, the values of *R/D* exhibit some minor deviations from the values of C_5 (with inverted sign).

A9. Equivalent reflection amplitude A_{eq}

It is a sort of reverberant-to-direct ratio, but expressed as signal amplitude ratio instead of employing the dB scale. Thus the definition is:

$$A_{eq} = \sqrt{\frac{\int_{\epsilon}^{\infty} h^2(\tau)\cdot d\tau}{\int_0^{\epsilon} h^2(\tau)\cdot d\tau}} \tag{a9}$$

where ϵ is the duration of the direct sound (see above).

A10. Speech transmission index (STI) and rapid speech transmission index (RASTI)

These parameters were defined by Houtgast and Steeneken [25,26]. They are computed from the values of the modulation transfer function (MTF). This quantity is defined by the ratio of the received modulation amplitude to the original modulation amplitude, making use of an excitation signal obtained by an octave-band filtered pink noise, having an energy envelope (squared amplitude) slowly modulated at frequency *F*.

At the receiving microphone, the sampled signal presents an energy envelope with reduced modulation depth. The ratio between the output modulation amplitude and the input modulation amplitude is the MTF,

$$m(F) = \frac{m_{out}}{m_{in}} \tag{a10}$$

The value of *m(F)* can be computed for each modulation frequency *F* in the one-third-octave increments from 0.63 to 12.5 Hz, covering the range of the human voice modulations. Furthermore, the octave-band filtered carrier signal can be produced for any octave band from 125 Hz to 8 kHz. Thus, a complete set of MTF values for seven octaves is obtained, which contains globally $7\cdot 14 = 98$ values.

The 14 MTF values coming from each carrier's octave band are first employed for computing an equivalent Signal-to-Noise ratio, clipped in the interval −15 to +15 dB. After this, a frequency-dependent STI value is obtained. Finally, the wide-band STI value can be obtained by a weighted average of the 7 single-frequency STI values. Different weights can be employed for male or female speakers.

Although the *STI* is an exhaustive acoustic parameter, describing the intelligibility of concert halls, Houtgast and Steeneken [27] found out that it was not necessary to obtain a complete set of 98 MTF values to quantify the intelligibility of an enclosed space. Hence they defined a shorter parameter, the Rapid Speech Transmission Index (RASTI), limiting the analysis to a restricted number of data. The procedure is very similar but the carrier's frequencies are referred only to the octave bands of 500 and 2000 Hz, and the modulation frequencies are only 4 at 500 Hz and 5 at 2000 Hz.

The values of $m(F)$ can be computed directly from the impulse response of a system, provided that it is a linear, passive, time-invariant system, as shown by Schroeder [28].

These hypotheses correspond to the assumption of a very little background noise, so that the modulation depth reduction is due only to the room's reflections, echoes and reverberation. This assumption is certainly met in this case, because the theatres were very silent, and no external noise could affect the listening conditions.

Since the effective signal-to-noise ratio is of no influence, there is no problem employing a sound source with a spectrum different from the normalized spectrum defined in the original works of Houtgast and Steeneken [25–27], or those defined in the new IEC 268-16-16 standard [29]. The only problem connected with the use of an impulsive source (a gun shot in this case) is the directivity of the source, which does not resemble the normalized directivity of the human voice since it is almost perfectly omnidirectional. Furthermore, the recording of the impulse response is not made with an omnidirectional microphone, but with a binaural one. These deviations are, to a certain degree, reciprocally compensating. Furthermore, the measurements under these conditions are probably more representative of the real listening conditions in a theatre, where the listeners are always facing the stage, while the speaker is moving around on it.

To compute each value of $m(F)$ from the impulse response $h(t)$, an octave-band filter is first applied to the impulse response, in order to select the carrier's frequency band f. Then $m(F)$ is obtained with the formula

$$m(F) = \frac{\int_0^{\infty} h_f^2(\tau)\cdot\exp(-j\cdot 2\cdot\pi\cdot F\cdot\tau)\cdot \mathrm{d}\tau}{\int_0^{\infty} h_f^2(\tau)\cdot \mathrm{d}\tau} \tag{a11}$$

A11. Lateral efficiency LE

This parameter was first introduced by Jordan [21], and is defined as the ratio between the late energy coming from lateral directions (measured with a figure-of-eight microphone aligned with the ears of the listener) and the total omnidirectional energy. Thus, the mathematical definition is:

$$LE = \frac{\int_{25\,\mathrm{ms}}^{80\,\mathrm{ms}} h_8^2(\tau)\cdot \mathrm{d}\tau}{\int_0^{80\,\mathrm{ms}} h_0^2(\tau)\cdot \mathrm{d}\tau} \tag{a12}$$

where h_8 denotes the impulse response measured with a figure-of-eight microphone, and h_0 is instead the normal impulse response measured with an omnidirectional microphone.

A12. Lateral fraction LF

The definition of *LF*, taken from ISO3382, is very similar to *LE*, but the ratio now includes also the early lateral reflections, which were explicitely excluded in *LE* by the use of the lower integration limit of 25 ms after the direct sound. Thus, the definition of *LF* is the following:

$$LF = \frac{\int_{5\,\mathrm{ms}}^{80\,\mathrm{ms}} h_8^2(\tau)\cdot \mathrm{d}\tau}{\int_{0}^{80\,\mathrm{ms}} h_0^2(\tau)\cdot \mathrm{d}\tau} \tag{a13}$$

This means that the value of *LF* is greater than *LE* (although always less than 1), particularly in rooms of limited size, where the first reflections on the lateral walls, occurring before the 25 ms limit, do not contribute to *LE* but are included in *LF*.

References

[1] Ando Y. Concert hall acoustic. Berlin: Springer Series in electrophysics, 1985.
[2] Muller, H.A. Room acoustical criteria and their meaning In: Proceeding of the International Conference Acoustics and Recovery of Spaces for Music, 1993, Ferrara, pp. 51–60
[3] Vorlaender M, Bietz H et al. Uncertainties of measurements in room acoustics. Acustica 1995;81:344–55.
[4] ISO 3382. Acoustics — measurement of the reverberation time of rooms with reference to other acoustical parameters. International Organisation for Standardisation, 1997
[5] CIARM (International Centre of Acoustics and Music Research). Guidelines for acoustical measurements inside historical opera houses: procedures and validation, Ferrara, Italy, November 1999. http://acustica.ing.unife.it/ciarm/news.htm.
[6] AURORA software home page – HTTP://www.ramsete.com/aurora.
[7] Farina A, Fausti P. Acoustic measurements in opera houses: comparison between different techniques and equipment. Journal of Sound and Vibration 2000;232(1):213–29.
[8] Vanderkooy J. Aspects of MLS measuring systems. JAES 1994;42(4):219–31.
[9] Farina A. Simultaneous measurement of impulse response and distortion with a swept-sine technique. 108th AES Convention, Paris 18–22 February 2000.
[10] Angelo Farina's collection of publicly available impulse responses — HTTP://pcangelo.eng.unipr.it/public/imp-resp.
[11] Wilkens H. Mehrdimensionale beschreibung subjektiver beurteilungen der akustic von konzertsälen. Acustica, 1977; 38 (10).
[12] Hidaka T, Beranek LL. Objective and subjective evaluations of twenty-three opera houses in Europe, Japan and the Americas. J. Acoust. Soc. Am. 2000 107 (1).
[13] Farina A, Ugolotti E. Subjective comparison of different car audio systems by the auralization technique. Pre-prints of the 103rd AES Convention, New York, pp. 26–29 September 1997.
[14] Farina A. Auralization software for the evaluation of a pyramid tracing code: results of subjective listening tests. ICA95 (International Conference on Acoustics), Trondheim, Norway 26–30 June 1995.
[15] Farina A, Righini F. Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering. Pre-prints of the 103rd AES Convention, New York, 26–29 September 1997.
[16] Farina A. An example of adding spatial impression to recorded music: signal convolution with binaural impulse responses. Proceedings of the International Conference Acoustics and Recovery of Spaces for Music, Ferrara 27–28 October 1993.
[17] Kirkeby O, Nelson PA, Hamada H. The "stereo dipole" — a virtual source imaging system using two closely spaced loudspeakers. JAES 1998;46(5):387–95.

[18] Kirkeby O, Nelson PA, Rubak P, Farina A. Design of cross-talk cancellation networks by using fast deconvolution. 106th AES Convention, Munich, 8–11 May 1999.
[19] Farina A, Ugolotti E. Subjective comparison between stereo dipole and 3D ambisonics surround systems for automotive applications. 16th AES Conference, Rovaniemi, Finland, 10–12 April 1999.
[20] Schroeder M. New method of measuring reverberation time. Journal of Acoustical Society of America 1965;37:409–12.
[21] Jordan VL. A group of objective acoustical criteria for concert halls. Applied Acoustics, 1981; 14.
[22] Kürer, R. Zur gewinnung von eizahlkriterien bei impulsmessungeg in der raumakustik. Acustica, 1969; 21.
[23] Beranek L. Music, acoustics and architecture. New York: John Wiley & Sons, 1962.
[24] Reichardt W, Abel Alim O, Schmidt W. Definition und Meβgrunglage eines objectiven Maβes zur Ermittlungder Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei musikdarbietung. Acustica 1975;32:126.
[25] Houtgast T, Steeneken HJM, Plomp R. Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics. Acustica 1980;16:60–72.
[26] Steeneken HJM, Houtgast T. A physical method for measuring speech-transmission quality. Journal of Acoustical Society of America 1980;61:318–26.
[27] Houtgast T, Steeneken HJM. A multi-language evaluation of the RASTI-method for estimating speech intelligibility in auditoria. Acustica 1984;54:185–99.
[28] Schroeder M. Modulation transfer functions: definition and measurement. Acustica 1981;49:179–82.
[29] IEC 268-16, Sound system equipment part 16, entitled: The objective rating of speech intelligibility by the Speech Transmission Index.